

**TAXONOMIC AND GENOMIC  
INVESTIGATION OF SOIL  
MICROORGANISMS ASSOCIATED  
WITH ACUTE OAK DECLINE**

**DANIEL WILLIAM MADDOCK BSc.**

A thesis submitted in partial fulfilment of the requirements of the  
University of the West of England, Bristol for the degree of Doctor of  
Philosophy

This research programme was carried out in collaboration with  
Forest Research and the Woodland Heritage

College of Health, Science and Society, The University of the West of  
England, Bristol

August 2023

## Author's declaration

This copy has been supplied on the understanding that it is copyright material and no quotation from the thesis may be published without proper acknowledgment.

## Abstract

Cases of the current episode of Acute Oak Decline (AOD) were first described in the UK in 2008 on native British oaks, but currently this disease is seen to affect most species of oak on a global scale. The disease has become a focus of forest research, due to the large number of abiotic and biotic components that contribute to AOD. Some of the most recent discoveries have shown the rhizosphere microbiome may play a significant role in the disease.

As such the work presented here utilised ten healthy and ten AOD symptomatic trees spread over the woodland and parkland areas of Hatchlands Park, Guilford, UK to look for functional differences within their rhizosphere. The first aim was to investigate if the rhizosphere soil could function as a reservoir of infection for the four bacteria currently associated with AOD, by using microbiological and rapid diagnostic molecular methods to detect their presence. The second aim was to isolate and classify potentially novel members of the Enterobacteriaceae isolated from the rhizosphere of both healthy and diseased oak. The final aim was to use 16S rRNA gene sequence analysis to identify differences in the microbiome community composition of the oak rhizosphere.

Using a modified *Enterobacteriaceae* enrichment method followed by undiluted spread plating on a range of media, HRM was used to test for AOD bacteria from both healthy and diseased oak rhizosphere samples. *Gibbsiella quercinecans* and *Rahnella victoriana* were identified in the rhizosphere microbiome of healthy and diseased oak using high resolution melt analysis, but *Brenneria goodwinii* and *Lonsdalea britannica* were not. However, *B. goodwinii* and *L. britannica* were identified in acorns from the same site, indicating that they may be endophytic members of the seed microbiome which is inherited between generations of oak.

Using the same enrichment method, 100's of isolates were collected from the same rhizosphere samples taken from Hatchlands and the use of polyphasic taxonomy allowed for the identification of two novel genera of bacteria and nine new species. These novel species appear to be potential plant pathogens and growth-promoting bacteria based on the genomic comparisons. This highlights the untapped potential that the oak rhizosphere microbiome offers in relation to bacterial taxonomy.

Finally, single gene community analysis via synthetic long read 16S rRNA gene sequencing was performed on DNA isolated from the Hatchlands rhizosphere samples. The analysis of sequencing data revealed important distinctions between samples, with the largest effect being seen between woodland and parkland samples. Parkland healthy and diseased trees presenting the second largest effect with different relative abundance and microbiome composition. However, this effect was not seen between single healthy and diseased trees with minimised sampling distances, though a difference between the exorhizosphere and endorhizosphere was recorded for both diseased and healthy trees. These differences indicate that the composition of the rhizosphere may be linked to health status of oak, and that the difference seen between the endosphere of diseased and healthy oak may be influenced by which bacteria are recruited from the rhizosphere.

## Acknowledgements

While I am immensely proud of everything that I have achieved in the course of completing my doctoral thesis, to say that I couldn't have achieved a fraction of it without the support of those around me would be an understatement. To my supervisor, friend, and mentor Dr Carrie Brady, I owe a debt of gratitude, for providing the best working environment I can imagine. Likewise, I would like to thank the other members of my supervisory team, Professor Dawn Arnold for teaching me so much and putting up with my terrible figure legends as well as many a fun evening at a conference and Dr Joel Allainguillaume for all his guidance.

An unfathomable amount of thanks is owed to all my friends at UWE, including Joshua Steven, Eva Perrin, Alex Copper, Savannah Britton, Angeliki Savvantoglou, Marina Kauffmanner, Rebecca O'Brien and Buffy Smith. I am very lucky to have taken this journey alongside you, thank you for all the support, respite and fun we've had! Likewise, I would like to thank all my oldest friends from Chandlers Ford, for always believing in me and looking after me on many fun weekends whilst still a poor student.

Acknowledgments must also be made to Dr Sandra Denman who helped secure funding, sampling permissions and has contributed her unparalleled knowledge in the field of forest health. Further thanks are due to Dr Bridgett Crampton for her training at Alice Holt and for allowing me to trial the LAMP method that ensured the validity of this work, Dr Dann Turner for assistance with the bioinformatics as well as Sally Simpson, Harry Kaluszniak, Susan Streeter, Lorna Hodgson and the extended AOD research group who made much of this work possible. I would like to acknowledge the financial, academic and technical support of the University of the West of England, Bristol, Alice Holt, Woodland Heritage and all their staff. This studentship was funded by Woodland Heritage (Charity no. 1041611) and the University of the West of England.

To my family I'm so grateful for the support, love and encouragement given to me during the course of this PhD. Though I'm not sure you ever understood what I was doing, I do appreciate you listening to me rattle on about it and trying to remember what was going on. The highest gratitude goes to my girlfriend Jenny, who not only supported, encouraged and put up with me, but did so while also completing her own PhD. This thesis is dedicated to you, you truly are an inspiration.

## Table of Contents

Author’s declaration .....	2
Abstract.....	3
Acknowledgements.....	5
Table of Contents.....	6
List of Figures .....	12
List of Tables .....	17
Abbreviations.....	21
Chapter 1. Introduction .....	24
1.1. Acute oak decline. ....	25
1.1.1. Description and occurrence. ....	25
1.1.2. Descriptive symptoms of AOD .....	26
1.2. Models of Decline diseases.....	28
1.3. Abiotic causes of AOD.....	31
1.4. Biotic causes - <i>Agrilus biguttatus</i> .....	32
1.5. Biotic factors - Bacterial components .....	34
1.6. Taxonomic classification of the microbiome .....	35
1.7. Understanding the role of the oak microbiome.....	37
1.8. Rapid identification methods for key members of the diseased microbiome. ....	39
1.9. In-depth analysis of key members of the AOD microbiome .....	42
1.10. Bulk Soil:.....	46
1.11. The rhizosphere: .....	48
1.11.1. Structure:.....	48
1.11.2. Colonization of the ectorrhizosphere and rhizoplane .....	49
1.11.3. Roles of rhizosphere colonisers in oak.....	50
1.11.4. Colonization of the endorhizosphere and endophytic spaces.....	51

1.12.	Plant pathogen interactions.....	52
1.13.	Aims and objectives: .....	54
Chapter 2.	Materials and methods .....	55
	Culturing methods: .....	56
2.1.1.	Sample collection.....	56
2.1.2.	Isolation of bacteria from bleeding lesions.....	56
2.1.3.	Culturing bacteria from mixed samples .....	56
2.1.4.	Isolation of bacteria in mixed samples.....	56
2.1.5.	Glycerol stocks .....	57
2.1.6.	Overnight batch cultures of bacterial strains.....	57
2.1.7.	Cultivation of bacteria from soil.....	57
2.1.8.	Enterobacteriaceae enrichment of bacteria from soil .....	58
2.1.9.	Growth curves.....	58
2.1.10.	Surface sterilisation of leaves and acorns .....	59
2.1.11.	Spiking of different microcosms with AOD bacteria .....	59
2.2.	Genotypic Methods: .....	60
2.2.1.	DNA extraction; colony in water .....	60
2.2.2.	DNA extraction; alkaline lysis.....	60
2.2.3.	Soil DNA extraction .....	60
2.2.4.	DNA extraction from pure cultures.....	60
2.2.5.	Root DNA extraction .....	61
2.2.6.	Polymerase chain reaction.....	61
2.2.7.	Agarose Gel electrophoresis .....	66
2.2.8.	ExoSap .....	66
2.2.9.	Sequencing .....	66
2.2.10.	NanoDrop DNA quantification .....	68

2.2.11.	Broad range Qubit DNA quantification .....	68
2.2.12.	Analysis of sequencing products .....	68
2.2.13.	High-Resolution Melt Analysis .....	69
2.2.14.	Spiking soil microcosms with AOD bacterial DNA .....	72
2.2.15.	Loop-mediated isothermal amplification (LAMP) .....	72
2.3.	Taxonomic Methods .....	73
2.3.1.	Phylogenetic analysis: .....	73
2.3.2.	DNA Fingerprinting: .....	73
2.3.3.	Genomic features .....	74
2.3.4.	Phylogenomic analysis .....	74
2.3.5.	Whole genome similarity analysis:.....	75
2.3.6.	Light microscopy .....	75
2.3.7.	Transmission electron microscopy.....	75
2.3.8.	Cell physiology .....	76
2.3.9.	Antibiotic resistance .....	76
2.3.10.	Phenotypic tests .....	76
2.3.11.	Fatty Acid Methyl Ester analysis .....	77
2.3.12.	Virulence feature .....	77
2.3.13.	PGP features .....	77
2.4.	Bioinformatics Methods .....	78
2.4.1.	Loop genomics sequencing .....	78
2.4.2.	QIIME OTU identification .....	79
2.4.3.	Krona plots.....	79
2.4.4.	Diversity statistics. ....	79
2.4.5.	PERMANOVA.....	79
2.4.6.	Taxonomic biomarkers and functional biomarkers.....	79

Chapter 3.	Optimisation of Bacterial and DNA isolation from Soil and Sample Collection.	81
3.2.	Introduction:	82
3.3.	Results	88
3.3.1.	Identification of bacteria from Forest Research samples	88
3.3.2.	Optimisation of the dispersal of bacteria from soil aggregates	93
3.3.3.	Optimisation of soil DNA extraction	98
3.3.4.	Soil storage optimisation	100
3.3.5.	Storage of DNA in DNA Shield	103
3.3.6.	Hatchlands Park Sampling	104
3.3.7.	LAMP confirmation of oak roots	107
3.3.8.	AOD confirmation of biological sample bleeds	109
3.4.	Discussion	110
Chapter 4.	Environmental reservoirs of the AOD bacteria.	114
4.2.	Introduction	115
4.3.	Results	117
4.3.1.	Detection of AOD bacterial DNA in soil	117
4.3.2.	EMB colour testing	118
4.3.3.	Growth curves for soil spiking experiments	122
4.3.4.	Survivability of AOD bacteria in soil	125
4.3.5.	Survivability of AOD bacteria in EE broth	127
4.3.6.	Survival and recovery of AOD bacteria oak-related niches	127
4.3.7.	Comparison of single isolate identities from both recovery methods	130
4.3.8.	HRM analysis of bacteria from Hatchlands Park samples	131
4.3.9.	<i>gyrB</i> identification of EE isolates	133
4.4.	Discussion	137
Chapter 5.	The Taxonomic Classification of Oak Rhizosphere Novel Species	142

5.1.	Introduction.....	144
5.2.	Results: .....	148
5.2.1.	Genotypic characterisation:.....	148
5.2.2.	Box PCR:.....	154
5.2.3.	ERIC PCR:.....	155
5.2.4.	Genomic characterisation:.....	156
5.2.5.	Cell imaging: .....	170
5.2.6.	Cell physiology: .....	174
5.2.7.	Antibiotic resistance: .....	175
5.2.8.	Phenotypic characterisation: .....	175
5.2.9.	Fatty acid and methyl ester analysis: .....	188
5.2.10.	Virulence genes analysis of <i>Scandinavium</i> species: .....	194
5.2.11.	Plant growth promoting genes analysis of <i>Leclercia</i> and <i>Silvania</i> species:..	195
5.3.	Conclusions.....	200
5.4.	Discussion .....	200
Chapter 6.	Long Read 16S rRNA Gene Sequencing of the Oak Rhizosphere Microbiome	202
6.1.	Introduction:.....	203
6.2.	Results .....	207
6.2.1.	Whole Site Composition .....	207
6.2.2.	Diversity Statistics:.....	218
6.2.3.	LEfSe analysis:.....	223
6.2.4.	PERMANOVA results .....	224
6.2.5.	Parkland paired samples .....	225
6.3.	Discussion:.....	247
Chapter 7.	General Discussion .....	251
7.1.	Conclusion .....	260

7.2. Future work .....	261
References.....	263
Supplementary data:.....	303
<i>Scandianvium</i> protologs:.....	320
<i>Dryocola</i> protologs:.....	323
<i>Leclercia</i> and <i>Silvania</i> protologs: .....	325

## List of Figures

Figure 1: Mapped locations affected by Acute Oak Decline in the UK from March 2006 to March 2021 (Forest Research, 2021). .....	25
Figure 2: Images of the three key features associated with the disease phenotype of AOD. 27	
Figure 3: The decline disease spiral.....	30
Figure 4: The updated decline disease spiral based on Acute Oak Decline as a model (Denman <i>et al.</i> , 2022). .....	31
Figure 5: The four horizons of soil and an example of micro-zones within these horizons. ..	47
Figure 6: LAMP reaction for Hatchlands Park tree 6 rhizosphere samples from each 4-cardinal point.....	73
Figure 7: The composition of a soil aggregate.....	84
Figure 8: Melt curve obtained from High resolution melt protocol 1 for the detection of AOD lesion bacteria.....	89
Figure 9: Melt curve obtained from high resolution melt protocol 2. ....	90
Figure 10: Isolation of bacteria from soil on a magnetic mixer. ....	93
Figure 11: The average morphologically distinguishable colonies from soil on magnetic stirrer. ....	94
Figure 12: Isolation of bacteria by disruption in an ultrasonic water bath. ....	95
Figure 13: The average morphologically distinguishable colonies from soil disrupted in an ultrasonic water bath.....	95
Figure 14: Isolation of bacteria in a shaking incubator.....	96
Figure 15: The average morphologically distinguishable colonies recovered by disruption in a shaking incubator.....	97
Figure 16: The amount of DNA (ng/ $\mu$ L) recovered from the soil at both different time frames in both the Disruptor genie and TissueLyser LT n=1.....	98
Figure 17: The purity of DNA from the A260/A280 ratio recovered from the soil at both different time frames in both the Disruptor genie and TissueLyser LT n=1.....	99
Figure 18: The purity of DNA from the A260/A280 ratio recovered from the soil at both different time frames in both the Disruptor genie and TissueLyser LT n=1.....	99
Figure 19: Different storage conditions effect on the number of bacteria recovered from soil. ....	100

Figure 20: Different storage conditions effect on the types of bacteria recovered from soil. .....	101
Figure 21: Different storage conditions effect on the number of bacteria recovered from soil. .....	101
Figure 22: Different storage conditions effect on the types of bacteria recovered from soil. .....	102
Figure 23: Map detailing the location of rhizosphere soil samples taken from Hatchlands Park site. ....	106
Figure 24: Results for Hatchlands H6 samples following Loop-mediated isothermal amplification. ....	107
Figure 25: Positive identification of <i>Brenneria goodwinii</i> (red), <i>Rahnella victoriana</i> (brown) and <i>Gibbsiella quercinecans</i> (green), using DNA extracted from pure cultures in a multiplex HRM analysis.....	118
Figure 26: Positive amplification of DNA from pure cultures of <i>Brenneria goodwinii</i> (red), <i>Rahnella victoriana</i> (brown) and <i>Gibbsiella quercinecans</i> (green), but negative amplification of the same respective species after spiking DNA into the soil and then performing DNA extraction on the whole soil sample. ....	118
Figure 27: Images of <i>Brenneria goodwinii</i> type strain FRB 141 <sup>T</sup> on EMB.....	120
Figure 28: Images of <i>Gibbsiella quercinecans</i> strain Gq1 on EMB.....	122
Figure 29: Growth of <i>Brenneria goodwinii</i> in batch culture over 26 hours. ....	123
Figure 30: Growth of <i>Gibbsiella quercinecans</i> in batch culture over 26 hours. ....	123
Figure 31: Growth of <i>Rahnella victoriana</i> in batch culture over 26 hours.....	124
Figure 32: Growth of <i>Pseudomonas laurylsulfatorans</i> in batch culture over 26 hours. ....	124
Figure 33: A) High-resolution melt curve for <i>Brenneria goodwinii</i> (red) <i>Rahnella victoriana</i> (brown), and <i>Gibbsiella quercinecans</i> (green) after the soil was spiked at 0 hours using method 2.4 protocol 1.....	126
Figure 34: Maximum likelihood tree based on the concatenated partial gene sequences of <i>infB</i> , <i>atpD</i> , and <i>gyrB</i> . Sequences included are for the genus <i>Scandinavium</i> , the three proposed novel species, <i>Scandinavium hiltneri</i> sp. nov., <i>Scandinavium manionii</i> sp. nov. and <i>Scandinavium tedordense</i> sp. nov., as well as close phylogenetic neighbours.....	150

Figure 35: Maximum likelihood tree based on the concatenated partial gene sequences of <i>fusA</i> , <i>leuS</i> , <i>pyrG</i> and <i>rpoB</i> from species of the proposed genus <i>Dryocola</i> gen. nov., and its closest phylogenetic neighbours <i>Cedecea</i> and <i>Buttiauxella</i> .....	151
Figure 36: Maximum likelihood tree based on the concatenated partial gene sequences of <i>atpD</i> , <i>infB</i> , <i>gyrB</i> and <i>rpoB</i> from species of the proposed genus <i>Silvania</i> gen. nov., the novel species <i>Leclercia tamurae</i> sp. nov. and their closest phylogenetic neighbours. ....	153
Figure 37: BOX PCR fingerprinting patterns generated from strains of the proposed novel species of <i>Scandinavium</i> . ....	154
Figure 38: BOX PCR fingerprinting patterns generated from strains of the proposed novel genus <i>Dryocola</i> .....	155
Figure 39: ERIC PCR fingerprinting patterns generated from strains of <i>Leclercia adecarboxylata</i> , <i>Leclercia tamurae</i> sp. nov. and the novel genus <i>Silvania</i> gen. nov. ....	156
Figure 40: Phylogenomic tree for the genus <i>Scandinavium</i> including the three proposed novel species, <i>Scandinavium hiltneri</i> sp. nov., <i>Scandinavium manionii</i> sp. nov. and <i>Scandinavium tedordense</i> sp. nov., and close phylogenetic neighbours.....	158
Figure 41: Phylogenomic tree of the proposed genus <i>Dryocola</i> gen. nov., and its closest phylogenetic neighbours.....	162
Figure 42: Phylogenomic tree of the proposed genus <i>Silvania</i> gen. nov., the novel species <i>Leclercia tamurae</i> sp. nov. and their closest phylogenetic neighbours. ....	166
Figure 43: Transmission electron microscopy of A) <i>Scandinavium goeteborgense</i> CCUG 66741 <sup>T</sup> , B) <i>Scandinavium manionii</i> sp. nov. H17S15 <sup>T</sup> , C) <i>Scandinavium hiltneri</i> sp. nov. H11S7 <sup>T</sup> and D) <i>Scandinavium tedordense</i> sp. nov. TWS1a <sup>T</sup> displaying their peritrichous flagella arrangement.....	170
Figure 44: Transmission electron microscopy of <i>Dryocola boscaweniae</i> sp. nov. H6W4 <sup>T</sup> (A, B, C) and <i>Dryocola clanedunesis</i> sp. nov. (D, E, F) displaying the peritrichous flagella, fimbriae and cell arrangement. ....	172
Figure 45: Transmission electron microscopy images of A) <i>Leclercia tamurae</i> H6S3 <sup>T</sup> B) <i>Leclercia tamurae</i> H6W5 C) <i>Silvania hatchlandensis</i> H19S6 <sup>T</sup> D) <i>Silvania confinis</i> H4N4 <sup>T</sup> displaying their peritrichous flagella arrangement.....	173
Figure 46 Krona plot representation of the major plant growth-promoting traits found in <i>Silvania hatchlandensis</i> sp. nov. (H19S6T) and <i>Silvania confinis</i> sp. nov. (H4N4T).....	197

Figure 47: Krona plot representation of the major plant growth-promoting traits found in <i>Leclercia adecarboxylata</i> (H10E4) and <i>Leclercia tamurae</i> sp. nov. (H6S3 <sup>T</sup> ).....	198
Figure 48: Phylogenetic trees on the identified taxonomy for each variable region (V1-9) of the 16S rRNA gene. ....	206
Figure 49: Krona plot of the woodland OTU identifications from the EzBiocloud Pipeline. ....	210
Figure 50 Krona plot of the parkland OTU identifications from the EzBiocloud Pipeline. ...	211
Figure 51: Krona plot of the woodland healthy OTU identifications from the EzBiocloud Pipeline. ....	213
Figure 52: Krona plot of the woodland diseased OTU identifications from the EzBiocloud Pipeline. ....	214
Figure 53: Krona plot of the parkland healthy OTU identifications from the EzBiocloud Pipeline. ....	216
Figure 54: Krona plot of the parkland diseased OTU identifications from the EzBiocloud Pipeline. ....	217
Figure 55 The alpha diversity statistics taken for the whole site comparisons, with health as the separator. ....	219
Figure 56: The beta diversity unweighted UniFrac principal coordinate analysis including unclassified reads for the full site analysis. ....	220
Figure 57: The beta diversity unweighted UniFrac principal coordinate analysis including unclassified reads for the woodland diseased and woodland healthy samples. ....	221
Figure 58: The beta diversity unweighted UniFrac principal coordinate analysis including unclassified reads for the parkland diseased and parkland healthy samples. ....	222
Figure 59: UPGMA unrooted UniFrac tree representation of the phylogenetic clustering of all samples. ....	222
Figure 60: Krona plots of the OTU identifications from the EzBiocloud Pipeline. A = parkland diseased sample H1 and B = parkland healthy sample H2. ....	227
Figure 61: The alpha and beta diversity stats generated from the H1 and H2 microbiome taxonomic profiles. ....	228
Figure 62: Krona plots of the OTU identifications from the EzBiocloud Pipeline. A = parkland diseased sample H3 and B = parkland healthy sample H4. ....	231
Figure 63 The alpha and beta diversity stats generated from the H3 and H4 microbiome taxonomic profiles. ....	232

Figure 64: Krona plots of the OTU identifications from the EzBiocloud Pipeline. A = parkland diseased sample H5 and B = parkland healthy sample H6. ....	235
Figure 65: The alpha and beta diversity stats generated from the H5 and H6 microbiome taxonomic profiles. ....	236
Figure 66: Krona plots of the OTU identifications from the EzBiocloud Pipeline. A = parkland diseased sample H7 and B = parkland healthy sample H8. ....	239
Figure 67: The alpha and beta diversity stats generated from the H7 and H8 microbiome taxonomic profiles. ....	240
Figure 68: Krona plots of the OTU identifications from the EzBiocloud Pipeline. A = parkland diseased sample H9 and B = parkland healthy sample H10. ....	244
Figure 69: The alpha and beta diversity stats generated from the H9, and H10 microbiome taxonomic profiles. ....	245
Suppl. Figure. S1. 16S rRNA gene maximum likelihood phylogenetic tree for <i>Scandinavium</i> species, the proposed novel species and their closest phylogenetic neighbours.....	310
Suppl. Figure. S2. Neighbour Joining 16S rRNA phylogenetic tree for the novel genus <i>Dryocola</i> and closest phylogenetic neighbours. ....	311
Supple. Figure. S3. Maximum Likelihood 16S rRNA gene phylogenetic tree for the novel genus <i>Dryocola</i> and closest phylogenetic neighbours. ....	312
Supple. Figure. S4. Maximum Likelihood phylogenetic tree based on 16S rRNA gene sequences for species of the novel genus <i>Silvania</i> gen. nov., <i>Leclercia</i> , the novel species <i>Leclercia tamurae</i> sp. nov. and several closest phylogenetic neighbours. ....	313

## List of Tables

Table 1: List of primers used for PCR amplifications detailing the gene amplified, the primers' official name, the primer sequence and the paper in which the primer sequence was originally published. ....	61
Table 2: PCR amplification program by gene and primer type. ....	64
Table 3: List of primers used for sequencing by Eurofins Genomics. ....	67
Table 4: Primer sequences used in HRM analysis for AOD bacteria and the reference for their origin. ....	69
Table 5: High-resolution melt conditions used for the identification of the members of the AOD lesion microbiome including primers, their concentrations and PCR protocol. ....	71
Table 6: The gene amplified in the LAMP for the identification of oak material, the primers used to disrupt and amplify that gene and the sequence of the primers used. ....	72
Table 7: The fingerprinting methods used for the identification of clonal isolates. ....	74
Table 8: Identities of bacterial spikes from Forest Research. ....	90
Table 9: Purity and concentration of DNA extracted from soil (n = 3). ....	103
Table 10: Purity and concentration of DNA extracted from soil after 7-day storage (n = 3) ....	103
Table 11: Hatchlands Park sample identities. Hatchlands Park samples divided by their original location, each tree is paired to the number in the adjacent column for that location with the health status of the tree indicated by the column name. ....	104
Table 12: Loop-mediated isothermal amplification results. ....	107
Table 13: High resolution melt results for Hatchlands lesion swabs. +ve indicates the presence of that species in the sample, while -ve indicates no detection. ....	109
Table 14: List of the strains of <i>Brenneria goodwinii</i> streaked on EMB agar and whether metallic green colonies were obtained after 48-hour incubation under aerobic and anaerobic conditions. ....	119
Table 15: List of the strains of <i>Gibbsiella quercinecans</i> streaked on EMB agar and whether metallic green colonies were obtained after 48-hour incubations under aerobic and anaerobic conditions. ....	121
Table 16: The colony forming units calculated from Miles & Misra plates (n=3) for each species used in the spiking experiments. ....	128

Table 17: The identities of AOD bacteria recovered from the spiked soil, acorns and leaves. .....	128
Table 18: The AOD bacteria identified in environmental samples from Hatchlands Park, Guildford, UK. Bg = <i>Brenneria goodwinii</i> , Gq = <i>Gibbsiella quercinecans</i> , Lb = <i>Lonsdalea britannica</i> and Rv = <i>Rahnella victoriana</i> .....	131
Table 19: The number of potentially novel isolates in each key group. ....	133
Table 20: Summary of ability of the 16S rRNA gene to identify clinically relevant Enterobacterales (formerly <i>Enterobacteriaceae</i> ).....	145
Table 21: <i>in silico</i> DNA - DNA Hybridisation d5 matrix ( <i>isDDH</i> – top right) and Average Nucleotide Identity based on BLAST (ANiB – bottom left) percentage values for novel <i>Scandinavium</i> species and <i>S. goeteborgense</i> . ....	159
Table 22: In silico DNA-DNA Hybridisation ( <i>isDDH</i> – top right) and Average Nucleotide Identities matrix ( <i>fastANI</i> – bottom left) percentage values for <i>Dryocola boscaweniae</i> sp. nov. and <i>Dryocola clanedunensis</i> sp. nov. and all species of <i>Cedecea</i> and <i>Buttiauxella</i> , the closest phylogenetic neighbours.....	163
Table 23: Genome comparison values for <i>in silico</i> DNA - DNA Hybridisation ( <i>isDDH</i> – top right) and Average Nucleotide Identity ( <i>fastANI</i> – bottom left).....	168
Table 24: Key phenotypic characteristics that allow for the differentiation of all known members of the genus <i>Scandinavium</i> . ....	175
Table 25: Key phenotypic characteristics that allow for the differentiation between <i>Dryocola boscaweniae</i> sp. nov. and <i>Dryocola clanedunensis</i> sp. nov. from each other, and existing species of <i>Cedecea</i> and <i>Buttiauxella</i> .. ....	178
Table 26: Key phenotypic characteristics that allow for the differentiation between (1) <i>Dryocola</i> gen. nov., (2) <i>Cedecea</i> and (3) <i>Buttiauxella</i> .....	181
Table 27: Key phenotypic characteristics for differentiation of (1) <i>Leclercia adecarboxylata</i> (n = 4), (2) <i>Leclercia pneumoniae</i> 49125 <sup>T</sup> and (3) <i>Leclercia tamurae</i> sp. nov. (n = 5).....	182
Table 28: Key phenotypic characteristics for differentiation of (1) <i>Silvania hatchlandensis</i> sp. nov. (n = 2) and (2) <i>Silvania confinis</i> sp. nov. H4N4 <sup>T</sup> .....	185
Table 29: Key phenotypic characteristics for differentiation between (1) <i>Leclercia</i> * (n = 9) and (2) <i>Silvania</i> gen. nov. (n = 3) .....	186
Table 30: The major fatty acid methyl ester (FAME) average % peaks and standard deviation recorded for species of <i>Scandinavium</i> . ....	189

Table 31: The average percentage of peak areas making up the fatty acid methyl ester composition of <i>Dryocola</i> gen. nov., <i>Cedecea</i> and <i>Buttiauxella</i> species. ....	191
Table 32: The major fatty acid methyl ester (FAME) average % peaks and standard deviation for <i>Leclercia</i> and <i>Silvania</i> gen.nov. ....	193
Table 33: The number of shared and unique taxa identified at each taxonomic rank for woodland and parkland samples. ....	208
Table 34: The PERMANOVA results for each pairwise comparison from the full site datasets, including the Pseudo-F and q values. ....	225
Table 35: The number of shared and unique taxa identified at each taxonomic rank for paired parkland samples H1 and H2. ....	226
Table 36: PERMANOVA results for paired samples H1 and H2. ....	229
Table 37: The number of shared and unique taxa identified at each taxonomic rank for paired parkland samples H3 and H4. ....	230
Table 38: PERMANOVA results for paired samples H3 and H4. ....	233
Table 39: The number of shared and unique taxa identified at each taxonomic rank for paired parkland samples H5 and H6. ....	234
Table 40: PERMANOVA results for paired samples H5 and H6. ....	237
Table 41: The number of shared and unique taxa identified at each taxonomic rank for paired parkland samples H7 and H8. ....	238
Table 42: PERMANOVA results for paired samples H7 and H8. ....	241
Table 43: The number of shared and unique taxa identified at each taxonomic rank for paired parkland samples H9 and H10. ....	242
Table 44: PERMANOVA results for paired samples H9 and H10. ....	246
Suppl. Table. S1. Comparison of the sequencing identifications o ....	303
Suppl. Table. S2. The whole genome sequence information for strains investigated in this study. ....	314
Suppl. Table. S3. Whole Genome sequence information of strains investigated in this study. ....	316
Suppl. Table. S4. Whole Genome sequence information of strains investigated in this study. ....	318

Suppl. Table. S5. The hyperlink to the interactive krona plots for the 16S rRNA data collected for each sample from Hatchlands park.....	317
--	-----

## Abbreviations

AAI = average amino identity

ANI = average nucleotide identity

AOB = ammonia oxidising bacteria

AOD = Acute oak decline

*atpD* = ATP synthase  $\beta$ -subunit

AVR = avirulence

BPW = Buffered Peptone Water

CBA = Columbia blood agar

CFU = colony forming units

COD = Chronic oak decline

DAI = Decline acuteness index

dNTPs = deoxyribonucleotide triphosphates

DRS = DNA/RNA shield

EE = Enterobacteriaceae Enrichment

EMB = eosin methylene blue

FAME = fatty acid methyl ester

FISH = fluorescence *in situ* hybridisation

*fusA* = elongation factor G

GAM = Gifu Anaerobic Medium

GBDP = Genome Blast Distance Phylogeny

GFP = Green fluorescent protein

*gyrB* = Gyrase  $\beta$

HR = hypersensitive reaction

HRM = High resolution melt

*infB* = translational initiation factor  $\beta$

*isDDH* = *in silico* DNA – DNA hybridisation

ITS1 = internal transcribed spacer region 1

LAMP = Loop mediated isothermal amplification

LB = Luria-Bertani

LEfSe = linear discriminant analysis (LDA) effect size  
*leuS* = leucine tRNA ligase  
MLSA = Multilocus sequencing analysis  
MTP = microbiome taxonomic profile  
NaOH = sodium hydroxide  
OD = optical density  
OTUs = Operational taxonomic units  
PAMPs = pathogen-associated molecular patterns  
PCoA = principle coordinate analysis  
PCR = Polymerase chain reaction  
PCWDE = plant cell wall degrading enzymes  
PDI = phenotypic decline index  
PERMANOVA = permutational multivariate ANOVA  
PGPB = Plant growth promoting bacteria  
PGPT = plant growth promoting traits  
PIECRUSt = phylogenetic investigation of communities by reconstruction of unobserved states  
PLaBAs = PLant-associated Bacteria web resource  
PTI = PAMP-triggered immunity  
*pyrG* = CTP synthase  
qPCR = quantitative PCR  
R2A = Reasoner's 2 agar  
RISA = Ribosomal intergenic spacer analysis  
RPM = revolutions per minute  
*rpoB* = RNA polymerase  $\beta$   
SAP = Shrimp Alkaline Phosphatase  
SAR = systemic acquired resistance  
SDS = sodium dodecyl sulfate  
SLR = synthetic long read  
SMS = smart model selection  
SPR = Subtree Pruning and Regrafting  
SPRI = Solid Phase Reversible Immobilisation

T2SS = type II secretion systems

T3SS = type III secretion systems

T6SS = Type VI Secretion System

TEM = Transmission electron microscopy

TNTC = Too numerous to count

TSA = Tryptone soy agar

TSB = tryptone soy broth

TSOB = two spotted oak borer

TYGS = Type (Strain) Genome Server

UPGMA = unweighted pair group method with arithmetic mean

VFDB = virulence factor database

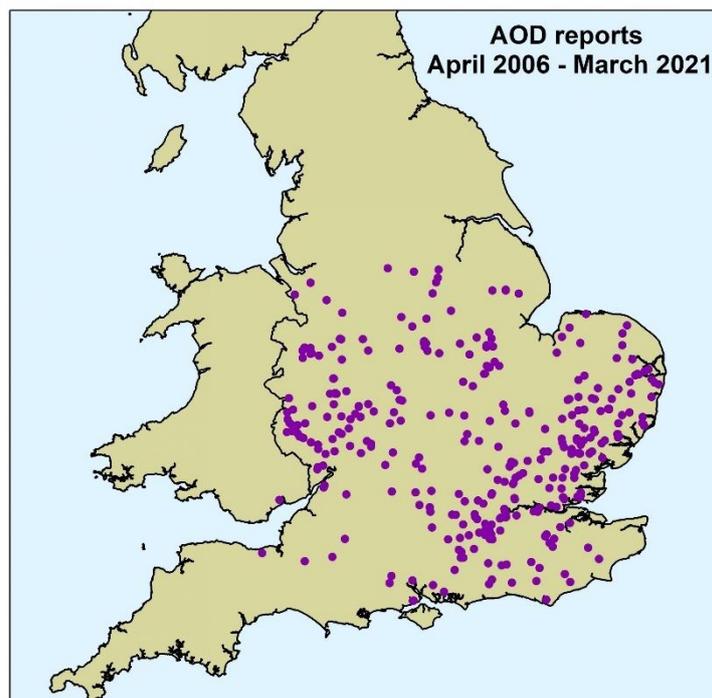
WGS = Whole genome sequence

# Chapter 1. Introduction

## 1.1. Acute oak decline.

### 1.1.1. Description and occurrence.

Acute oak decline (AOD) is a dieback disease seen to originate on native British oak trees *Quercus robur* and *petraea*. Symptoms have been observed since the 1980s, with the disease being officially defined in 2009 by Forest Research (Denman and Webber, 2009). The disease differs from the other major dieback diseases affecting oak trees in the last hundred years, including Sudden Oak Death (SOD) and Chronic Oak Decline (COD). The key differences are the rapid rate at which disease symptoms develop, the higher rate of mortality and the speed at which tree death occurs, 4-6 years (Denman and Webber, 2009). The disease was first recorded in Southeast Anglia, spreading west to The Welsh border and down to the south coast. A map of disease distribution can be seen in Figure 1 (Brown *et al.*, 2017a; Denman, Kirk and Webber, 2010).



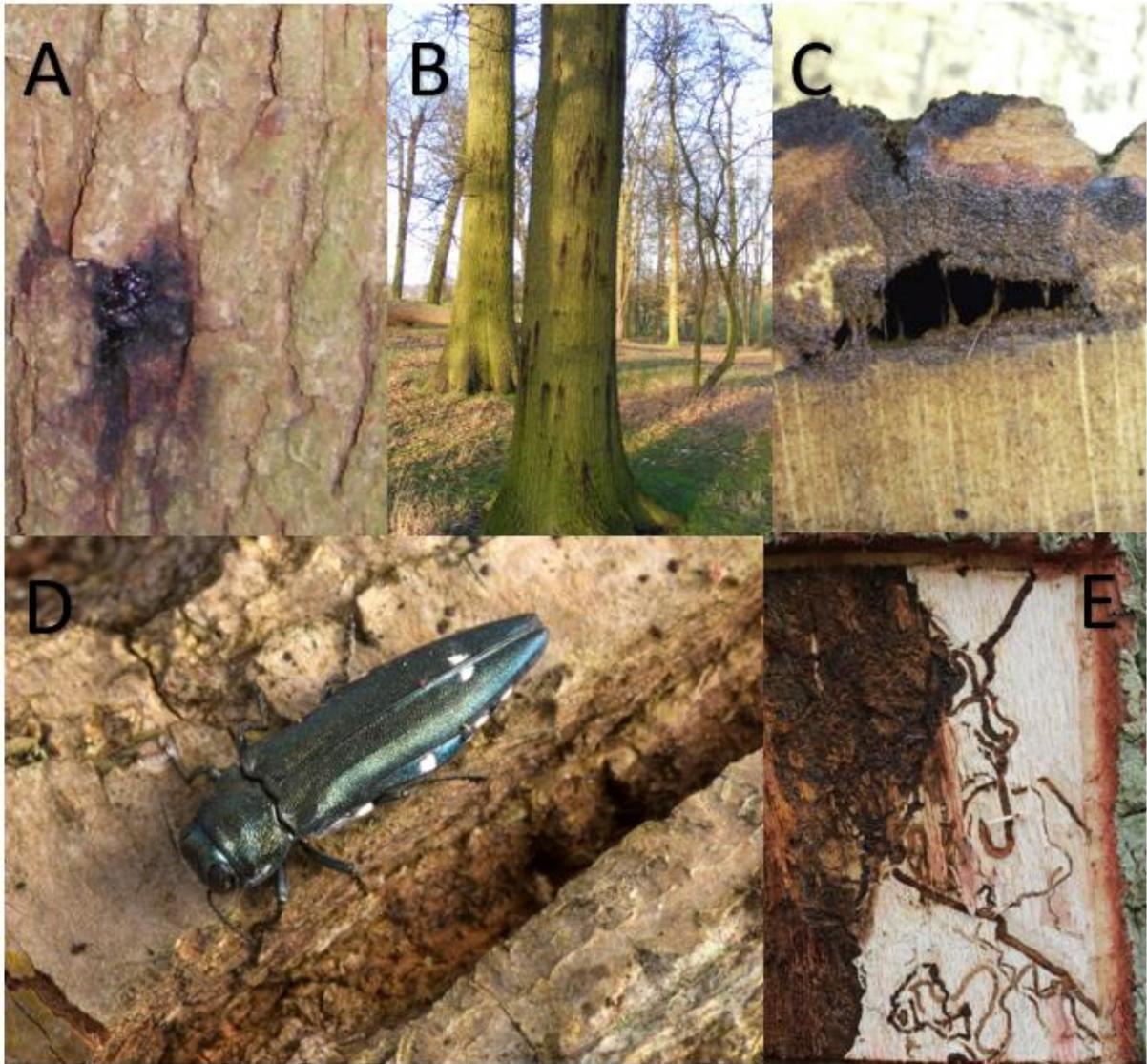
**Figure 1:** Mapped locations affected by Acute Oak Decline in the UK from March 2006 to March 2021 Reproduced from (Forest Research, 2021).

Symptomatic AOD trees are seen to be distributed in nearly all habitats, including urban, parkland, farmland and woodland environments. Sites tend to follow a pattern of infection in which a small number of trees show symptoms, with numbers rapidly increasing until more than half of the oaks are afflicted (Denman, Kirk and Webber, 2010).

Throughout history, decline of both native British oak and many different species (*Quercus; frainetto Ten, ilex L, suber L, pubescens and cerris L*) with similar symptoms has been seen throughout the majority of Europe (Thomas, Blank and Hartmann, 2002). This current episode of decline is no different, with disease now being recorded on a range of non-native oak in the UK including Bali oak (*Quercus fabri*), holm oak (*Quercus ilex*), oriental white oak (*Quercus aliena var. acutiserrata*), pin oak (*Quercus palustris*), Pyrenean oak (*Quercus pyrenaica*), red oak (*Quercus rubra*), scarlet oak (*Quercus coccinea*), Turkey oak (*Quercus cerris*), water oak (*Quercus nigra*), chestnut-leaved oak (*Quercus castaneifolia*), Persian oak (*Quercus brantii*) and downy oak (*Quercus pubescens*) (Crampton, Brady and Denman, 2022). The disease has also been seen in several other countries, most recently in Spain, Switzerland, Portugal, Latvia and Iran (Fernandes *et al.*, 2022; Zalkalns and Celma, 2021; González and Ciordia, 2020; Ruffner *et al.*, 2020; Moradi-Amirabad *et al.*, 2019).

#### 1.1.2. Descriptive symptoms of AOD

While both native UK species of oak are affected by AOD, the sessile oak (*Q. petraea*) is less commonly afflicted than the pedunculate (*Q. robur*). Mature trees over the age of 50 years are more vulnerable, though occurrence on young trees has also been observed (Brown *et al.*, 2016). The disease can be identified through three key symptoms which include stem bleeding, where clear dark exudate can be seen from patches on the trunk of the tree, unless yeast colonises the infection leading to the exudate becoming cloudy. Weeping predominantly occurs through the months of March-June and October-November. The positioning of stem bleeds varies vertically and horizontally from 5-22 cm apart (Brady *et al.*, 2017). The second symptom, seen in more advanced cases is 3-15 vertical cracks between the bark plates, from which stem bleeding occurs. Underneath the cracked outer bark, the tissue is stained and necrotic, leading to lesions that penetrate the sapwood, but never as far as the heartwood. The third characterisation that is seen to later into the disease is the presence of larval galleries of the two-spotted oak borer (*Agrilus biguttatus*) close to the previously described lesions (Denman *et al.*, 2014). These larval galleries are not consistently recorded, and as such their role in AOD has not been confirmed (Denman *et al.*, 2016). These phenotypic features associated with AOD can be seen below (Figure 2).



**Figure 2:** Images of the three key features associated with the disease phenotype of AOD. **A**, **B** external symptom of weeping stem bleeds. **C** side view of underlying necrotic lesion causing a crack in the bark plate. **D** mature *Agrilus biguttatus* beetle. **E** necrotic lesions and closely situated larval galleries of *Agrilus biguttatus* (Brady *et al.*, 2017).

The symptoms of AOD show phenotypic similarities to previous cases of decline in oak with crown thinning, twig abscission in the upper canopy and bud dieback observed. Leaves are reduced in size, discoloured, and sometimes yellowing is observed (Thomas, 2008). Necrosis of bark and cambium, with weeping and reduction in size and diameter, also occurs (Thomas and Hartmann, 1996). Despite being a symptom in most cases of decline, crown thinning is normally a response to periods of severe drought, frost, defoliation or reduced water uptake due to impaired root growth (Thomas and Hartmann, 1996). Canopy condition, though nonspecific to AOD is a good indicator of tree health and may show levels of deterioration,

however, in many cases this is not notable till death occurs (Forest Research, 2020; Brown *et al.*, 2016). The phenotypic profiles based on the symptoms above and other symptoms of diseased oak from 174 *Q. robur* across nine sites have been used to construct a Phenotypic Decline Index (PDI) and Decline Acuteness Index (DAI). Both indexes can be used for the quantification and differentiation of AOD and COD. PDI relies largely on the crown size and quality while tree stature and stem bleeds contribute to the DAI (Finch *et al.*, 2021). These indexes demonstrate how phenotypic traits are specifically tied to decline diseases and can be used for monitoring and management.

## 1.2. Models of Decline diseases

Diseases with complex aetiology that cannot be defined as biotic or abiotic by cause are often referred to as decline diseases. This is a catchall term which covers the events, combined or successive, which disrupt the essential processes of the tree until loss of vitality leads to death (Shigo, 1986). The fundamentals of decline rely on Sinclair's 'predisposition and contributing factors' that lead to the inability to process stress due to the weakened state of the tree. Predisposition factors are stresses responsible for the initial development of decline, while secondary 'contributing' factors like opportunistic pathogens cause the symptoms and death of the tree (Ostry, Venette and Juzwik, 2011; Wargo, 1996). Trees constantly interact with both positive and negative abiotic and biotic forces that impact growth and development. An injury, infection or stress that stops the tree from responding to positive interactions, will also stop the resistance to negative interactions leading to decline. Decline is used to explain disorders which are not fully understood and are caused by multiple factors. Four models have been used to explain the main causes of decline (Sinclair and Hudler, 1988; Sinclair, 1965, 1967).

- i. Decline that is caused primarily by one consistent stress factor. This covers the slow-acting parasitic/viral action of some phytopathogens.
- ii. Decline caused by a primary injury and secondary stress. This covers a short-term primary event, followed by an opportunistic secondary event, preventing remission from the initial damaged. Alone, neither factor would cause decline.

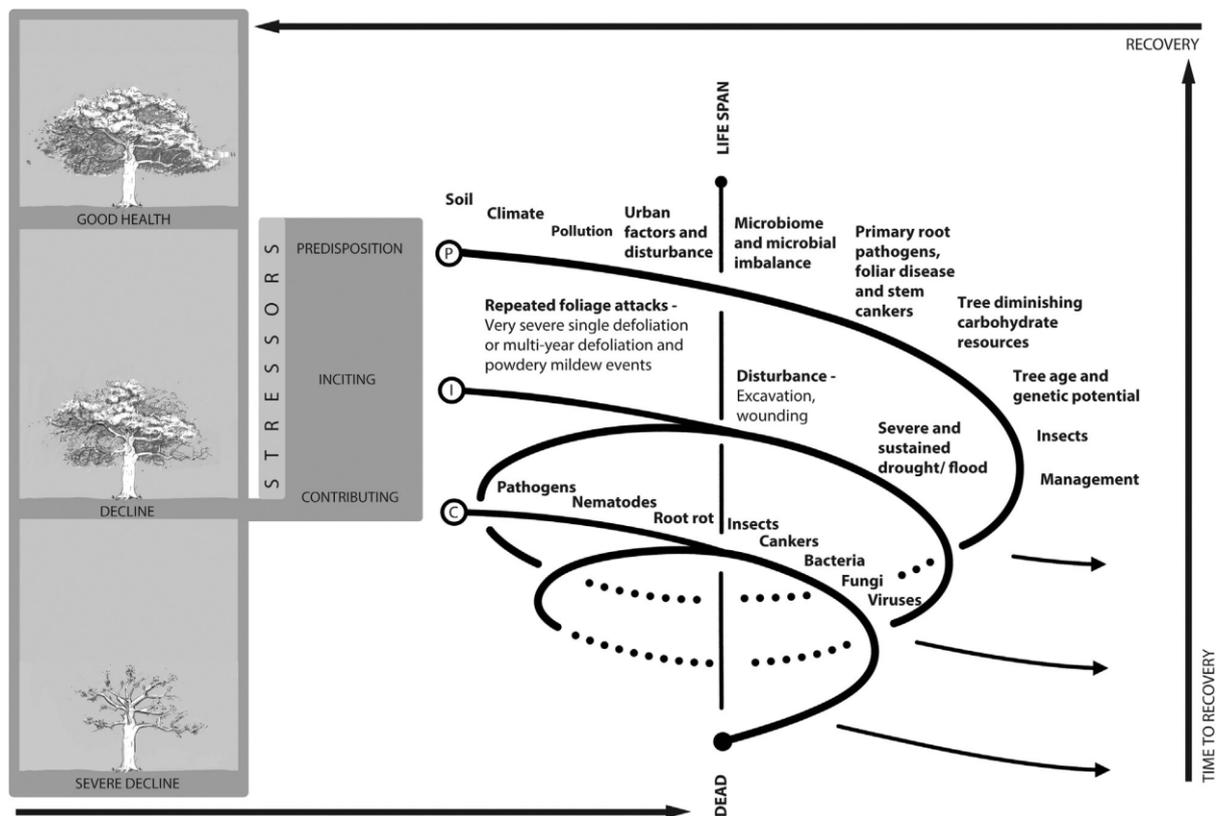
- iii. Decline caused by interchangeable predisposing, inciting and contributing factors. Inciting factors are short-term events like drought or insect defoliation, which cause initial decline symptoms, reducing stress tolerance. Contributing factors are any other factors that contribute to decline, such as age, genetics and further biotic/abiotic events.
- iv. Cohort sensing follows model iii, but predisposition is caused by age combined with tree number exceeding sustainable levels for the site enabling other factors to cause decline.

Sinclair's model of decline offers a useful framework, with multiple concepts making it versatile and adaptable, as decline diseases do not suit a 'one size fits all' diagnosis. Model iii can be applied to AOD, as it suffers from a complex aetiology, with no single causative agent, and an interchangeable sequence of events. However, other models of decline exist such as Houston's Decline concept, in which environmental stress (abiotic factors) weakens the tree allowing for secondary attack from an otherwise unthreatening organism (biotic factors) (Houston, 1987). The order of events in Houston's Decline concept is specific, which in AOD is currently unclear. The main model of decline for AOD is attributed to Manion who introduced the decline spiral shown (Figure 3), combining the same combination of interacting factors but with a key focus on the temporal sequence and their interaction (Ostry, Venette and Juzwik, 2011; Manion, 1981)



**Figure 3:** The decline disease spiral. The series of events is divided by predisposing, inciting and contributing factors that lead to the death of a tree adapted from Manion, 1981.

The decline disease spiral has recently been updated with several new features found to contribute to the development of disease and overall death of a tree, including roots and the rhizosphere (Figure 4). These updates use AOD and its associated research performed in the last 15 years as a model (Denman *et al.*, 2022). One conclusion reached in the updated model is the requirement for a holistic research approach which fully appreciates the interaction between both the biotic and abiotic factors which contribute to decline diseases.



<b>Soil</b>	<b>Climate</b>	<b>Pollution</b>	<b>Microbiome and microbial imbalance</b>	<b>Primary root pathogens, foliar disease and stem cankers</b>	<b>Insects</b>	<b>Management</b>
Compaction.	Low rainfall.	Excess atmospheric and soil pollutants (e.g. nitrous oxides).	Rhizosphere soil – imbalances in bioconversion of nutrients (e.g. Nitrifying microbes).	Buttress root rots (e.g. <i>Phytophthora</i> , <i>Gymnopus</i> , <i>Armillaria</i> ).	Repeated Insect attacks by defoliators and live-bark boring beetles.	Stock on land.
Low moisture holding capacity.	High temperatures.	Imbalances of nutrients (e.g. P- shortage).	Disruption to bioconversion of nutrients (e.g. Mycorrhizae).	Feeder root rots (e.g. <i>Phytophthora</i> , <i>Pythium</i> , <i>Ilyonectria</i> )		Planting density.
Poor soil drainage Fluctuating water table.	Flooding.		Disruption to protective and growth promoting activities.	Foliar – Powdery mildew - repeated outbreaks.		Thinning.
Soil acidity.	Extreme events.					Mixtures.
Poor fertility.	Frosts.					Social pressures (e.g. foot fall, pathways).
Nutrient imbalances.	Sustained strong wind.					

**Figure 4:** The updated decline disease spiral based on Acute Oak Decline as a model (Denman *et al.*, 2022).

### 1.3. Abiotic causes of AOD

The abiotic environmental stresses that can be considered predisposition factors in the decline of oak are large and may be acting together or sequentially. Predisposition occurs before decline and is marked by a reduction in the growth of trees, demonstrating the inability to respond to positive environmental stimuli (Thomas, 2008; Thomas, Blank and Hartmann, 2002). Between 2013 - 2014 extensive surveys of 544 locations were used to assess the effect of soil type, climate and pollutant depositions on predisposition. These factors influence water availability and drought-prone areas are correlated with increased AOD occurrence

(Forest Research, 2014). Although oaks are considered drought tolerant, droughts still limit carbon assimilation due to stomatal closure, photosynthesis, and the transport of sugars due to disruption of the vascular system. Alternatively, waterlogged soils are detrimental to tree growth due to limiting respiration in hypoxic conditions (Allen *et al.*, 2010). This is directly affected by soil type, with clay soils prone to waterlogging, and shallow and sandy soils limiting water storage (Thomas, 2008). The trends identified relate to drier, warmer, and longer growing periods, with low elevation being significant due to these conditions. These conditions lead to more dry deposition of chemicals in which nitrogen levels were higher at AOD sites and sulphur levels lower. A total of 39 % of AOD cases occurred on brown earth, though no relation to soil drainage was identified. The current geographic restriction of AOD to up to the Welsh boarder and southern England to the midlands indicate a limited temperature range. However, this is poorly understood with models predicting a wider range of distribution (Brown *et al.*, 2018). The correlation of AOD to warmer temperatures may be due to the increased suitability for defoliating insects such as *Agrilus biguttatus* that cause D-shaped exit holes (Brown *et al.*, 2017b). The importance of biotic features was also observed in 115-260 mature oaks between 2009 and 2012. The study found that affected trees tended to occur in non-randomly localised clusters within sites, suggesting a biotic cause of the spread and development of AOD rather than wide-scale environmental effects (Brown *et al.*, 2016).

#### 1.4. Biotic causes - *Agrilus biguttatus*

One biotic factor noted in AOD is *Agrilus biguttatus* or the two-spotted oak borer (TSOB), a European bark-boring beetle, known for playing out its larval stage in the vascular system of oak trees where it feeds. Of all four species of *Agrilus*, only *A. biguttatus* was consistently recorded, with 90 % of AOD symptomatic oaks having larval galleries internally adjacent to lesions and 30-33 % of diseased trees having the D-shaped exit holes of mature beetles (Reed *et al.*, 2018; Brown *et al.*, 2015, 2017b).

*A. biguttatus* was thought to be a declining species, but numbers and distribution have increased drastically since the 1980s, making them an expanding species. Expansion is possible because of climate change, with warm weather pushing the northern limit further and the ready availability of hosts due to oak dieback diseases (Alexander, 2003a, 2003b). *A.*

*biguttatus* association with dying oaks has been known for over 100 years but the increased rate of correlation between the beetle with oak dieback, made it a high-risk invasive pest (Yemshanov *et al.*, 2013; Gibbs and Greig, 1997). *A. biguttatus* and native oaks have co-evolved (Brown *et al.*, 2015), because of this oaks must be in a weakened state before an attack can occur (Manion, 1981). *A. biguttatus* finds suitable host oaks through the use of attractive volatile compounds that influence their olfactometers (Vuts *et al.*, 2016).

Lab cultivation was used to understand the cryptic life cycle of *A. biguttatus* which could not be observed in nature (Sallé, Nageleisen and Lieutier, 2014). Mature beetles were collected from felled AOD symptomatic oaks containing overwintering pupal stage *A. biguttatus*. Beetles were bred, and their eggs were collected and placed on felled oak stems to observe larval development at a range of temperatures. Results showed the lifespan of female beetles lasted 22-162 days.  $25 \pm 2.2$  days after emergence, oviposition took place between bark plates in cracks and crevices providing larvae with a short route to the cambial zone where they developed galleries (Brown *et al.*, 2015; Vansteenkiste *et al.*, 2004). These oviposition sites are picked through the detection of volatiles produce by oaks irrespective of health status, in the same way as host trees are selected (Vuts *et al.*, 2016). The eggs appeared to hatch in  $3 \pm 0.4$  days independent of temperature (Reed *et al.*, 2018).

The larval galleries of *A. biguttatus* (where a large part of their development takes place) are known to cause further weakening of the tree due to disruption of the phloem and xylem (Evans, Moraal and Pajares, 2004). The phloem is an essential feature in trees as it allows the transportation of carbohydrates, nutrients and signalling molecules (Liesche and Patrick, 2017). Xylem disruption or girdling stops the movement of water and essential nutrients from the roots to the crown (Vansteenkiste *et al.*, 2004). Disruption of either the xylem or phloem limits the growth of the tree, which then cannot respond favourably to positive external influences on the tree.

Larvae take 1 – 2 years to emerge, during which galleries can reach 1.5 m long and 3 – 4 mm wide (Vansteenkiste *et al.*, 2004). Before emergence larvae move near the surface, for overwintering and pupal diapause (Brown *et al.*, 2015). Without the cold period, a 100 % mortality rate of eggs is recorded in most *Agrilus* species (Reed *et al.*, 2018; Liang and Fei, 2014; Saunders, 2002). Warmer weather then wakes and matures beetles, with colder

temperatures leading to lower numbers of emerging beetles. This may explain the lower number of D-shaped emergence holes than galleries recorded and why the feature is associated with south-facing tree stems, where the temperature is higher (Brown *et al.*, 2016, 2017b; Denman *et al.*, 2014). Following emergence, mature beetles are recorded feeding on foliage in the oak crown (Habermann and Preller, 2003). The severity of stem bleeds is known to be more severe in oaks with weakened crowns correlating the stages of the lifecycle of *A. biguttatus* with AOD symptoms (Brown *et al.*, 2016). As such emergence holes and stem bleeds occur in a non-independently distributed fashion. Emergence holes are more common on trees which had stem bleeds the previous year, indicating a role in the weakened response state (lowering defensive ability of host) in facilitating full development of AOD (Brown *et al.*, 2017b; Dunn, Potter and Kimmerer, 1990).

It has been hypothesised that the beetle could contribute to the spread of the AOD lesion bacteria. Alternatively, larval feeding may provide a suitable area of disruption in tissue for bacteria to establish themselves. Whether acting as a vector or through/by providing a site for bacterial infection, *A. biguttatus* contribute to the formation of lesions under the bark plate, thus contributing to the mortality of trees suffering from AOD, making *A. biguttatus* an essential feature of the decline complex (Brown *et al.*, 2017b; Denman *et al.*, 2014; Vansteenkiste *et al.*, 2004).

### 1.5. Biotic factors - Bacterial components

One essential feature that distinguishes AOD from other current cases of oak decline, is the role of several bacterial species in the formation of severe symptoms such as necrosis, lesions and stem bleeds. These are known to be key features of the lesions, as if the cause was purely caused by the *A. biguttatus* wounds would be expected to heal and callus over quickly. This is not the case with AOD in which lesions are seen to have expanding margins until remission or death occurs, which is most commonly explained by a casual microbial agent (Denman *et al.*, 2016; Donaubauer, 1987).

To further understand the biotic component of AOD the key research outputs then focused on the taxonomic classification of the microbiome, the detection of the AOD pathogens and understanding their roles within AOD lesions.

## 1.6. Taxonomic classification of the microbiome

The identification of bacterial isolates from necrotic tissue of symptomatic oak began in 2009. The isolates were expected to be closely related to *Brenneria quercina* or *Serratia*, which had been suggested as the causal agent of declining Mediterranean oak isolates from Spain (Poza-Carrión *et al.*, 2008; Biosca *et al.*, 2003). Phenotypic tests placed the isolates in the *Enterobacteriaceae*, with 16S rRNA gene sequencing dividing the isolates into two groups showing close relation to *Brenneria* and *Serratia* species. Due to the poor resolution of the 16S rRNA gene, DNA gyrase (*gyrB*) and RNA polymerase  $\beta$  (*rpoB*) genes were used for further identification (Brady *et al.*, 2010, 2012). Further investigation employed polyphasic taxonomy utilising DNA-DNA hybridization, G+C content, standard phenotypic and biochemical metabolism which were used to classify species at the time (Gevers *et al.*, 2005). Based on these tests clear phenotypic and phylogenetic differences distinguished the species from its closest phylogenetic neighbours, and as such the name *Gibbsiella quercinecans* was proposed for the group of isolates related to *Serratia* species (Brady *et al.*, 2010).

Meanwhile, a group of *Brenneria quercina*-like isolates from Spain and Britain proved harder to classify, 16S rRNA gene comparisons and phylogenetic trees clustered the isolates with *B. quercina* but were far removed from the remaining *Brenneria* species. Other studies had shown that monophyletic clades did not form for *B. quercina* when phylogenetic trees were constructed, a common issue for the *Enterobacteriaceae*, meaning isolates required investigation by multilocus sequencing analysis (MLSA) (Brady *et al.*, 2012). The method was first utilised by Maiden *et al.* (1998) and has since been used for the delineation of the highly recombinogenic bacterial genus *Neisseria*, where individual loci could not distinguish between species (Hanage, Fraser and Spratt, 2005). Brady *et al.* (2008) had previously used this method for a taxonomic re-evaluation of the genus *Pantoea*, which led to the description of nine novel *Pantoea* species and the transfer of *Pectobacterium cypripedii* to the now emended genus *Pantoea* (Brady *et al.*, 2009, 2010a, 2010b).

The four genes (*rpoB*, *atpD*, *gyrB* and *infB*) used in the *Pantoea* studies were later applied to a range of phytopathogenic bacteria allowing for the taxonomic examination of several genera and species, including the *B. quercina* like isolates. The four genes were used to build a concatenated nucleotide phylogenetic tree where the *B. quercina*-like isolates were

removed from existing *Brenneria* species. Polyphasic taxonomic results supported by MLSA gave evidence to support the creation of a novel genus *Lonsdalea* and emendations to the *Brenneria* genus and *Dickeya dadantii* and *dieffenbachiae*. *B. quercina* was reclassified as *Lonsdalea quercina* and the Spanish and British oak isolates were classified as subspecies namely *L. quercina* subsp. *iberica* and *L. quercina* subsp. *Britannica*, respectively (Brady *et al.*, 2012). The two subspecies were later elevated to species level following the identification of *Lonsdalea populi* isolated from bark canker of poplar trees which provided further taxonomic resolution to the genus (Li *et al.*, 2017). The MLSA scheme created by Brady *et al.* 2008 allowed for further classification of novel isolates from AOD symptomatic oak and is noted for its application in the re-evaluation of the prokaryotic genera, especially with regards to the *Enterobacteriaceae* (Brady *et al.*, 2017; Glaeser and Kämpfer, 2015a). The polyphasic approach used for *Lonsdalea* was later used to classify the remaining *Brenneria*-like isolates leading to the proposal of *Brenneria goodwinii* (Denman *et al.*, 2012).

As AOD occurrence increased, so did sampling frequency and the number of isolates that could not be classified. These isolates consisted of two groups, consisting of a distinct and likely novel *Brenneria* species and several species related to *Rahnella aquatilis*, at the time the only species in the genus *Rahnella* (Brady *et al.*, 2014a, 2014b). The novel *Brenneria* species had been isolated from symptomatic oaks in both the UK and the USA and contained some minor genetic differences and high spatial separation. This led to the classification of *Brenneria roseae* and subspecies *B. roseae* subsp. *roseae* and *B. roseae* subsp. *americana* (Brady *et al.*, 2014a). The *Rahnella* species had been isolated from a wide range of locations, including alder and walnut logs, AOD symptomatic oaks and beetles and were distinct from *R. aquatilis*. The names given to these species were, *Rahnella inusitata*, *Rahnella bruchi*, *Rahnella woolbedingensis*, *Rahnella variigena* and *Rahnella victoriana*, with the final two being commonly isolated species seen in AOD microbiome (Brady *et al.*, 2014b).

The lesion pathobiome is seen to be dominated by these members of the order *Enterobacterales*, which are known to consist of commensals, mutualists and phytopathogens (Adeolu *et al.*, 2016a). However a group of non-fermentative Gram-negative bacilli, belonging to the genus *Pseudomonas* have also been identified (Wisplinghoff, 2017). These *Pseudomonas* species were consistently isolated from AOD affected bark and the TSOB larval galleries in a later study which aimed to investigate and compare the composition of the

symptomatic vs the non-symptomatic microbiome (Denman *et al.*, 2016). These isolates were classified as *Pseudomonas daroniae* and *Pseudomonas dryadis* (Bueno-Gonzalez *et al.*, 2019).

The initial bulk of the taxonomy surrounding the AOD-associated microbiome had taken eight years (not including the more recent description of *Pseudomonas* species). Within those eight years, high numbers of novel species and discoveries and emendations of new and existing genera had taken place, highlighting the poor understanding of the microbiome of the native oak species before the occurrence of AOD. In light of this, it was hard to ascribe roles in the disease to the wide range of species discovered so far (Brady *et al.*, 2017).

### 1.7. Understanding the role of the oak microbiome

Understanding the composition of the plant microbiome is crucial as it supports essential functions with the host, as seen in most eukaryotic organisms. The associate microbes found within both the endo and ectosphere of plants have received much attention as growth-promoting and biological control agents to be utilised in agriculture. This is because they are seen to play active roles in resource acquisition, biological control of phytopathogens and promote stress tolerance in plants, with a direct and overall contribution to the growth of plants (Berg *et al.*, 2013, 2014; Compant, Clément and Sessitsch, 2010). Endophytic microbes are generally considered to be non-pathogenic, however, they should be considered latent pathogens, which given the correct environmental stimulus or genomic nuance of the host may trigger disease (James and Olivares, 2010).

The V3-V5 region 16S rRNA gene is a gold standard taxonomic marker found in both bacteria and archaea (Clarridge., 2004), which Sapp *et al.* (2016) used for their metabarcoding study of oak endophytes. Samples showed high variation in the number of sequences obtained, and as such rarefaction was applied to allow comparison of libraries (Sapp *et al.*, 2016). However, single taxa were not rarefied as they are known to play important roles in both the environment and host-associated microbiomes (Jousset *et al.*, 2017). For example, rare taxa in the plant microbiome have been recorded to play important roles in the protection of the host by the production of antagonistic volatiles that kill pathogens (Hol *et al.*, 2015). No clear shift in microbiome was identified in diseased tissue, but two taxa in the Gammaproteobacteria were associated with healthy tissue and one phylotype from the *Enterobacteriaceae* with diseased (Sapp *et al.*, 2016).

A second study investigated the origin, biology, and pathogenicity of the lesion bacteria. A range of molecular methods was used, but to find appropriate bacteria to test Koch's postulates, biological and phenotypic information are needed, which required culturing (Denman *et al.*, 2016). Koch's postulates are a set of criteria that prove a pathogenic microbe is the causal agent in a specific disease and are as follows (Koch, 1884):

1. The casual microorganism is seen as present in every diseased host and not present in every healthy host.
2. The microbe must be isolated and cultured from the diseased host.
3. Inoculation of the microbe into a healthy host causes the disease.
4. The identical microbe must be reisolated from the experimentally infected host.

Falkow, Colwell and others are recognized for their critiques and adaptations of the criteria, with even Koch appreciating its shortcomings concerning cholera not consistently fulfilling the third postulate (Ingles, 2007). Nevertheless, the criteria can be used to understand the root microbial cause of disease.

Bacteria were isolated from the outer and inner bark, sap and heartwood of 10 diseased and five healthy trees from five sites. 16S rRNA and *gyrB* gene PCR and sequencing were performed on single colony isolate subcultures, with BlastN used to identify them. Multivariate analysis was performed on bacterial communities at the species level, to understand composition based on different sampling combinations. Diseased tissue was observed to contain significantly higher numbers of cultivable bacteria than other tissue types. The phloem and sapwood also had higher numbers of cultivable bacteria than the outer bark, but no relation to tissue condition and tissue position was found. At all sites, symptomatic tree bacterial communities were dominated by the *Enterobacteriales*, namely, *B. goodwinii*, *G. quercinecans*, *R. victoriana*, *Erwinia billingiae*, *P. daroniae* and *P. dryadis*. The composition of the Gram-negative microbiome of oak in decline differed largely from the Gram-negative microbiome seen in healthy oaks. Further differences were found based on the severity of AOD, with a higher abundance *G. quercinecans* representing the early stages

of the disease and *B. goodwinii* the latter. Overall, the disease microbiome displayed higher diversity, but species overlap in healthy and diseased samples was observed, which was in line with the results from Sapp *et al.* (2016). As such the study suggested that the necrogenic potential of *B. goodwinii*, *G. quercinecans* and *R. victoriana* should be investigated (Denman *et al.*, 2016).

A separate study which investigated the effects of host age and spatial location on the microbiome of *Q. robur* also supported these results indirectly. 16S rRNA gene amplification of 192 samples showed that alpha species diversity decreased with age. The key addition to the understanding of AOD was that no *Brenneria* species were found, indicating *B. goodwinii* was only found in the disease microbiome. However, the study only gained taxonomic resolution of microbiome composition to a class level and the same trends were not found in the diversity of OTUs (Meaden, Metcalf and Koskella, 2016).

A trend had begun to appear with *B. goodwinii*, *G. quercinecans*, *R. victoriana* and *L. quercina*, to a lesser extent, at the heart of the issue. One focus would aim for rapid identification of the AOD lesion associated bacteria, allowing for the description of AOD-afflicted trees, although this work had begun simultaneously with the investigation of the microbiome shift. Another focus would be to test further aspects of Koch's postulates, such as infection, re-isolation and looking for specific genes associated with plant pathogenicity.

#### 1.8. Rapid identification methods for key members of the diseased microbiome.

Due to the lesion microbiome being dominated by the *Enterobacterales*, phenotypic differentiation and identification of isolates could not be achieved (Brown and Benson, 2009). The 16S rRNA gene is also not sufficient for the differentiation of the *Enterobacterales* (Naum, Brown and Mason-Gamer, 2008). MLSA was the only reliable way to identify the AOD lesion bacteria, which was time-consuming and laborious (Brady *et al.*, 2016, 2017). The search for rapid identification methods began.

In 2014 Doonan *et al.*, began to look for a cost-effective, rapid and accurate method to screen for *B. goodwinii* and *G. quercinecans* in potential AOD symptomatic tissue. The internal transcribed spacer region 1 (ITS1), was a strong candidate, often used for species recognition in a range of environmental samples (Weig *et al.*, 2013; Guasp *et al.*, 2000). It is suitable as it

varies in length and sequence allowing differentiation of both distant and closely related species (Scheinert *et al.*, 1996). Species differentiation after amplification can be seen via amplicon patterns visualised in agarose gels, because of the large difference in the length of fragments from 2 base pairs to 301 bp (Doonan *et al.*, 2015; Fisher and Triplett, 1999). Two strains of *B. goodwinii*, *G. quercinecans* and *Rahnella* species were tested, with identities confirmed using *gyrB* sequencing. In 3 % w/v agarose, reproducibility of results was inconsistent and in polyacrylamide gel electrophoresis, manual validation was required and *gyrB* sequencing level resolution was not obtained. However, for certain strains the higher copy number of ITS1 revealed differences that were not present in the single copy *gyrB* gene, allowing for greater phylogenetic resolution. Overall the method was deemed insufficient for genus-level understanding of AOD bacteria, due to the variability of the marker which limited comparisons within a single genus (Doonan *et al.*, 2015).

The second method tested by Brady *et al.* was a real-time PCR method called high-resolution melt (HRM) analysis. The method identifies single nucleotide polymorphisms in short areas of target genes without the need for sequencing. HRM analysis is based on the binding of fluorescent dyes which actively intercalate with double-stranded DNA. When exposed to increasing temperatures, DNA is seen to dissociate from double to the single-stranded form, causing the detachment of... and loss of fluorescence which leads to a signal change (Er and Chang, 2012). The G + C content and length of the amplified sequence control the rate at which fluorescence is lost. A melt curve profile is generated which can distinguish variation in nucleic acid bases due to its specificity and sensitivity. This is because variations in the target gene alter the temperature at which dissociation occurs, as such differences in the melt curve demonstrate sequence variants (Brady *et al.*, 2016; Garritano *et al.*, 2009).

Use of reference-based curves can then be used to genotype these single nucleotide polymorphisms, which has yielded a wide range of microbiological applications, such as the identification of yeast, mycobacteria as well as differentiation of bacterial species and strains (Słomka *et al.*, 2017). A short section of the *atpD* gene was selected as the target as it showed the most sequence variation within MLSA. The melt curve generated separated *G. quercinecans*, *B. goodwinii*, *B. roseae* subsp. *roseae* and *L. quercina* subsp. *britannica* into distinct peaks based on their genera, species and subspecies. The method remained effective up to  $10^{-4}$  dilution. It was hoped with modifications, environmental swabs taken by citizen

scientists could be directly used. This would remove the need to culture before identification, allowing rapid identification of bacteria from AOD tissue, and aiding understanding of both their role in the disease and the overall spread and occurrence of AOD (Brady *et al.*, 2016).

Simultaneously, Forest Research was investigating the TaqMan probe method for the rapid identification of low levels of AOD-related bacteria from environmental samples. TaqMan qPCR is a real-time PCR method which utilises non-extendable hydrolysis probes. The short oligonucleotide probe binds to a specific region which is controlled by the primers. The 5' terminus of the probe is labelled with a fluorescent dye while the 3' end is bound to a quencher, which together forms a donor-acceptor fluorescence resonance energy transfer pair, the same principle as that of an eclipse probe but with termini reversed (Stevenson, Hymas and Hillyard, 2005). The probe binds downstream from one primer at the target sequence as amplification proceeds. As elongation from the upstream primer reaches the 5' end of the probe, hydrolysis takes place. This disrupts the fluorescence resonance energy transfer pair, triggering the release of the dye from the 3' quencher. Dyes with specific wavelengths can be bound to the probe, meaning fluorescent illumination from the release of the dye can show the presence of specific species when detected. Moreover, the amount of fluorescence detected is proportional to the amount of PCR product in the mix, meaning it can be used quantitatively for single-species detection (Nagy *et al.*, 2017). Crampton *et al.*, develop a TaqMan PCR assay that could detect the presence of *B. goodwinii*, *R. victoriana* and *L. britannica* by their *gyrB* gene and *G. quercinecans* by its *rpoB* gene, simultaneously in single samples and quantify the number of species present. These results were confirmed by standard culturing and BlastN identification of the *gyrB* gene. The detection protocol was also non-destructive and rapid, requiring no culturing of species before detection (Crampton *et al.*, 2020).

As such there are currently two viable options for the detection of bacteria associated with the oak microbiome, both HRM analysis and TaqMan probe. The TaqMan analysis offers highly sensitive, rapid and quantitative detection of bacteria. However, the pitfalls lie mainly in the depth of resolution, with *Gibbsiella* only being identified to the genus level and *R. victoriana* and *R. variigena* being indistinguishable using currently available probes (Crampton *et al.*, 2020). Meanwhile, HRM analysis allows for species-specific detection within highly related clades and comparatively is the more affordable option, but suffers from the

requirement for single colony isolation, limiting the speed and low-level detection of bacteria from samples (Brady *et al.*, 2016).

### 1.9. In-depth analysis of key members of the AOD microbiome

Due to the consistent isolation of *B. goodwinii*, *G. quercinecans*, *R. victoriana* and *L. britannica* from AOD lesions, further investigation to elucidate their role in the development of the disease was investigated. This included single and multispecies inoculation to assess Koch's postulates and genomic comparisons to understand the virulence gene pools the species contained. MLSA of larger numbers of housekeeping genes has been used to understand strain-level evolution and ecology, as shown in *Streptomyces griseus* (Rong *et al.*, 2010). Seven genes were used to investigate 44 strains of *B. goodwinii* isolated from seven locations. No geographical pattern was identified, instead, strains were primarily clonal, and their genetic structure and evolution are shaped by mutation and recombination (Dykhuizen and Kalia, 2007). This implies *B. goodwinii* is likely an endemic species to oak, and recombination leads to new alleles which then become the dominant genotype in their spatial position and host (Kaczmarek *et al.*, 2017). A similar story exists with *P. syringae* which can be endemic in plants but also a devastating phytopathogen (Sarkar and Guttman, 2004). Around 60 pathovars of *P. syringae* have been identified and recombination allows them to inhabit and dominate their specific niches (Xin, Kvitko and He, 2018). Recombination events provide novel virulence factors and metabolic capabilities that alter their pathogenicity, creating short- and long-term evolutionary lineages of species (Feil and Spratt, 2001). The study of the 44 *B. goodwinii* strains using multilocus sequence typing (MLST) by Kaczmarek *et al.* (2017) provided insight into how *B. goodwinii* could evolve in oak trees and then shift to become a causal agent of AOD.

The second investigation characterised the metagenome and metatranscriptome from both healthy and AOD symptomatic oaks. *B. goodwinii*, *G. quercinecans* and *R. victoriana* were inoculated as pure cultures, polymicrobial mixtures and in combination with and without *A. biguttatus* eggs to simulate the symptoms of AOD. Inoculations of *Q. robur* logs in controlled chambers and 25-year-old oaks in the field took place over three years (Denman *et al.*, 2018). Genome analysis showed *G. quercinecans* and *R. victoriana* contained phytopathogenic genes related to soft rot *Enterobacteriaceae* phytopathogens (Denman *et al.*, 2018). The group was

of interest due to their taxonomic relation and shifting relationship from commensal to necrotrophic pathogens containing virulence genes. These include plant cell wall degrading enzymes (PCWDE) to digest the polysaccharides that form plant cell walls, type III secretion systems (T3SS), phytotoxins, and adhesins which help in plant cell association (Pritchard *et al.*, 2016; Charkowski *et al.*, 2012). The genome-wide comparison of the AOD bacteria to the soft rot *Enterobacteriaceae* revealed pectinases, cellulases and tannases, these PCWDEs were all confirmed via phenotypic tests. Furthermore, the presence of type II secretion systems (T2SS) were seen in *G. quercinecans* and *R. victoriana*, while *B. goodwinii* contained genes coding for the T3SS.

There are six well defined secretion systems in Gram-negative bacteria, which inject virulence factors into the environment or directly into hosts. Most pathogenic bacteria encode one or more secretion systems to allow invasion of hosts, but the T3SS is the trans-kingdom protein transport system associated with pathogenic effector proteins which alter host cells (Buttner, 2012). The T2SS is seen in many Gram-negative bacteria but is most notably present in pathogenic bacteria. T2SS dependant enzymes are virulent, and loss of these enzymes or the system impairs virulence, with detrimental effects on the inhabitation of environmental niches (Cianciotto, 2005). Inoculation showed that combinations of *G. quercinecans* and *B. goodwinii* caused wounds significantly larger than those of the control, and the negative control *Erwinia billingiae* (Denman *et al.*, 2018). The results fulfilled three of Koch's postulates with presence of the AOD bacteria recorded on AOD symptomatic trees, but not at sites where AOD was absent. They were isolated and cultivated from diseased trees and back isolated from inoculated trees. However, no singular microbial cause was identified, as combinations of bacteria and *A. biguttatus* showed significantly larger wounds than single microbial inoculations. With all three bacteria causing necrosis, an adaptation of Koch postulates implicating a polymicrobial role in the disease was suggested (Denman *et al.*, 2018). Koch's postulates have previously been adapted for cases such as periodontal disease in humans, and olive knot disease which has increased severity when *Pseudomonas savastanoi* pv. *savastanoi* interacts with other endophytes in olive trees (Buonaurio *et al.*, 2015; Antiabong, Boardman and Ball, 2014). The study revealed the importance of the polymicrobial lesion microbiome and *A. biguttatus* in AOD development and symptoms, giving further direction for pathogenic gene investigation.

Host-microbiota interactions were assessed via the metagenome, metatranscriptome and metaproteome of inner bark tissue from both AOD symptomatic and symptom-free trees. Virulence factors such as PCWDEs, flagella and reactive oxygen species were identified. Host tree defence systems were also seen, including cell wall modifications, reactive oxygen species and defence regulators. *B. goodwinii* was observed as the most abundant species in the AOD lesion, where core pathogenic proteins were exclusively found. This activity by *B. goodwinii* was seen actively aided by both Gram-positive and Gram-negative species of interest. The Gram-negative species based on the activity of their metatranscriptome and metaproteome were *G. quercinecans* and *R. victoriana*. Their main role was interbacterial actions and stress responses, aiding the virulence of *B. goodwinii* and the Gram-positive species. The Gram-positive species had draft genomes assigned to *Clostridioides* and *Carnobacterium* which have not currently been isolated from AOD symptomatic oak. The study showed promise in multi-omic analysis in understanding the oak hologenome in the disease and provided evidence of *B. goodwinii* as the primary disease agent (Broberg *et al.*, 2018).

In 2019 whole genome sequencing was used to find virulence mechanisms used by members of the AOD pathobiome. This approach can both identify the pathogenic mechanisms single species use, and can aid in finding tools to control the disease (Toth, Pritchard and Birch, 2006). Virulence factors including PCWDEs, secretion systems and their effectors (namely the T3SS) were chosen as key targets to detect the pathogenic potential. These genes or homologs of these can be present in non-virulent species, which can hinder understanding of the pathobiome. Symbionts have these genes for a range of reasons, including being evolutionarily conserved features or essential features for host interaction (Nishiguchi *et al.*, 2008). Recombination disrupts the genes' homeostasis, allowing bacteria to become pathogenic and occupy new niches, as suggested for *B. goodwinii* (Denman *et al.*, 2018). However, basic differences in these virulence factors exist between symbionts and pathogens. For one, the locations of unique pathogenic genes are in pathogenicity virulence islands. Domain number and domain coverage value also differ, with unique virulence genes being more compacted and less complex than common highly conserved factors (Niu *et al.*, 2013). Other issues included hemibiotrophic pathogens, which can be pathogenic or asymptomatic biotrophs depending on conditions, and saprophytes which can cause disease

at high concentrations. As such a rigorous approach to pathogenic potential would be required to understand the pathobiome-mediated disease and the virulence factors that cause AOD tissue necrosis. *B. goodwinii* was shown to have significant associations with necrotrophic virulence PCWDE homologues as well as the T3SS, associated harpins and effectors found in hemibiotrophs such as *P. syringae*. Evidence implies *B. goodwinii*'s associated pathogenic genes cause the development of tissue necrosis in the AOD lesion microbiome, which was also demonstrated by the upregulation of these genes in the lesion metatranscriptome (Doonan *et al.*, 2019; Broberg *et al.*, 2018). *G. quercinecans* virulence orthologues appeared closer to saprophytes than pathogens, which fits previous data from log inoculations and another study that isolated *G. quercinecans* from decaying wood (suggesting a saprophytic role) (Denman *et al.*, 2018; Geider *et al.*, 2015). *Rahnella* species showed high similarity to *G. quercinecans* pathogenic potential, lacking T3SS and key PCWDEs. The virulence genes they contain imply a secondary role in lesion formation, similar to endophytic species found in olive knot disease (Buonauro *et al.*, 2015). Despite consistent isolation of all three species, *B. goodwinii* contains the key virulence genes, implicating it as the causal agent of AOD, while *R. victoriana* and *G. quercinecans* likely contribute a secondary role in disease development (Doonan *et al.*, 2019).

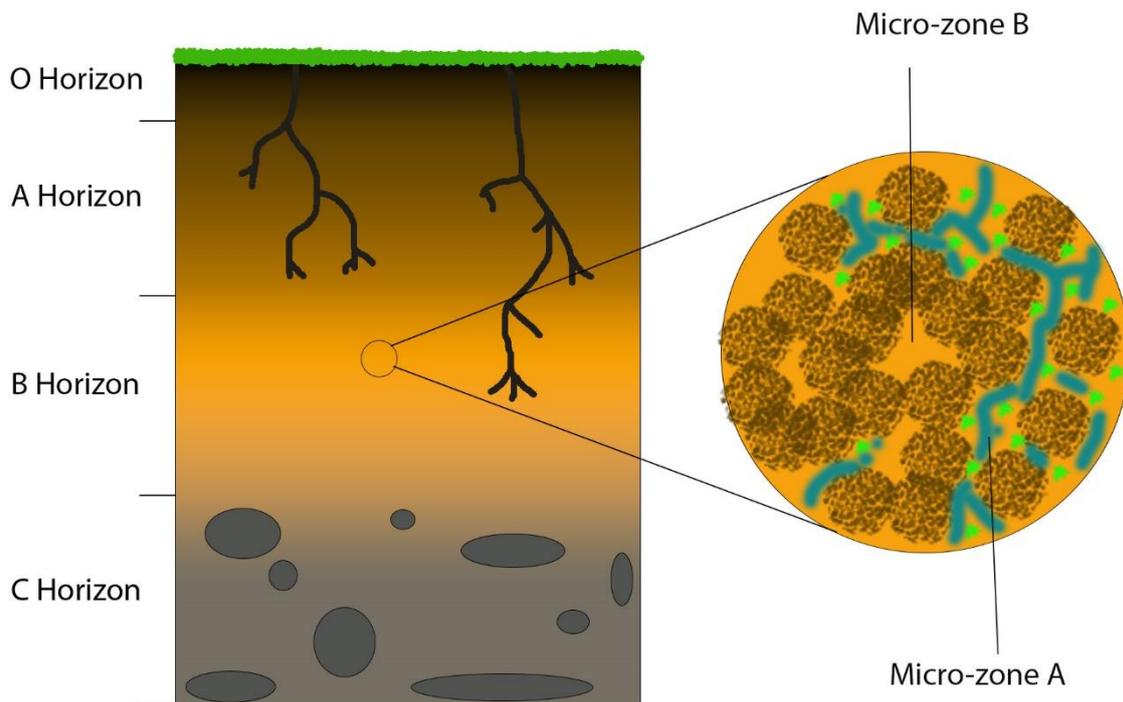
In a recent study, non-oak reads were extracted from whole genome sequences of the oak phyllosphere and queried against bacterial databases. All the AOD bacteria had reads identified in these sequences, regardless of site or tree health status. This supports the conclusions of several other studies that the AOD lesion bacteria are endophytic in oak, but predisposition from decline is required for their pathogenicity (Gathercole *et al.*, 2021). However, some questions remain unanswered. If endophytic, where do these bacteria originate from? Secondly, what role does the microbial community of the soil associated with the roots of oaks play? Thirdly, is there a difference in the composition of the rhizosphere associated with AOD trees compared with their non-symptomatic counterparts? And finally, could this large reservoir of microbial activity be the answer to the first question?

#### 1.10. Bulk Soil:

Soil is the physical support for plants, the functions of both the living and non-living components are essential for their survival. This dynamic substrate is made of different fractions of mineral particles (which differ in size), an array of components, such as exopolysaccharides, cell debris, secretions suspended or dissolved in water, other organic matter and a staggering amount of biota. This reservoir of biotic and abiotic components works together to maintain a continuous flow of energy and nutrients, supporting plant life.

Interaction with soil biota is required for nutrients such as nitrogen and phosphate which bacteria transform from base elements, and minerals bound together from the mucous secretion of bacteria. Microorganisms also perform decomposition, breaking down 95 % of green plant matter. Their importance is undeniable with a hectare of soil potentially containing two tonnes of live weight bacteria (Dawkins and Ellwood, 2015). Tens of thousands of species may compromise that number (though a full understanding of the extent of this diversity is unknown), each with its unique niche which the microenvironments in the soil provide (Simon and Daniel, 2011).

Soil exists in micro-zones over four horizons, in which one micro-zone containing high levels of oxygen, water and nutrients may be millimetres from one containing low levels as demonstrated in Figure 5 (Voroney, 2007). Soil horizons are layers which develop during soil formation and display varying thicknesses and blurred boundaries. The layers develop via alteration in the parent material by accumulation of organic matter from the humus of plant residue, soluble and colloidal organics, inorganics moving down from the surface and the accumulation of both organic and inorganic precipitates. Temperature and water both fluctuate seasonally and based on the soil horizon (Voroney, 2007).



**Figure 5:** The four horizons of soil and an example of micro-zones within these horizons. The O horizon is made up of organic fallen plant material. The A horizon is the topsoil which is especially rich inorganic humus which causes the dark colours. The B horizon (subsoil) is mainly enriched by precipitates, while the C horizon is the least weathered of the base material that forms the soil. Within these horizons are micro-zones which vary by pore size. Micro-zone A demonstrates a zone with large pores that allows air, water (blue) and nutrient (green) movement. Micro-zone B demonstrates small pores which limit movement of air water and nutrients, potentially excluding them.

The formation of the horizons is complex and reliant on chemical, physical and biological processes. Soil microbes cause the majority of the biological transformations, and their presence leads to the accumulation of carbon, nitrogen and other nutrients in stable accessible forms which facilitates the establishment of terrestrial plant life in that medium (Schulz *et al.*, 2013). The true understanding of the role of soil microbiology is complex, due to the  $\mu\text{m}^3$  scale communities function in the soil matrix, where they form biogeochemical interfaces, and shape their environment (Totsche *et al.*, 2010). The addition of abiotic factors further complicates this, with environmental conditions altering phenotypes within minute time frames. Some transcripts are seen to have half times lasting seconds to minutes based on external stimuli, limiting understanding of their roles in terrestrial ecosystems (Sharma *et*

*al.*, 2012). Finally, the timescales in which soil communities develop and stabilise are long with soil formation alone taking centuries.

To summarise, bacteria exist in predictable spatial patterns over wide scales. Distinct compositions of bacterial communities exist in their specific location due to the original base strata, horizon and current environmental conditions. Key functions are performed in their niche and understanding bulk soil microbiological diversity helps us understand how specific members of the rhizosphere community are recruited.

### 1.11. The rhizosphere:

In 1904 Lorenz Hiltner defined the rhizosphere as the region of soil which immediately surrounds the root hairs and surface of the roots, which remains after shaking off loose soil (Hartmann, Rothballer and Schmid, 2008). This is the first region where bacterial communities that play essential roles in soil have a direct impact on the health of plants, both positive and negative (Varnam and Evans, 1995). Staining has revealed a high concentration of root exudates and rhizodeposits, alongside bacteria (Hartmann, Rothballer and Schmid, 2008). Rhizosphere research received little attention for over 50 years before becoming a frontier of research thanks to the increased accessibility of metagenomic and metabolic methods. It has become a focus required for understanding plant health and growth, and a vital player in the next green revolution (Brink, 2016).

#### 1.11.1. Structure:

The rhizosphere encompasses the complete plant root interface and was officially described as the area surrounding the root, which is inhabited by a unique subset of microorganisms, that plants recruit from bulk soil via chemicals released from their roots (Hartmann, Rothballer and Schmid, 2008; Hiltner, L., 1904). Because of the highly variant nature of root systems, the rhizosphere cannot be defined by a set radius, instead it is a radially and longitudinally functioning gradient of chemical, biological and physical properties. This gradient is divided into three zones, based on distance and influence over the root. The ectorrhizosphere is the furthest from the root starting in the bulk soil and extending into the rhizoplane. The rhizoplane is the medial zone seen between the roots and soil, including the root epidermis and mucilage. Finally, the endorrhizosphere is the interior portion of the roots, including the cortex and endodermis, where microbes and cations move into free space

(McNear, 2013). The endorhizosphere is the point where root-associated microbes become endophytic within the plant. The movement from bulk soil into the endorhizosphere is the route which some pathogenic bacteria utilise to colonise plant prior to causing disease, and as such is of key interest.

#### 1.11.2. Colonization of the ectorrhizosphere and rhizoplane

Outside of laboratory or greenhouse conditions, poor colonisation of the rhizosphere by bacteria due to competition from native microbes and predations is frequent, driving the need to understand the mechanisms involved in colonisation (Kumar and Dubey, 2020; Lugtenberg, Kravchenko and Simons, 1999). Root exudates are key, as between 5 - 21 % of carbon and 15 % nitrogen produced in the aerial section of the plant are secreted, causing a 10 to 1000-fold increase in the bacterial concentration from bulk soil to the rhizosphere (Walker *et al.*, 2003; Merbach *et al.*, 1999). Low weight molecular compounds such as sugars, uronic acids, and organic acids are also deposited into the ectorrhizosphere (Haichar *et al.*, 2014; Fan *et al.*, 1997; Vančura, 1964). Exudates act as gradient signals which cause chemoattraction of microorganisms which then utilise the readily available nutrients, encouraging colonisation and multiplication (Lugtenberg and Kamilova, 2009).

Exudate production is specific to cultivar, paternal accession and the growth stage (age) in *Arabidopsis* (Monchgesang *et al.*, 2016). Environmental stress also exhibits a large effect on the exudates deposited into the rhizosphere (Chaparro, Badri and Vivanco, 2014; Haichar *et al.*, 2008). Different exudates influence the colonization process to specifically enhance adaptation, by overcoming nutrient deficiencies/external stresses via members of the rhizosphere (Pii *et al.*, 2015; Lugtenberg, Dekkers and Bloemberg, 2001). This process is not always used for positive recruitment, with some exudates acting as antimicrobials, fungicidal, insecticidal and having nematocidal mediation (Haichar *et al.*, 2014; Compant, Clément and Sessitsch, 2010). Exudate secretion is dynamic based on the plant's requirements, occurring non-homogeneously in concentrated zones where they may also be reabsorbed by the plant (Sasse, Martinoia and Northen, 2018; Compant, Clément and Sessitsch, 2010).

Because rhizosphere colonisation is selective and requires motility for chemotaxis, only 7 % of bulk soil microbes manage to colonise (DeAngelis *et al.*, 2009), shifting the number of taxa from hundreds of thousands identified in the bulk soil to hundreds in the rhizosphere (Sasse,

Martinoia and Northen, 2018). However, most bacteria exhibit motility (Taktikos, Stark and Zaburdaev, 2013), so colonisation is a highly competitive process. Biocontrol secondary metabolites such as siderophores, lytic enzymes and antibiotics amongst other features that control phytopathogen growth are essential when colonising the rhizosphere (Kumar, Dubey and Maheshwari, 2012; Tariq, Yasmin and Hafeez, 2010; Kinsella *et al.*, 2009). Biocontrol elements are often regulated by abiotic and biotic features such as exudates, the presence of pathogens and cell density, as demonstrated by the *GacA/S* system in *P. fluorescens* (Martínez-Granero, Rivilla and Martín, 2006; Raaijmakers, Vlami and de Souza, 2002). Disruption of these genes reduced competitive ability in many bacteria trying to colonise the rhizosphere. Both rhizosphere plant growth promoting bacteria (PGPB) and phytopathogens contain genetic homologs of these genes, implicating essential conserved features for root colonisation (Mark *et al.*, 2005).

Another reduction in the number and composition of bacteria is observed from the rhizosphere soil ( $10^7 - 10^9$  colony forming units (CFU g<sup>-1</sup>), to the rhizoplane ( $10^5 - 10^7$  CFU g<sup>-1</sup>) (Compant, Clément and Sessitsch, 2010; Benizri, Baudoin and Guckert, 2001). Investigation of the rhizoplane is complicated, requiring rhizosphere sterilisation followed by the removal of root surface colonisers which requires harsh chemical treatment and/or sonication (Richter-Heitmann *et al.*, 2016; Buesing and Gessner, 2002). Nevertheless, green fluorescent protein (GFP) and *gusA*-labelled strains in gnotobiotic systems have shown that movement of bacteria from the ectorrhizosphere to the rhizoplane occurs after colonisation of the rhizosphere (Gamalero *et al.*, 2003). Bacterial cells then attach to the root surfaces, after which doublets attach to the rhizodermis, forming a string of bacteria. Finally, colonisation takes place across whole sections of the rhizodermal layer, potentially as microcolonies or biofilms (Compant, Clément and Sessitsch, 2010; Benizri, Baudoin and Guckert, 2001).

#### 1.11.3. Roles of rhizosphere colonisers in oak

The role of rhizosphere colonisers is multi-faceted, with the main effects being nutrient mobilisation, production of phytohormones and phytopathogen antagonism with direct effects varying from increased biomass to plant death (Liang *et al.*, 2014; Blom *et al.*, 2011). The oak microbiome has received attention in recent years indicating the potential role of rhizosphere colonisers in oak health. Tree health, rhizosphere properties and microbiome composition have been shown to shift at sites suffering from AOD. Healthy trees exhibit less

extreme soil conditions and a wealth of PGPB while AOD oaks suffered from acidic soils with distinct microbiomes (Pinho *et al.*, 2020). An assessment of the nitrogen-fixing members of the rhizosphere revealed the association of ammonia-oxidising bacteria with healthy oaks increasing nitrogen content for the alleviation of stress (Scarlett *et al.*, 2021).

AOD-independent studies have shown that oaks utilise rhizospheric PGPB to enhance arial growth, with species such as the holm oak utilising *Arthrobacter* for the degradation of organic polymers. The plant growth-promoting effect of these bacteria has been observed in the recovery of holm oak after being damaged by forest fires (Fernández-González *et al.*, 2017). Abiotic disturbances have also been noted to alter the rhizosphere of cork oak, leading to worse outcomes for those suffering from decline (Maghnia *et al.*, 2019). Clearly, the oak rhizosphere is of key importance, but how does it affect the endophytic composition of the oak microbiome which leads to AOD symptoms?

#### 1.11.4. Colonization of the endorhizosphere and endophytic spaces

The endorhizosphere covers the internal structures of the root, consisting of cortical cells and the endodermis until the vascular tissue (McNear, 2013). Bacteria colonise and proliferate in the endorhizosphere due to the high levels of carbon compounds released from root epidermal and cortical cells (Haichar *et al.*, 2014). The colonisation of the endorhizosphere is how many bacteria become members of the endosphere, as shown through their highly similar taxonomy. Members of the endosphere are mostly non-pathogenic and beneficial to the host, despite their invasive origin (Kandel, Joubert and Doty, 2017; Hinsinger *et al.*, 2009). Penetration of the root is the first step in endophytic colonisation. The process can be passive, with bacteria moving through cracks caused by the emergence of roots at lateral root junctions, damage caused by deleterious microorganisms such as nematodes or at points such as the lenticels which have parenchymal cells with larger intracellular gaps and unsubsided cell walls (Compant, Clément and Sessitsch, 2010; Huang, 1986). Other bacteria secrete PCWDEs allowing them to break through the root surface for endophytic colonisation (Monteiro *et al.*, 2012). *Burkholderia* sp. strain PsJN tagged with GFP has been visualised colonising root surfaces, where it produces endoglucanases and endopolyglucanases. They then colonise the internal root tissue before spreading to the phyllosphere, where they proliferate. Localised host defences such as strengthening of cortical and exodermis cell walls are seen in response to this method of colonisation (Compant *et al.*, 2005).

The endorhizosphere is filled with subpopulations of rhizosphere colonisers, to concentrations of  $10^5 - 10^7$  CFU  $g^{-1}$  (Compant, Clément and Sessitsch, 2010; Hallman, 2001). Endophytic bacteria are thought to represent a subpopulation of those inhabiting the rhizosphere, that are able to enter the plant and adapt to the internal environment through altered metabolism (Bulgarelli *et al.*, 2012; Compant, Clément and Sessitsch, 2010). Inoculation of PGPB around the root system to establish endophytic colonies has been used to confer phenotypic traits including stress tolerance and growth enhancement (Patel and Archana, 2017; Tian *et al.*, 2017). Lavender shoot endophytes show a similar composition of genera with reduced species diversity to those of the root endophytes, implying they are composed of opportunistic members of the endorhizosphere (Pereira *et al.*, 2016). However, in *Populus deltoids* there is little overlap between operational taxonomic units (OTUs) seen in the rhizosphere and endosphere, suggesting that the endosphere offers a unique niche, occupied by distinct assemblages of bacteria rather than random opportunistic members of the rhizosphere (Gottel *et al.*, 2011). As such it seems that while some members of the endophytic bacteria can be tracked through the roots, other species must utilise other routes. Some are inherited as core members of the seed microbiome (Rahman *et al.*, 2018) or enter through leaves, where they remain close to the surface in the phyllosphere (Compant *et al.*, 2021).

#### 1.12. Plant pathogen interactions

Due to the microorganism-rich environment in which plants have evolved and the close symbiotic relationship with many of these microorganisms, plants have had to evolve immunity mechanisms to pathogens which also inhabit these environments. Basal defences are standard features in plants such as cell walls, bark or waxy cuticles and are enough to stop pathogen interaction with unsuitable hosts. However, when interactions between a susceptible host and virulent pathogen occur more specialised defences are required (Soosaar, Burch-Smith and Dinesh-Kumar, 2005).

Innate immunity refers to non-specific proteins and other cellular mechanisms which can identify conserved pathogen-associated molecular patterns (PAMPs) to inactivate them before infection occurs (Alberts *et al.*, 2002). PAMPs are molecular features associated with or produced by pathogens, that are not found within host cells but are recognised within hosts by pathogen recognition receptors (PRR) which are both membrane bound and

cytoplasmic. PAMP recognition causes the activation of PAMP-triggered immunity (PTI), which is the first line of defence (Jones and Dangl, 2006). Many phytopathogens can interfere with PAMP recognition or suppress PTI by effector proteins, and as such plants require specialized defence mechanisms (Chisholm *et al.*, 2006). Pathogens produce general elicitors that cause a defence response in all plants, or specific avirulence (AVR) proteins that only trigger a response if the complimentary resistance (R) gene is present within the host (Hammond-Kosack and Jones, 1996). AVR proteins are most commonly delivered through the T3SS which is coded by the *hrp* gene cluster (He, Huang and Collmer, 1993). Inside the cell these effector proteins suppress the hypersensitive reaction (HR) and cell wall defences, demonstrating their essential role in pathogenicity (Mansfield, 2009).

R gene activation can prevent colonisation, but limited pathogen proliferation may still occur. The HR or programmed cell death, where the infection site becomes separate necrotic tissue limiting nutrient supply and spread of pathogens is the first R gene phenotype (Soosaar, Burch-Smith and Dinesh-Kumar, 2005). Because necrotrophic and hemibiotrophic pathogens can still obtain nutrients (Balint-Kurti, 2019), harmful substances are also released from the vacuole, with phytoalexins reaching inhibitory concentrations for pathogens (Hammond-Kosack and Jones, 1997; Osbourn, 1996). Other responses such as production of reactive oxygen species, lipid peroxidation and systemic acquired resistance (SAR) which can last for weeks are also recorded (Balint-Kurti, 2019; Soosaar, Burch-Smith and Dinesh-Kumar, 2005). If the HR is triggered, then an incompatible plant-pathogen interaction occurs. However, if the innate immune system is overcome and HR is not triggered then the pathogen will successfully invade, colonise and cause disease in a compatible plant-pathogen interaction (Van Der Biezen and Jones, 1998; Hammond-Kosack and Jones, 1997).

As *B. goodwinii* encodes a T3SS and associated avirulence genes (Doonan *et al.*, 2019), could it be an invasive phytopathogen? Do these genes elicit the HR response in oak trees, initiating necrosis and lesion formation, which necrotrophic/hemibiotrophic microbes such as *B. goodwinii* and *G. quercinecans* feed on, expanding the lesion? Or are these microbes new phytopathogens to oaks, and afflicted trees have no corresponding R genes, preventing their initiation of HR, leading to the formation of infection sites, necrosis and lesions?

### 1.13. Aims and objectives:

Overall, this project aimed to investigate the poorly understood role of the microbial community of the soil associated with the roots of oak trees, with and without symptoms of AOD. Which involved finding if the rhizosphere/rhizoplane and under-canopy bulk soil microbial populations differed in symptomatic and non-symptomatic oaks, identify the dominant bacterial genera and species in the soils of AOD and non-symptomatic trees; and investigate if the key AOD lesion bacteria were present to determine if soil is functions as a reservoir for infection. The key functions of the groups of bacteria that differ in the rhizosphere soils of AOD and non-symptomatic trees was also examined.

Objective 1: Identify the best, sampling, storage, an DNA/bacterial extraction method for rhizosphere samples.

Objective 2: Confirm that rhizosphere samples originate from oak.

Objective 3: Screen the samples for AOD bacteria using high resolution melt analysis.

Objective 4: Isolate bacteria from the rhizosphere for the identification of potentially novel species.

Objective 5: Perform taxonomic classification of novel isolates.

Objective 6: Analyse the soil microbiome to determine if there is a correlation between the occurrence of AOD symptoms and the bacteria present in the soil. Short read sequencing is then performed for each barcoded section.

Objective 7: Determine the potential function of differently abundant groups of bacteria specifically associated with diseased/healthy rhizosphere samples.

# Chapter 2. Materials and methods

## Culturing methods:

### 2.1.1. Sample collection

For the collection of rhizosphere soil samples from Hatchlands park a standardised method was taken. Suitable healthy and AOD symptomatic oaks were identified and assigned a number in which 1 – 10 indicated parkland samples and 11 – 20 indicated woodland samples, while odd numbers indicated AOD symptomatic and even numbers indicated healthy oaks. 2-meter perimeters were then marked around trees and holes of ~ 6-inch radius and 1 ft depth were made. Roots and the adhered rhizosphere soil that remained after shaking were then taken and stored in sterile zip-lock bags and returned to the laboratory where they were processing began immediately.

### 2.1.2. Isolation of bacteria from bleeding lesions

To isolate bacteria from bleeding lesions a standardised swabbing method was employed. All samples taken from bleeding lesions on *Tilia* species from Tidworth, Wiltshire and Westonbirt Arboretum, Gloucestershire, UK and *Quercus* species from Hatchlands Park were collected using a sterile swab which was rubbed on bleed spots (both dry and wet). These were suspended in 5 mL of ¼ strength Ringer's solution (Oxoid) once returned to the lab and then plated on Luria-Bertani (LB) agar and incubated anaerobically for 48 h at 35 °C (Kile et al., 2022). This temperature was used due to present conditions of the anaerobic cabinet, which is also used for growth of anaerobic human pathogens.

### 2.1.3. Culturing bacteria from mixed samples

Following isolation cultures were recovered from a range of sources, including a variety of swabs and agar stabs sent by Forest Research and a range of frozen glycerol reference stocks using a standardised method. From the original samples (bacteria inoculated into a slide, glycerol stocks, etc.) a sterile loop full of bacteria or 100 µL of suspension was placed onto LB agar in triplicate. Plates were streaked out for single colonies and then incubated aerobically at 28 °C or anaerobically at 35 °C for 48 hours.

### 2.1.4. Isolation of bacteria in mixed samples

To obtain pure single colonies from environmental samples which could be stored and further investigates later a standard isolation method was used. Agar plates containing mixed colonies had individual colonies picked up with a sterile loop and were quadrant streaked for

single isolates. Plates were then incubated at 28 °C for 48 hours to allow for the growth of single colonies. Pure single colonies were then used to create stocks and suspensions for further use.

#### 2.1.5. Glycerol stocks

Once pure colonies had been obtained from samples, isolated needed to be stored to allow for revival when required. As such glycerol stocks were made by suspending a single colony of bacteria in 200 µl of LB broth (Fisher Scientific, UK) (see Appendix), which was then placed in a shaking incubator at 25 °C (170 RPM) until growth was observed by an increase in turbidity. Following the identification of growth 200 µl of 40 % glycerol was then added, allowing stocks to be stored at -80 °C (Howard, 1956).

#### 2.1.6. Overnight batch cultures of bacterial strains

Liquid cultures of bacteria incubated overnight were routinely used to start experiments and for DNA extraction, due to the ability to adjust the cell density. These overnight batch cultures were made with single colonies placed in 10 ml of LB broth and incubated for a minimum of 12 hours at 25 °C in a shaking incubator at 170 RPM.

#### 2.1.7. Cultivation of bacteria from soil

Bacteria were routinely isolated from rhizosphere soil samples for both the identification of AOD related lesion pathogens and to collect potentially novel isolates. To extract the highest number of bacteria from the soil, harsh physical treatment separating cells from small pores in the matrix was required (Janssen *et al.*, 2002; Buesing and Gessner, 2002). The soil was collected and then sieved in autoclaved and UV-treated 20 µm sieves under aseptic conditions to remove debris and homogenise soil particle size. A 10 g quantity was weighed and suspended in 90 ml of ¼ strength Ringer's solution, creating a 1/10 dilution. A sterile Teflon-coated stirrer was added to the suspension. The suspension was mechanically disrupted at 1,150 RPM for 10 minutes. Further dilutions were then made to 10<sup>-5</sup> and 10<sup>-6</sup> for anaerobic and aerobic culture conditions, respectively. A 100 µL aliquot of the dilutions were then taken and spread plated, after which plates were incubated at 28 °C to allow colonies to develop (Pepper *et al.*, 2015). Anaerobic plates were incubated in a 3.5 L (Thermo-Fisher, UK) anaerobic jar, with a CO<sub>2</sub> gen (Oxoid, UK) packet to generate an anaerobic environment.

#### 2.1.8. Enterobacteriaceae enrichment of bacteria from soil

To specifically isolate members of the current order Enterobacterales from samples, Enterobacteriaceae Enrichment (EE) broth (Sigma) was used. Ten grams of soil was suspended in 95 mL of EE broth and mechanically disrupted at 1,150 RPM with a Teflon-coated bar on a magnetic stirrer for 10 minutes. For other materials such as leaves and acorns that were to be enriched, samples were pre-processed by grinding in an autoclaved, UV treated mortar and pestle. Ground samples were then suspended in EE broth. EE suspensions were placed in a shaking incubator for 48 hours at 28 °C at 190 RPM after which appropriate dilutions were made ( $10^{-5}$  and  $10^{-6}$  for anaerobic and aerobic, respectively due to the lower recovery of colonies under anaerobic conditions) and 100  $\mu$ L was plated on the eosin methylene blue (EMB, Sigma), LB agar, Gifu Anaerobic Medium (GAM, Trafalgar Scientific Ltd) agar and Reasoner's 2 agar (R2A) (Thermo-Fisher) agar and incubated for 48 hours at 28 °C under both aerobic and anaerobic conditions. Anaerobic plates were incubated in a 3.5 L (Fisher, UK) anaerobic jar, with a CO<sub>2</sub> gen (Oxoid, UK) packet to generate an anaerobic environment.

For recovery of more sensitive isolates (namely *Brenneria* and *Lonsdalea*), pre-enrichment in Buffered Peptone Water (BPW) was performed. Samples were disrupted in the same fashion as the EE pathway but suspended in 95 mL of BPW (Oxoid). BPW suspensions were then placed in a shaking incubator for 4 hours at 28 °C at 190 RPM after which 10 mL was transferred into 90 mL of EE broth. At this point, samples in the EE broth followed the same method as previously described, with samples being incubated for 48 hours.

#### 2.1.9. Growth curves

To determine the growth of bacterial isolates optical density (OD) readings at 600 nm of batch cultures were taken at hourly intervals, this data was used to determine appropriate periods of growth to get bacteria to a particular OD. Batch cultures were prepared by centrifuging overnight cultures at 8346 x *g* for 10 minutes to pellet cells, removing the supernatant and resuspending in 10 mL of fresh LB broth. This process was repeated with the final resuspension of the pellet in 5 mL to concentrate bacteria. The OD<sub>600</sub> was measured and the correct volume was added to 129 mL of LB broth to give an initial optical density of 0.05. The new suspension was then placed in a shaking incubator (170 RPM) at 25 °C and removed hourly to have the OD measured. Once the OD had plateaued for 3 hours, measurements

were halted, as bacteria had reached the stationary phase of their growth curve (Maier *et al.*, 2000).

#### 2.1.10. Surface sterilisation of leaves and acorns

To ensure only endophytic bacteria were identified in leaves and acorns while trying to identify the presence of AOD bacteria, surface sterilisation was performed. Samples were submerged in 70 % ethanol for 1 minute followed by 10 % sodium hypochlorite for 1 minute. Samples were then washed twice in sterile distilled water to remove sterilising agents before use.

#### 2.1.11. Spiking of different microcosms with AOD bacteria

To assess the environmental range of the bacteria associated with AOD lesions they were artificially inoculated into different environments and left over a six-week period with weekly recoveries performed. For this work an adapted spiking method from Pettifor *et al.*, (2020) was used to spike single isolates into different microcosms. Bacteria were grown overnight to an OD<sub>600</sub> of 0.5 in 10 ml of LB to ensure they were in the log phase, and therefore in a sufficient state to inoculate into a new medium. Overnight broths were then centrifuged at 8,346 x g to pellet bacteria, the supernatant removed, and the pellet resuspended in ¼ strength Ringer's. This process was repeated to ensure all LB media was removed before final resuspension. The optical density of the pure washed culture was adjusted to an OD of 0.5 by diluting with ¼ strength Ringer's.

For soil, 0.5 mL of pure washed bacterial culture was pipetted into the middle of a falcon tube containing 10 g of soil. Soil microcosms were then vortexed and hand-shaken to ensure the dispersal of bacteria throughout the soil. For leaves, hypodermic needles were used to inject 20 µL of pure-washed bacterial culture into the petiole and midrib at three different points and four points on the blade, near the veins. Acorns had four, 5 µL injections made with a hypodermic needle that was used to break through the pericarp and testa.

All microcosms were stored at 8 °C, the average annual UK forest temperature (The Met Office, 2019), until extraction of bacteria was performed and dilutions were plated on EMB, LB, GAM and R2A agar.

## 2.2. Genotypic Methods:

### 2.2.1. DNA extraction; colony in water

Isolation of genomic DNA for PCR, HRM and sequencing was initially performed by taking a single bacterial colony which was placed in an Eppendorf with 250 µl of molecular grade (DNA/RNA free water, Thermo-Fisher Scientific) and vortexed, cells were then lysed in water via boiling.

### 2.2.2. DNA extraction; alkaline lysis

Most of this work used alkaline lysis for crude extractions of DNA for subsequent PCR, HRM and sequencing methods. For alkaline lysis (Niemann *et al.*, 1997) a single colony of bacteria was taken and placed in a microcentrifuge tube, followed by 20 µl of alkaline lysis buffer, a combination of 0.05 M sodium hydroxide (NaOH) and 0.25 % sodium dodecyl sulfate (SDS) which was filter sterilised before use. The 20 µl suspensions were incubated at 95 °C for 15 minutes, after which they were quickly centrifuged to remove evaporation from the lid. A 180 µl volume of molecular grade water was then added to dilute suspension 10-fold, followed by 5 minutes of centrifugation at 17,000 x g to pellet lysed cells. Isolated DNA was then stored in -20 °C freezers before further use.

### 2.2.3. Soil DNA extraction

To extract DNA from soil for 16S rRNA gene community analysis of the oak rhizosphere a range of commercial kits were tested. The Qiagen Power Soil Pro Kit (Qiagen) was chosen as a suitable kit for the extraction of DNA from soil following optimisation in chapter one. A weight of 250 mg of soil for immediate extraction or up to 1 g of soil stored in 1 mL DNA/RNA shield (Cambridge Bioscience) could be processed by the kit. Extractions were performed as per the manufacturer's protocol with DNA being frozen after for downstream application.

### 2.2.4. DNA extraction from pure cultures

When high volumes of pure DNA were required for experiments such as artificial spiking of soil with DNA, extraction of DNA from liquid cultures was performed. The Easy Powerlyze Kit (Qiagen) was used for the extraction of DNA from pure cultures. Cultures were grown overnight in 10 mL of LB broth in a shaking incubator at 25 °C (170 RPM). Cultures were removed and centrifuged for 10 minutes at 8,346 x g to pellet cells. The supernatant was removed and the pellet was resuspended in 5 mL of ¼ strength Ringer's to concentrate them.

This was then used in the kit as per the manufacturer’s protocol. DNA recovered was stored on ice or frozen at -20 °C for further use.

#### 2.2.5. Root DNA extraction

The Extract ‘n Amp™ Plant PCR Kit (XNAP2; Sigma) was used as a rapid method for the extraction of DNA from fine roots collected with rhizosphere samples to ensure the samples contained oak roots. A fine root was selected from rhizosphere samples and approximately 1.5 cm was cut from the end of the root. This was placed in 100 µL of extraction buffer provided in the kit and ground with a disposable pellete pestle (Sigma) which was washed and autoclaved between uses. The sample was vortexed and placed at 95 °C for 10 minutes. Samples were briefly centrifuged and 100 mL of dilution buffer was added. Extracted DNA was then stored at -20 °C until further use.

#### 2.2.6. Polymerase chain reaction

Polymerase chain reaction (PCR) was used to amplify extracted DNA which was subsequently used for the identification and classification of single isolates. Initially, a mastermix was made of 12.5 µL of 2x Taq PCR Master Mix (QIAGEN, UK) 1 µL of both the appropriate forward and reverse primers at 10 pmol (Table 1) (Eurofins Genomic, UK), and finally 11.5 µL of molecular grade water. The mix was then distributed into 0.2 mL PCR tubes so that each one contained 24 µL/23 µL of PCR mix after which 2 µL of DNA was added to their reaction tubes respectively. Each PCR was performed with the positive control (*Gibbsiella quercinecans* strain Gq4) and negative control (molecular grade H<sub>2</sub>O). Tubes were placed in a PCR machine (Techne, Flexigene) and the PCR amplification was performed using the specific conditions detailed in Table 2.

**Table 1:** List of primers used for PCR amplifications detailing the gene amplified, the primers' official name, the primer sequence and the paper in which the primer sequence was originally published. \* indicates primers that were use for amplification and sequencing.

Gene	Primer name	Sequence (5'->3')	Reference
16S rRNA	PAF	AGA GTT TGA TCC TGG CTC AG	Coenye <i>et al.</i> , 1999

16S rRNA	PHR	AAG GAG GTG ATC CAG CCG CA	Coenye <i>et al.</i> , 1999
<i>gyrB</i>	<i>gyrB</i> 01-F	TAA RTT YGA YGA YAA CTC YTA YAA AGT	Brady <i>et al.</i> , 2008
<i>gyrB</i>	<i>gyrB</i> 02-R	CMC CYT CCA CCA RGT AMA GTT	Brady <i>et al.</i> , 2008
<i>rpoB</i>	CM 7-F	AAC CAG TTC CGC GTT GGC CTG	Brady <i>et al.</i> , 2008
<i>rpoB</i>	CM 31b-R	CCT GAA CAA CAC GCT CGG A	Brady <i>et al.</i> , 2008
<i>atpD</i>	<i>atpD</i> 01-F	RTA ATY GGM GCS GTR GTN GAY GT	Brady <i>et al.</i> , 2008
<i>atpD</i>	<i>atpD</i> 02-R	TCA TCC GCM GGW ACR TAW AYN GCC TG	Brady <i>et al.</i> , 2008
<i>infB</i>	<i>infB</i> 01-F	ATY ATG GGH CAY GTH GAY CA	Brady <i>et al.</i> , 2008
<i>infB</i>	<i>infB</i> 02-R	ACK GAG TAR TAA CGC AGA TCC A	Brady <i>et al.</i> , 2008
<i>fusA</i>	<i>fusA</i> 3*	CAT CGG TAT CAG TGC KCA CAT CGA	(Delétoile <i>et al.</i> , 2009)
<i>fusA</i>	<i>fusA</i> 4*	CAG CAT CGC CTG AAC RCC TTT GTT	Delétoile <i>et al.</i> , 2009

<i>leuS</i>	leuS3*	CAG ACC GTG CTG GCC AAC GAR CAR GT	Delétoile <i>et al.</i> , 2009
<i>leuS</i>	leuS4*	CGG CGC GCC CCA RTA RCG CT	Delétoile <i>et al.</i> , 2009
<i>pyrG</i>	pyrG3*	GGG GTC GTA TCC TCT CTG GGT AAA GG	Delétoile <i>et al.</i> , 2009
<i>pyrG</i>	pyrG4*	GGA ACG GCA GGG ATT CGA TAT CNC CKA	Delétoile <i>et al.</i> , 2009
<i>rpoB</i>	Vic3*	GGC GAA ATG GCW GAG AAC CA	Delétoile <i>et al.</i> , 2009
<i>rpoB</i>	Vic2*	GAG TCT TCG AAG TTG TAA CC	Delétoile <i>et al.</i> , 2009

**Table 2:** PCR amplification program by gene and primer type. 1) Initial denaturation. 2) Denaturation. 3) Annealing temperature of the primer. 4) Temperature required for extension. 5) The temperature of the final extension. For *FusA*, *LeuS* and *PyrG* 5) Second denaturation. 6) Second annealing temperature 7) Second temperature required for extension. 8) The temperature of the final extension. <sup>a</sup> indicated the template corresponds to the internal portion of the PCR product used for sequence comparison.

Gene/primer	Number of Cycles	1	2	3	4	5	6	7	8	Template size (bp) <sup>a</sup>	Reference
16S rRNA PAF/PHR	30	94°, 5 min	94°, 1 min	55°, 1 min	72°, 1.5 min	72°, 10 min	N/A	N/A	N/A	1533	Coenye <i>et al.</i> , 1999
<i>rpoB</i> CM 7-F/ CM 31b-R	40	94°, 1 min	94°, 10 sec	58°, 20 sec	72°, 50 sec	72°, 5 min	N/A	N/A	N/A	637	Brady <i>et al.</i> , 2008
<i>gyrB</i> 01-F/ 02-R	30	94°, 5 min	94°, 1 min	46°, 1 min	72°, 1.5 min	72°, 10 min	N/A	N/A	N/A	742	Brady <i>et al.</i> , 2008
<i>atpD</i> 01- F/02-R	40	94°, 1 min	94°, 10 sec	58°, 20 sec	72°, 50 sec	72°, 5 min	N/A	N/A	N/A	657	Brady <i>et al.</i> , 2008

<i>infB</i> 01-F/02-R	40	94°, 1 min	94°, 10 sec	58°, 20 sec	72°, 50 sec	72°, 5 min	N/A	N/A	N/A	615	Brady <i>et al.</i> , 2008
<i>rpoB</i>	2-4 = 30 cycles	94°, 4 mins	94°, 30 sec	50°, 30 sec	72°, 30 sec	72°, 5 mins	N/A	N/A	N/A	501	Delétoile <i>et al.</i> , 2009
<i>fusA</i> <i>fusA3/A4</i>	2-4 = 10 cycles  6-7 = 21 cycles	94° 2 min	94°, 1 min	60°, 1 min	72°, 1 min	94°, 1 min	50°, 1 min	72°, 1 min	72°, 5 mins	633	Delétoile <i>et al.</i> , 2009
<i>leuS</i> <i>leuS3/S4</i>	2-4 = 10 cycles  6-7 = 21 cycles	94° 2 min	94°, 1 min	60°, 1 min	72°, 1 min	94°, 1 min	50°, 1 min	72°, 1 min	72°, 5 mins	642	Delétoile <i>et al.</i> , 2009
<i>pyrG</i> <i>pyrG3/G4</i>	2-4 = 10 cycles  6-7 = 21 cycles	94°, 2 min	94°, 1 min	60°, 1 min	72°, 1 min	94°, 1 min	50°, 1 min	72°, 1 min	72°, 5 mins	306	Delétoile <i>et al.</i> , 2009

#### 2.2.7. Agarose Gel electrophoresis

To confirm the amplification of DNA following PCR reactions, agarose gel electrophoresis was used to visualise and analyse samples. Gels were made to a 1 % w/v concentration of agarose powder (Fisher Scientific, UK) using 1x TAE buffer (Appendix), followed by microwaving (Samsung) at 1 minute per 50 mL at 850 W to dissolve agarose. SYBR safe (Thermo Fisher Scientific) or GelGreen (Cambridge Bioscience) diluted 1:10 with TAE were added as intercalating gel stains. To each well, a mix of 2  $\mu$ L of DNA loading buffer (Bioline, UK) to 5 $\mu$ L of amplified DNA was loaded. The appropriate Hyperladder (Bioline, UK) (100 bp-1 Kb) was added as a size marker for PCR products. Gels were run for an appropriate time based on the volume of gel, the size of PCR products and the resolution required, with 50 mL gels run at 8 V/cm for 40 minutes. After which gels were visualised (Syngene U:Genius).

#### 2.2.8. ExoSap

For the removal of excess nucleotides PCR products were cleaned prior to sequencing, Exonuclease 1 was used to digest any remaining primers and Shrimp Alkaline Phosphatase (SAP) removed leftover deoxyribonucleotide triphosphates (dNTPs). An individual reaction was made up of 0.025  $\mu$ L of Exonuclease, 0.25  $\mu$ L of SAP (Thermo Fisher, UK) and 9.725  $\mu$ L of molecular-grade water. Due to the low volumes used a minimum of ten reactions was made. 10  $\mu$ L was added to 20  $\mu$ L of remaining PCR products, giving a 30  $\mu$ L reaction volume. Samples were incubated at 37 °C for 30 minutes followed by 95 °C for 5 minutes in a PCR machine and then stored at -20 °C.

#### 2.2.9. Sequencing

All sequencing of amplified PCR products was performed by Eurofins Genomics (Ebersberg, Germany) using the Mix2Seq kits per the manufacturer's instruction with user-supplied sequencing primers listed in Table 3. The sequencing primers for *fusA*, *leuS*, *pyrG* and *rpoB* are the same as the amplification primers from Table 2.

**Table 3:** List of primers used for sequencing by Eurofins Genomics. Gene sequenced is subdivided by the forward and reverse sequences, indicated by F and R respectively. 16S rRNA primers also use '\*' to indicate a forward primer as in the original paper by Coenye *et al.*

Gene sequenced	Primer name	Sequence (5'->3')	Reference
16S rRNA F	16F358/*Gamma	CTC CTA CGG GAG GCA GCA GT	Coenye <i>et al.</i> , 1999
16S rRNA F	16F536/*PD	CAG CAG CCG CGG TAA TAC	Coenye <i>et al.</i> , 1999
16S rRNA F	16F926/*O	AAC TCA AAG GAA TTG ACG G	Coenye <i>et al.</i> , 1999
16S rRNA F	161112/*3	AGT CCC GCA ACG AGC GCA AC	Coenye <i>et al.</i> , 1999
16S rRNA F	16F1241/*R	GCT ACA CAC GTG CTA CAA TG ACT	Coenye <i>et al.</i> , 1999
16S rRNA R	16R339/Gamma	ACT GCT GCC TCC CGT AGG AG	Coenye <i>et al.</i> , 1999
16S rRNA R	16R519/PD	GTA TTA CCG CGG CTG CTG	Coenye <i>et al.</i> , 1999
16S rRNA R	16R1093/3	GTT GCG CTC GTT GCG GGA CT	Coenye <i>et al.</i> , 1999
<i>rpoB</i> F	<i>rpoB</i> CM81-F	CAG TTC CGC GTT GGC CTG	Brady <i>et al.</i> , 2008
<i>rpoB</i> F	<i>rpoB</i> CM81b-F	TGA TCA ACG CCA AGC C	Brady <i>et al.</i> , 2008
<i>rpoB</i> R	<i>rpoB</i> CM32b-R	CGG ACC GGC CTG ACG TTG CAT	Brady <i>et al.</i> , 2008
<i>gyrB</i> F	<i>gyrB</i> 07-F	GTV CGT TTC TGG CCV AG	Brady <i>et al.</i> , 2008
<i>gyrB</i> R	<i>gyrB</i> 08-R	CTT TAC GRC GKG TCA TWT CAC	Brady <i>et al.</i> , 2008
<i>atpD</i> F	<i>atpD</i> 03-F	TGC TGG AAG TKC AGC ARC AG	Brady <i>et al.</i> , 2008

<i>atpD</i> R	atpD 04-R	CCM AGY ART GCG GAT ACT TC ACG	Brady <i>et al.</i> , 2008
<i>infB</i> F	infB 03-F	ACG GBA TGA TYA CST TCC TGG	Brady <i>et al.</i> , 2008
<i>infB</i> R	infB 04-R	AGY TTA GAT TTC TGC TGA CG	Brady <i>et al.</i> , 2008

#### 2.2.10. NanoDrop DNA quantification

To check the purity and concentration (ng/ $\mu$ L) of DNA a NanoDrop 1000 spectrophotometer (Thermo-Fisher, UK) was used. The DNA absorbance ratio at both A260/A280 and A260/A230 was assessed, with measurements below 1.8 indicating protein contamination or the presence of other contaminants respectively for each absorbance.

#### 2.2.11. Broad range Qubit DNA quantification

To accurately check the concentration of extracted DNA, Qubit broad range DNA selective dye fluorescence (Thermo-Fisher, UK) was used to measure the concentration of DNA (ng/ $\mu$ L) between 4-2000  $\mu$ L. A Qubit working solution was made by diluting 1  $\mu$ L of Qubit reagent in 199  $\mu$ L of Qubit buffer for each individual reaction. For the standards, 190  $\mu$ L of the working solution was added to 10  $\mu$ L of both the lower and upper standards, which when read in the Qubit 3 Fluorometer set the range. 199  $\mu$ L of the working solution was added to 1  $\mu$ L of DNA extract and then read in the Qubit 3 Fluorometer giving a concentration for the original sample. The method tolerates salts, free nucleotides, solvents, detergents, and proteins giving accurate readings of the concentration of DNA.

#### 2.2.12. Analysis of sequencing products

Sequencing results from Eurofins were first manually checked for quality to ensure accuracy, then forward and reverse contigs assembled using the CAP3 program via UGENE V 38.1. exported FASTA alignment files were then run through Nucleotide BLAST against the NCBI database (Altschul *et al.*, 1990) where a list of the most closely related taxa was generated. The highest % identity sequence was accepted as the most likely identification. 16S rRNA gene sequences were also run through the EzBioCloud server (Yoon *et al.*, 2017). Using information from both, the highest % similarity match was chosen.

### 2.2.13. High-Resolution Melt Analysis

High-resolution melt (HRM) analysis was performed on extracted DNA to rapidly identify the four bacteria most commonly associated with AOD lesions from a range of samples. Protocol parameters and primers are listed in Tables 4 and 5. Primers for the multiplex analysis were added in a mix with 336  $\mu\text{L}$  of molecular grade water and 28  $\mu\text{L}$  of each primer to give a working volume of 560  $\mu\text{L}$ , with each primer at a 5  $\mu\text{M}$  working concentration. Other protocols used individual primers at a 10  $\mu\text{M}$  working concentration. For each reaction 7.5  $\mu\text{L}$  of sensiFAST HRM kit Master Mix (Bioline, UK), 1  $\mu\text{L}$  of both forward and reverse primers (or 2  $\mu\text{L}$  of the primer mix) and 4.5  $\mu\text{L}$  of molecular grade water per 1  $\mu\text{L}$  of DNA giving a final reaction volume of 15  $\mu\text{L}$ . Samples were then placed in a Rotorgene real-time PCR machine (Qiagen) and the appropriate programme ran, shown in Table 5, generating species-specific melt curves on the identification of AOD lesion bacteria.

**Table 4:** Primer sequences used in HRM analysis for AOD bacteria and the reference for their origin.

Primer	Sequence (5'->3')	Reference
Bgi2F	CGTCAAAC TATTTGCTTCCACCCATC	Bueno-Gonzales, 2022
Bgi2R	CGGTATGGGTCTGGGACATTTG	Bueno-Gonzales, 2022
Bgi3F	CATCGCGTCCAGCGTCTG	Bueno-Gonzales, 2022
Bgi3R	GCCTATTGCGTGAACGAACTGGATAG	Bueno-Gonzales, 2022
Gqi3F	GCATACGCCTGGTACAGCGC	Bueno-Gonzales, 2022
Gqi3R	CCTTGGCGGGACAGTCTTGC	Bueno-Gonzales, 2022

Rvii1F	GCATCTCGCAGATCGCTGAAAC	Bueno- Gonzales, 2022
Rvii1R	TGGAAGCGGCGGCTGAC	Bueno- Gonzales, 2022
Lbi2F	GGAATCGCTTTACCGTCGCTATTG	Bueno- Gonzales, 2022
Lbi2R	CAAGGTGGTGATGGTGGTCGATC	Bueno- Gonzales, 2022
Gq6Bf	GGCAACCCATCGACATGAA	Brady <i>et al.</i> , 2016

**Table 5:** High-resolution melt conditions used for the identification of the members of the AOD lesion microbiome including primers, their concentrations and PCR protocol. 1) The hold time. 2) The number of cycles. 3) The denaturation temperature. 4) The annealing temperature. 5) the elongation temperature. 6) the HRM ramp. 7) The time for completion of the PCR.

Protocol Name	Primers	Primers conc	1	2	3	4	5	6	7
Multiplex HRM*	<b>B2-G3-R1-L2</b> (Bqi2F-Bqi2R- Gqi3F-Gqi3R- Rvii1F-Rvii1R- Lbi2F-Lbi2R)	10µM	95°C, 5 mins	30	95°C, 5 secs	75°C, 7 secs	73-93°C	0.5°C	37 mins
<i>Brenneria</i> test	<b>B3</b> (Bqi3F-Bqi3R)	10µM	95°C, 5mins	30	95°C, 5 secs	65°C, 7 secs	78-91°C	0.3°C	44 mins
<i>Rahnella</i> test Differentiates Rvi from the rest	<b>R1</b> (Rvii1F-Rvii1R)	10µM	95°C, 5mins	30	95°C, 5 secs	68°C, 7 secs	70-99°C	0.5°C	45 mins

#### 2.2.14. Spiking soil microcosms with AOD bacterial DNA

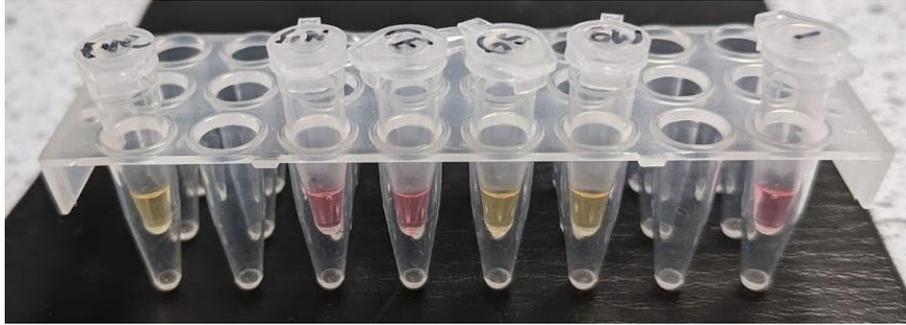
To check if DNA extracted from soil could be used for the detection of the AOD pathogenic bacteria, their pure DNA was spiked into soil in artificially high volumes followed by subsequent extraction and HRM analysis. Bacterial DNA was isolated using the Easy Powerlyze Kit (QIAGEN, UK) and checked for purity (section 2.8) and identities were confirmed using the HRM multiplex method. 50 µL of DNA was spiked into 250 mg microcosms of soil and then incubated at 8 °C. These 250 mg microcosms then entered the soil DNA extraction protocol (2.4) and the pure eluted DNA from the soil was assessed using HRM multiplex analysis, with the isolated DNA being used as positive controls.

#### 2.2.15. Loop-mediated isothermal amplification (LAMP)

LAMP reactions were used to confirm the presence of oak roots in rhizosphere samples. Each 25 µL LAMP reaction contained 1X WarmStart® Colorimetric LAMP Master Mix (New England Biolabs), 1X LAMP primer stock (1.6 µM FIP, 1.6 µM BIP, 0.2 µM F3 and 0.2 µM B3) (Table 6) and 1 µL of extracted subterranean root DNA (section 2.5) diluted two-fold. Samples were incubated at 65 °C for 60 minutes in a PCR machine (TECHNE) and reactions were subsequently assessed for colour change. Lamp reactions were positive when a colour change from pink to yellow was seen. Negative reactions had no colour change due to lack of amplification and as such remained pink (Figure 6).

**Table 6:** The gene amplified in the LAMP for the identification of oak material, the primers used to disrupt and amplify that gene and the sequence of the primers used.

Gene	Primer name	Sequence (5'->3')
Actin	Actin F3	AGTTCTTTAAGGACGCCAC
	Actin B3	CCTTGAGGTACTTGCCATG
	Actin FIP	TACTTTTTCTTCGTCGTCTTCAGCAGCTGTAAGACTCAGGAGT
	Actin BIP	CTCAAACAGAGCGTTTTGGACCGCAGCTTTTCAAGCGGATA



**Figure 6:** LAMP reaction for Hatchlands Park tree 6 rhizosphere samples from each 4-cardinal point.

### 2.3. Taxonomic Methods

#### 2.3.1. Phylogenetic analysis:

All additional sequences for the closest phylogenetic neighbours were downloaded from Genbank via BLAST (Benson *et al.*, 2013). Sequences were aligned and trimmed in MEGA X v10.0 (Tamura, Stecher and Kumar, 2021) to the length listed in Table 2. Both the housekeeping genes and the 16S rRNA gene contigs were aligned using contig assembly program 3 (CAP3) in UGENE V 38.1 (Okonechnikov *et al.*, 2012) to ensure coverage in both directions. Concatenated datasets of the four housekeeping genes were made and Smart Model Selection (SMS) (Lefort, Longueville and Gascuel, 2017) was performed for both datasets on the online PhyML server (Guindon and Gascuel, 2003). The output was then used to inform the maximum likelihood phylogenetic analysis used in MEGA. The reliability of the clades generated in the phylogenetic tree was tested through 1,000 bootstrap replications. Finally, the EzBioCloud server was used to calculate the 16S rRNA gene pairwise similarity of the proposed type strains (Yoon *et al.*, 2017).

#### 2.3.2. DNA Fingerprinting:

To ensure that novel species were not clonal the genetic diversity of all novel strains included in this study was assessed using either BOX or ERIC fingerprinting method, with the BOX A1R or ERIC 1 and 2 primers and protocol shown in Table 7 (Versalovic *et al.*, 1994; De Bruijn, 1992). The amplified products were separated in 1.5 % agarose at 50 V for 3 h.

**Table 7:** The fingerprinting methods used for the identification of clonal isolates. All methods are broken down by the primer used, primer sequence and protocol along with their original publication.

Fingerprinting method	Primer name	Sequence (5'->3')	Protocol	Reference
ERIC	ERIC1R ERIC2	ATGTAAGCTCCTGGGGATTAC AAGTAAGTGACTGGGGTGAGCG	95° 7mins, (94°C 1 min, 52°C for 1 min, and 65°C 8 min) 30 cycles, 65° 16 min	(De Bruijn, 1992)
BOX	BOX A1R	CTACGGCAAGGCGACGCTGACG	95° 7mins, (90°, 30 sec, 95° 1 min, 58° 1 min, 65° 8 min) 30 cycles, 65° 16 min	(Versalovic <i>et al.</i> , 1994)

### 2.3.3. Genomic features

To understand how the novel species relate to each other and the type species of the genus at the genomic level, representative strains were chosen for whole genome sequencing. DNA was extracted by cell lysis and purified on Solid Phase Reversible Immobilisation beads (SPRI) followed by sequencing on the Illumina HiSeq platform by Microbes-NG (Birmingham, UK) using Nextera library preparation kit, with a read length of 2x250 bp paired end reads. Trimmomatic 3.0 was used to trim adapters with a sliding window quality cut-off of Q15 from reads (Bolger, Lohse and Usadel, 2014). SPAdes 3.11.1 was used for the *de novo* assembly, while Prokka 1.11 was used to annotate the assembled contigs (Seemann, 2014; Nurk *et al.*, 2013). Contamination of the whole genome sequences was checked by aligning the 16S rRNA gene sequences obtained via Sanger sequencing to the whole genome sequences in Codon Code version 10.0.2.

### 2.3.4. Phylogenomic analysis

The phylogenomic distance between strains was calculated through pairwise comparisons of genomes using Genome Blast Distance Phylogeny (GBDP) with the Type (Strain) Genome

Server (TYGS) (Meier-Kolthoff and Goker, 2019). The intergenomic distance between a number of reference genomes was calculated using 100 replications of the distance formula  $d_5$  with the algorithm 'trimming' applied (Meier-Kolthoff *et al.*, 2013). The calculated  $d_5$  distance was then used to draw a genome caption tree with scaled branch lengths using FastME 2.1.6.1 (Lefort, Desper and Gascuel, 2015). Subtree Pruning and Regrafting (SPR) was applied to the dataset, before being rooted at the midpoint (Farris, 1972).

#### 2.3.5. Whole genome similarity analysis:

To further understand the relationships between the potential novel species and their closest phylogenomic neighbours, whole genome comparisons of the strains were made using *in silico* DNA – DNA hybridisation (*isDDH*), average nucleotide identity (ANI) and average amino identity (AAI). *isDDH* results were calculated using the Genome-to-Genome Distance calculator, which expresses  $d_5$  values with a cut-off point <70 % indicating a different species (Goris *et al.*, 2007). ANI values were calculated using FastANI (Jain *et al.*, 2018) while AAI was calculated through the Genome-based distance matrix calculator from Kostas lab (Rodriguez-R and Konstantinidis, 2016). In the case of *Scandinavium* where the use of FastANI was disputed by reviewers of the submitted manuscript, the JSpecies server was used to calculate ANIb values (Richter *et al.*, 2016).

#### 2.3.6. Light microscopy

CellSens Version 1.11 imaging software coupled with an Olympus SC180 (Olympus Life Science, Tokyo, Japan) microscope was used for all cell size, morphology and motility assessments.

#### 2.3.7. Transmission electron microscopy

Transmission electron microscopy (FEI Tecnai 12 120kV BioTwin Spirit TEM) was used to assess flagella arrangements for negatively stained novel strains. Negative stains were made by floating grids on a mid-log phase bacterial suspension for 2 minutes, followed by triplicate washing, floating grids in a 3 % w/v uranyl acetate suspension for 30 seconds, another triplicate wash step after which excess liquid was wicked away and grids left to air dry.

#### 2.3.8. Cell physiology

Growth at 4, 10, 25, 28, 30, 37, and 41 °C was assessed in triplicate on tryptone soy agar (TSA, Sigma) to find a suitable temperature range for growth. Colony morphology was determined on TSA for all genera excluding *Scandinavium* which was assessed on Colombia Blood Agar (CBA, Oxoid) after 24 hours incubation at 30 °C as per the type species description (Marathe *et al.*, 2019).

All strains including reference strains were tested in triplicate for both pH and salt tolerance by inoculation of broths with mid-log range overnight cultures that were incubated for 24 hours shaking at 180 RPM in a 37 °C growth cabinet. The pH tolerance was tested from 4 – 10 in increasing increments of 1 pH unit in tryptone soy broth (TSB, Sigma) after altering the original pH using sodium acetate/acetic acid and carbonate/bicarbonate buffers. Saline-free nutrient broth (3 g l<sup>-1</sup> beef extract, 5 g l<sup>-1</sup>) was used to assess the salt tolerance, by adjusting the salt concentration from 1 – 7 % w/v with the addition of 1 % w/v NaCl.

#### 2.3.9. Antibiotic resistance

Antibiotic resistance of novel species was tested against ampicillin, chloramphenicol, colistin, gentamicin, kanamycin, penicillin G, penicillin V, streptomycin, tetracycline, cefoxamine ciproflaxin and tetracycline depending on the genus. Mid-log range overnight cultures were used to make bacterial lawns on TSA after which six antibiotic discs were applied equidistant to each other using a disc dispenser (Oxoid). Plates were incubated at 30 °C for 24 hours and the zone of clearance around the antibiotic disc was assessed. Growth over the antibiotic disk indicated resistance to the antibiotic while any zone of clearance surrounding the disk indicated sensitivity to the antibiotic.

#### 2.3.10. Phenotypic tests

Phenotypic testing was performed using a range of commercial kits including API 20 E and 50 CHB/E (bioMérieux), GEN III GN/GP assays (Biolog) and ID 32 (bioMérieux). All commercial assays were conducted according to the manufacturer's instructions. API 20 E and ID 32 E galleries were scored after 24 hours incubation at 37 °C. API 50 CHB/E galleries and GEN III plates were incubated at 30 °C after which both were scored twice, at 24 and 48 hours, and 16 hours and 24 hours, respectively. Production of bubbles in 3 % v/v H<sub>2</sub>O<sub>2</sub> was assessed for

catalase activity and Kovács reagent (1 % tetra-methyl-p-phenylenediamine dihydrochloride) for oxidase activity was also performed on all strains.

#### 2.3.11. Fatty Acid Methyl Ester analysis

FERA Science Ltd performed Fatty Acid Methyl Ester (FAME) analysis on a range of strains from each proposed novel species. The Sherlock Microbial Identification System Version 6.4 protocol (MIDI Inc.) was followed after strains were grown for 24 hours on TSA at 30 °C. Results were referenced against the RTSBA6 6.21 comparisons library.

#### 2.3.12. Virulence feature

To identify genes that could promote pathogenicity traits, whole genome sequences were processed using the prokaryotic genome annotation pipeline (PGAP) during GenBank submission (Tatusova et al., 2016), and queried against several databases. DIAMOND v2.0.11.149 (Buchfink et al., 2021) was used to query annotated genomes against the Virulence Factor Database (VFDB; Liu et al., 2022), accessed 26 July 2022, via the Blast P command. To identify high sequence identity alignments between the genomes and the VFDB, a query cut-off of 97 % coverage for each alignment and a percentage identity equal or greater to than 50 were used (Doonan et al., 2019). For the identification of plant cell wall degrading enzymes and other virulence factors, the KEGG orthology online search tool was used (Aramaki et al., 2020). OrthoFinder was used to determine the conservation of virulence genes within the genomes by homology searching using ortholog on the subset of genes identified from the VFDB (Emms and Kelly, 2019). Further investigation of the Type VI Secretion System (T6SS) was performed using the SecReT6 v3 database, which was last updated on 15 November 2021 (Li et al., 2015). The T6SS gene cluster protein database was downloaded, and annotated genomes were queried using DIAMOND with the same query parameters as used for virulence factor identification.

#### 2.3.13. PGP features

To assess how novel isolates interacted with plants an online database comparison was utilised. The protein annotations produced from PGAP (Tatusova *et al.*, 2016) were queried against the PLant-associated BActeria web resource (PLaBAse) database using the DIAMOND MEGAN pipeline (Bağcı, Patz and Huson, 2021). First, the PLaBAse PGPT-db from 01/02/2022 was downloaded and used to build a protein database in DIAMOND v2.0.11.149 (Buchfink,

Reuter and Drost, 2021). Each annotated protein file was compared to the database using the BlastP command. To identify high sequence identity alignments between the genomes and the PGPT-db, a query cut-off of 97 % coverage for each alignment and percentage identity equal or greater to than 50 were used. These cut-offs were originally designed for high sequence identity alignments of virulence genes against virulence factors within the same pipeline (Doonan *et al.*, 2019). The alignments output was then entered into the MEGAN pipeline and mapped against the corresponding mgPGPT-mapping-db in MEGAN version 6.24.0. community edition (Bağcı, Patz and Huson, 2021; Huson *et al.*, 2016).

Krona plots were created to visualise the plant growth promoting trait (PGPT) genes identified as groupings defined by their interaction with plants (direct/indirect) and further specific roles (Ondov, Bergman and Phillippy, 2011). The annotated protein sequences were uploaded to the PGPT-pred online tool (available <https://plabase.informatik.uni-tuebingen.de/pb/form.php?var=PGPT-Pred>) and queried against the BlastP+HMMER Aligner/Mapper. Finally, to determine if novel isolates are plant-associated bacteria the PIFAR-BASE was used to identify 'plant bacterial only interaction factors' from the annotated protein files for each isolate using the BlastP+HMMER Aligner/Mapper.

## 2.4. Bioinformatics Methods

### 2.4.1. Loop genomics sequencing

To assess the community composition of the rhizosphere samples to identify structural differences between the bacteria in the healthy and AOD symptomatic rhizosphere single gene community analysis was utilised. Synthetic long read 16S rRNA gene sequencing was performed by Loop Genomics (Element biosciences). 50 µL of extracted DNA were shipped on dry ice to Loop where they underwent QC, dilution and enrichment prior to sequencing. Samples are then processed in the Loop genomics long read workflow. First each sample is exposed to millions of unique barcodes, with a singular barcode attaching to each strand of DNA. The DNA and barcode are then amplified by PCR using universal primers. Each strand of DNA then has the unique barcode randomly distributed throughout the sequence. Short read sequences with the same barcode are then reassembled using linked read *de novo* assembly to form full length molecules.

#### 2.4.2. QIIME OTU identification

To filter and identify unique sequencing variants which would later be assigned to species level identifications OTUs were first identified in sequencing outputs. Fastq files were uploaded to the EzBioCloud Microbiome Taxonomic profile (MTP) pipeline which utilises QIIME 2 for quality filtering and trimming of sequences. Pre-filtering of sequencing data was used to remove low quality, non-target and chimeric amplicons to leave total valid reads. OTU alignment to the EzBioCloud database was then performed. All steps were performed with the standard recommended parameters (Bolyen *et al.*, 2019; Yoon *et al.*, 2017).

#### 2.4.3. Krona plots

To visualise the community structure of rhizosphere samples interactive krona plots were made which assign colours based on hierarchy. Output OTU files were plotted using the `ktImportText` command (<https://github.com/marbl/Krona/wiki/Importing-text-and-XML-data>) (Ondov, Bergman and Phillippy, 2011).

#### 2.4.4. Diversity statistics.

To identify differences in the abundance, diversity, and composition of rhizosphere samples a range of different diversity statistics were tested on samples OTU outputs. All diversity statistics were initially calculated using Phyloseq in R using the command `plot_richness` (McMurdie and Holmes, 2013). Additional statistics were also calculated using the EzBioCloud Microbiome Taxonomic Profiling (MTP) 16S rRNA gene-based online pipeline which utilises the MOTHUR pipeline (Yoon *et al.*, 2017; Schloss *et al.*, 2009). Beta-diversity for grouping levels based on disease profile and location were visualised using the Bray-Curtis metric and 3D principal coordinate analysis at the species level through EzBioCloudMTP pipeline.

#### 2.4.5. PERMANOVA

To identify differences in the beta diversity of samples Pairwise PERMANOVA analysis was performed in the EzBioCloud MTP pipeline using the QIIME 2 pipeline. 999 permutations were used for each comparison with P-values, q-values and PSuedo-F values generated for each pair. Significance was set to  $P < 0.05$ .

#### 2.4.6. Taxonomic biomarkers and functional biomarkers

To identify statistically significant differences in specific bacteria at different taxonomic levels and protein pathways which are associated with those species taxonomic and functional

biomarkers were utilised. All biomarkers were identified by making microbiome taxonomic profile sets using grouped levels which were then processed in the Comparative MTP analyser on the EzBioCloudserver. The taxonomic biomarkers utilised LEfSe analysis in which Kruskal Wallis test is applied to the full datasets to look for differentially distributed classes within the data. Identifications that are differentially distributed then undergo pairwise comparisons under by Wilcoxon test. This is then used to build a linear discriminant analysis (LDA) effect size (LEfSe) model. This generates a list of taxa which are discriminative in respect to their sample of origin (Segata *et al.*, 2011). The same approach is taken to the functional biomarkers, however after the LEfSe analysis is performed the differentially abundant 16S rRNA genes are entered into the phylogenetic investigation of communities by reconstruction of unobserved states (PIECRUST) pipeline. Here 16S rRNA genes are used to predict metagenomes which then cluster KEGG protein sequences based on gene families, thus giving a functional profile for the dataset (Langille *et al.*, 2013).

# Chapter 3. Optimisation of Bacterial and DNA isolation from Soil and Sample Collection.

### 3.2. Introduction:

Soil is a complex environment, which is thought to be the largest reservoir for bacteria on earth, with  $2.6 \times 10^{29}$  predicted prokaryotic cells globally, consisting of  $4 \times 10^6$  different taxa (Long and Or, 2005). Ribosomal intergenic spacer analysis (RISA) fingerprinting results conservatively predict that 2,000 - 18,000 prokaryotic genomes may be present in one gram of soil (Xu *et al.*, 2014; Delmont *et al.*, 2011). This high diversity is attributed to the rate of mutation leading to a rate of speciation faster than the rate of extinction (Torsvik, Øvreås and Thingstad, 2002; Dykhuizen, 1998). Metagenomic and non-coding RNA analysis have shown significant differences can be seen in the genome of the same species, driven by their response to their ecological microclimatic differences that drives sympatric speciation in soil bacteria (Mukherjee *et al.*, 2022). However, despite the high diversity and presence of bacteria in the soil, only 0.1 – 1 % can be cultured under laboratory conditions through standard procedures (Aslam *et al.*, 2010). This has been demonstrated using methods such as fluorescence *in situ* hybridisation (FISH) which exceeds CFU counts by 30 – 1,000-fold (Bakken and Olsen, 1987).

The difficulty with culturing microorganisms from soil can be attributed to several factors. When looking at soil on a microscopic scale most literature discusses small functional units called aggregates. One aggregate contains Bacteria, Archaea and Eukarya as well as the six critical elements hydrogen, carbon, nitrogen, phosphorus, oxygen, and sulphur required for life. The soil biota found in an aggregate are responsible for biogeochemical cycling and as such has been referred to as a snapshot of the biological universe (Fortuna, 2012; Pii *et al.*, 2015). Aggregate composition and size have direct effects on how bacteria are recovered, with other features including biological secretions, plant roots and organic matter binding

together in different amounts altering how they must be handled (



**Micro-aggregate < 250  $\mu\text{m}$  > Macro-aggregates**

Figure 7).



## Micro-aggregate < 250 $\mu\text{m}$ > Macro-aggregates

**Figure 7:** The composition of a soil aggregate. Aggregates are formed by the following six features combining into a matrix which then forms the overall soil complex. These aggregates exist as micro and macro aggregates based on their size. Sand, clay and silt are present depending on the base strata, the environmental organic matter is compromised on surface litter that is gradually degraded, the fungal, plant roots and microorganisms are all dependant on what is present based on the above and below ground communities, as are the biological secretions such as mucus, exopolysaccharides and other cellular excretions.

Aggregate size impacts the community structure of the associated microorganisms which in turn dictate the key functions of the overall soil (Fortuna, 2012; Blaud *et al.*, 2017). Further variation is caused by the availability of nutrients, minerals, water as well as changes in the organic matter, pH, temperature, and base stratum. All these features change with soil horizon and distance, increasing variation on a 3D scale (Crowther and Grossart, 2015). As such, bacterial diversity also functions on a larger scale from macro-environments up to geographical environments. It is well established that microorganisms are non-randomly and non-uniformly distributed throughout a space, with a meta-analysis showing that much of a community composition can be explained in a manner analogous to plants and animals by the

effects of the inhabited environment and the geographical distance (Hanson *et al.*, 2012; Franklin and Mills, 2003).

Endemism, the isolation of bacteria to specific geographic locations, was shown through the isolation of 85 unique fluorescent pseudomonads from 38 samples over 10 sites in four different continents. Of the 85 species, there was no crossover of genotypes at continental or site level, suggesting strong endemism (Cho and Tiedje, 2000). This indicates how the high level of spatial heterogeneity provided creates niches (microhabitats) in which unique bacterial communities can exist and evolve (Tilman, 1994; Gause, 1934). Microhabitats exist on such small scales that large samples can easily underrepresent the actual diversity in samples and as such, small-scale sampling with high numbers of repeats is recommended to help assess the community composition of these microhabitats (Kirk *et al.*, 2004). Disruption of these microhabitats when culturing in suspension creates difficulties in culturing, as exposure to other communities quickly increases competition which favours more abundant bacterial species, leading to the extinction of smaller fractions of the community. This has been demonstrated through the use of sub-sampling to limit cell interaction, allowing increased recovery of soil microorganisms, with sets of four *Nitrobacter* serotypes being recovered in much higher percentages from small  $10^{-3}$  mm<sup>3</sup> volume soil samples (Grundmann and Gourbière, 1999).

Therefore, a specific sampling method that encompasses large scale site differences is required. This can utilise covering two different soil habitats for different baselines. However the sampling should also consider small scales for singular samples by using multiple sampling points per biological replicate. These considerations to sampling are required to allow accurate representation of the bacteria present which is the aim for this overall body of work.

Another issue arises from the complexity of soil aggregates, in which bacteria become entrapped and the Methods achieving significant disruption are required to achieve disaggregation is required for accurate assessment of microorganisms. (Lindahl, 1995). The strong binding that occurs between bacteria and soil particles involves a variety of mechanisms, the disruption of which may lead to severe cell damage and confound downstream analyses (Lindahl and Bakken, 1995). A range of methods have been assessed for disaggregation including simple methods such as mechanical disruption with a range of

instruments, ultrasonic disruption and more complicated methods such as density gradient centrifugation, with no clear consensus on which represents the most appropriate method (Bakken, 1985; Martin and Macdonald, 1981; Lindahl, 1995).

The complexity of the habitat is not the only limiting factor to consider. The poor nutrient content of the soil leaves many bacteria in a nutrient-starved state, which when exposed to the high nutrient compositions of normal laboratory agars and broths, leads to nutrient shock and death whilst also selecting for fast-growing bacteria which can form colonies/biofilms (Aslam *et al.* 2010. Ferrari *et al.* 2005). Low nutrient, highly diluted or novel soil extraction agars which mimic these bacteria's natural environment have been successfully used in the past for the recovery of the hard to cultivate members of the soil microbiome (Davis, Joseph and Janssen, 2005; Hamaki *et al.*, 2005). A more novel approach is the use of isolation chip (iChip) which utilises single cell diffusion chambers in situ for the high throughput recovery of novel isolates from the environment (Nichols *et al.*, 2010). Cell size distribution in taxa also leads to bias as most colony-forming units are mostly derived from large cells; the frequency of small cells in microscopic counts negatively correlates with percentage viability of plate and microcolony counts (Schulz and Jørgensen, 2001; Bakken and Olsen, 1987). However cell size is a relatively plastic trait that has been linked to taxa but also growth conditions and the growth phase can also effect cultivation of the soil microbiome (Portillo *et al.*, 2013). When alternative agars are coupled with long growth periods with the removal of competitive bacteria from agar, high levels of previously uncultured bacteria have been recovered (Eilers *et al.*, 2001; Campoverde, 2015).

Further effects on the recovery of bacteria can be caused by the sample handling method, with soil traditionally being oven-dried and then sieved to make it easier to work with. Although homogeneity makes working with soil easier, it also affects key sets of soil parameters with a direct impact on microbial community structure (Blaud *et al.*, 2017; Thomson *et al.*, 2010). Traditional culturing methods reveal a significant loss in microbial community members due to soil drying (Mikha, Rice and Milliken, 2005; Griffiths *et al.*, 2003). This is exacerbated by rewetting soil which increases respiration, due to the mobilisation of nutrients that favour certain species leading to alterations in the community composition (Clein and Schimel, 1994). As such it is preferable to perform both DNA and culturing examination of soil communities with freshly sieved undried soil (Thomson *et al.*, 2010).

The issues caused by the complexity of the soil environment listed above are seen as a range of explanations as to why, at most, only ~1 % of the soil microbiome has been cultured (Aslam *et al.*, 2010). As such many studies have attempted to resolve these limitations using metagenomic based approaches, by assessing the full community composition using metagenome shotgun sequencing or via other taxonomic resolution methods such as 16S rRNA gene sequencing (Popescu and Cao, 2018). However, DNA based approaches face similar biases due to limitations caused by the soil matrix, as well as biases caused by different DNA extraction methods and amplification biases (Delmont *et al.*, 2011; Carrigg *et al.*, 2007). Out of vertical soil sampling, cell separation by density, cell lysis protocol, and distribution DNA fragment via size, the effect of cell lysis was shown to have the largest effect on the number of operational taxonomic reads, particularly for rare taxa (Delmont *et al.*, 2011). Even if sample sizes, handling and DNA extraction were standardised, use of different quality filtering tools and assemblers used post-sequencing are also known to introduce bias that can lead to non-replicable and comparable results even within the same sample (Bharti and Grimm, 2021).

A further consideration when looking at the rhizosphere of a particular plant from a natural environment is how to ensure that the root originates from the plant of interest. If it does not, the rhizosphere analysis is not relevant. Root identification is traditionally achieved by microscopy differentiating root characteristics such as diameter, colour and texture which is enough to distinguish between genera of trees (Biddle, 1998; Cutler *et al.*, 1987). However, many rhizosphere studies do not check whether the roots originate from the species of interest, preventing the ability to conclude that results are relevant to their investigation. To overcome this, an unpublished LAMP assay was developed at Forest Research to confirm oak origins for rhizosphere samples based on the isolated, amplification and subsequent colour change in the presence of an intercalating dye (Bridget Crampton, personal communication, 2021).

Because of the complex nature of the soil, sampling, sample transport, storage and processing must all be considered as during this time a range of chemical changes can take place. As discussed here, changes in the chemistry of soil have effects on the microbiome of a sample, so samples must be processed and stored using consistent parameters to ensure minimal variations to microbiome composition. Processing of samples before suspension and

enumeration of the microbiome on agar is another point in which the biochemistry of samples may change drastically enough to affect community composition. It is also the first point in the process in which small aggregates will be exposed to each other, allowing for competition to occur, which will affect both DNA extraction and physical culturing.

The aims of the chapter were:

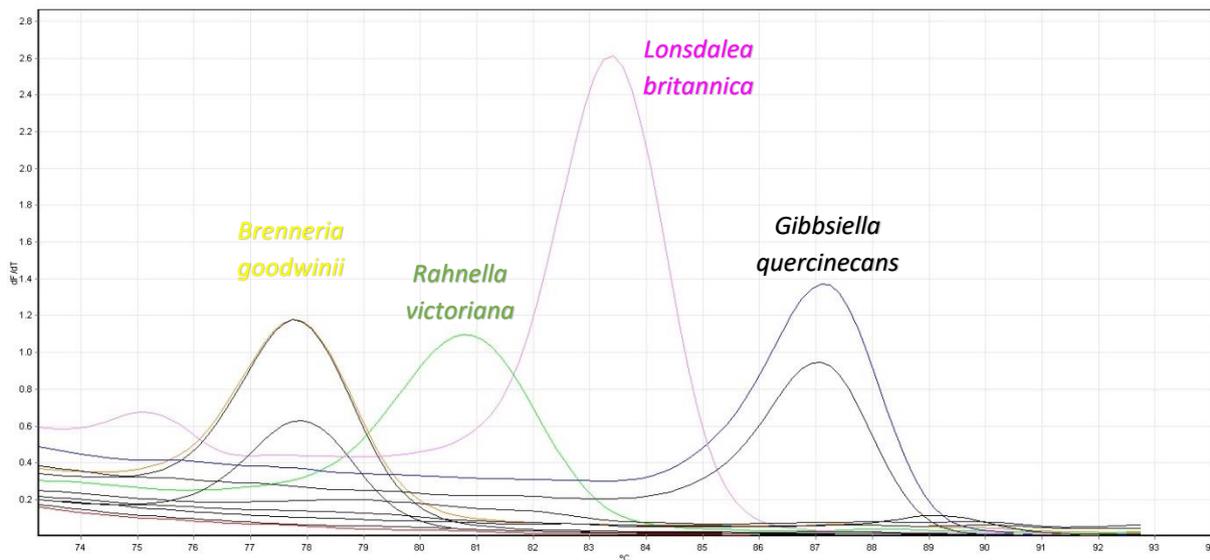
- To optimise the method used to screen soil samples for AOD bacteria.
- To observe the effect of different storage conditions on the number of different bacteria recovered and the amount and purity of DNA recovered from soil samples.
- To increase the diversity and number of culturable cells recovered from soil samples by optimising the method used for the disaggregation of cells.
- To identify the most proficient method to obtain high-quality DNA for further analysis rhizosphere soil.
- To pick the most suitable sampling method to reduce the effect spatial variation on samples.
- To use molecular methods to identify rhizosphere samples originating from oak, and to confirm that the diseased trees are suffering from AOD to ensure samples are relevant to the study.

### 3.3. Results

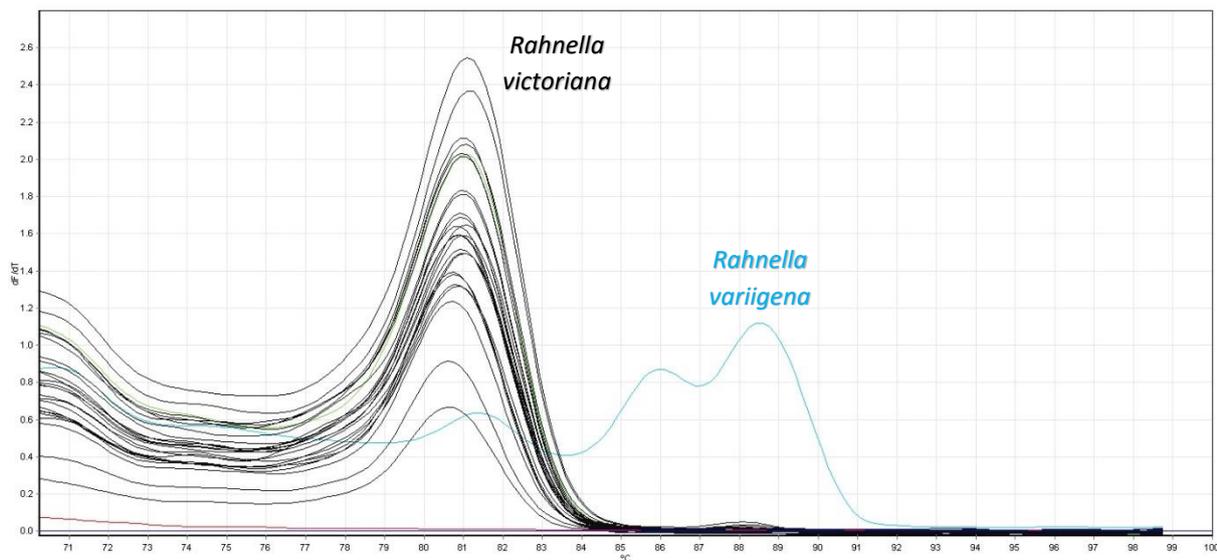
#### 3.3.1. Identification of bacteria from Forest Research samples

A selection of bacterial isolates, mainly from oak and Tilia (lime) trees with various bleeding cankers collected from a range of different sites, were provided by Forest Research. The aim was to find an appropriate method to identify AOD bacteria and, later, other bacterial isolates of interest, which could be used when analysing samples collected from Hatchlands. Firstly, High-Resolution Melt (HRM) analysis was performed on the samples (2.2.12) to allow for the identification of any of the bacteria which are commonly isolated from the AOD lesion microbiome. Isolated DNA was extracted from single colonies (2.1.2; 2.2.1) prior to PCR amplification of the 16S rRNA gene (2.2.6) and the bacteria were identified from the sequencing products (2.2.11). Isolates of interest underwent *gyrB* and *rpoB* gene sequencing to allow for more accurate identification. Finally, results from both sequencing and HRM analysis were compared to the TaqMan probe results from Forest Research, allowing

comparison between the application of the three methods. An example confirmation of the identification of AOD bacteria using the HRM method are shown below (Figure 8, Figure 9). Comparative identification showed similar results for the three identification methods, though the Taq-man probe identified multiple AOD isolates, while HRM identified one and 16S rRNA sequencing was more suitable for the identification of other isolates as shown in Table 8.



**Figure 8:** Melt curve obtained from multiplex High resolution melt, in which species specific primers are used to amplify DNA and then 0.5 °C temperature increases are used to generate fluorescent based curves based on the denaturation of the PCR products for the detection of AOD lesion bacteria. Samples originated from a set of Forest Research swabs taken from oaks and lime trees from a number of different sites across the UK. Positive controls are depicted by coloured peaks where orange, green, purple and blue represent *Brenneria goodwinii*, *Rahnella victoriana*, *Lonsdalea britannica* and *Gibbsiella quercinecans*, respectively. Black peaks indicate the presence of both *Brenneria* and *Gibbsiella* in these samples.



**Figure 9:** Melt curve obtained from high resolution melt protocol 2 which uses standard PCR followed by temperature increases of 0.5 °C from 68 °C for the differentiation of *Rahnella* species, using one primer set that binds to all species but with different melting temperatures. The samples originated from Forest Research swabs. 16S rRNA gene sequencing gave mixed results of *R. variigena* and *R. victoriana*. HRM identified isolates as *R. victoriana*, shown by the black peaks, the purple line shows the negative control and the blue peak shows *R. variigena* positive control.

**Table 8:** Identities of bacterial spikes from Forest Research. Taqman probe results were obtained at Forest Research. High Resolution Melt (HRM) analysis was used to assess agar spikes for AOD bacteria. Non-AOD lesion isolates were identified by 16S rRNA gene sequencing, with *gyrB* and *rpoB* sequencing used to confirm the identification. - = no identification of AOD bacteria. No growth indicates that no isolates were recovered on solid media.

Location	Culture	Taqman probe ID	HRM ID	16S rRNA ID	<i>gyrB</i> ID	<i>rpoB</i> ID
Shenley	Shen 1	-		<i>Serratia glossinae</i>		
	Shen 2	-		<i>Serratia glossinae</i>		
	Shen 3	-		No Growth		
	Shen 4	-		No Growth		
	Shen 5a	<i>B. goodwinii</i>		No Growth		

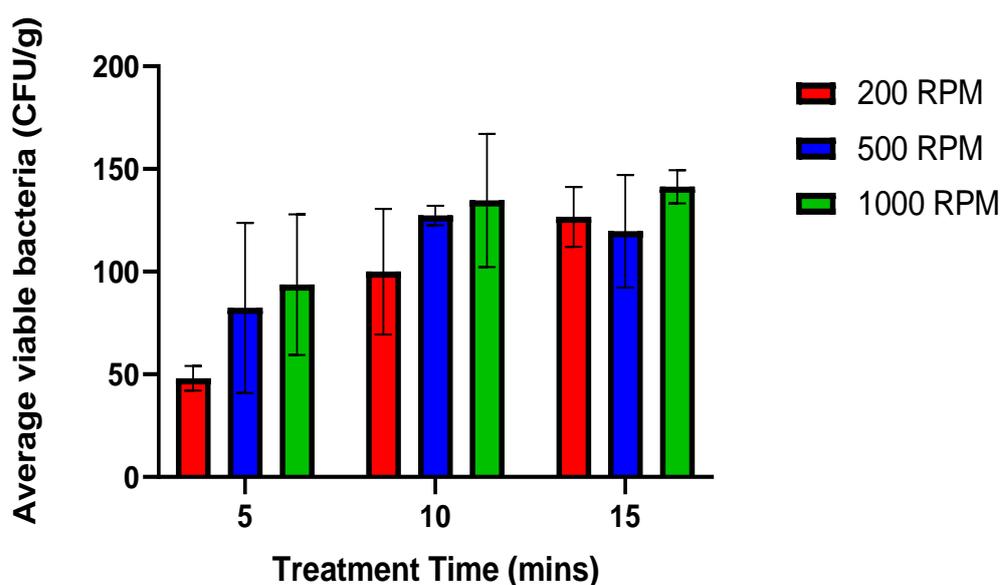
	Shen 5b	<i>B. goodwinii</i>		No Amplification		
	Shen 6	-		No Amplification		
	Shen 7	-		No Amplification		
	Shen 8	<i>B. goodwinii</i>		No Amplification		
	Shen 9	-		No Amplification		
	Shen 10a	-		No amplification		
	Shen 10b	-		No amplification		
Tilia	Tilia 1	<i>B. goodwinii</i>		<i>Erwinia toletana</i>		<i>E. toletana</i>
	Tilia 2	-		<i>Erwinia toletana</i>		<i>E. toletana</i>
	Tilia 3	-		<i>Erwinia toletana</i>		<i>E. toletana</i>
	Tilia 4	-		<i>Erwinia toletana</i>		<i>E. toletana</i>
	Tilia 5	-		<i>Erwinia toletana</i>		<i>E. toletana</i>
	Tilia 6	-		<i>Erwinia toletana</i>		<i>E. toletana</i>

Wilderness	Wild 1	<i>B. goodwinii</i> , <i>G. quercinencans</i> , <i>R. victoriana</i>		No Growth		
	Wild 2a-c	<i>B. goodwinii</i> , <i>G. quercinecans</i> , <i>R. victoriana</i>		<i>Rahnella</i>	<i>R. victoriana</i>	
	Wild 3	<i>B. goodwinii</i> , <i>G. quercinecans</i>	<i>B. goodwinii</i>	<i>B. goodwinii</i>	<i>B. goodwinii</i>	
	Wild 4	<i>B. goodwinii</i> , <i>G. quercinecans</i>		No Growth		
	Wild 5	<i>B. goodwinii</i> , <i>G. quercinecans</i> , <i>R. victoriana</i>	<i>B. goodwinii</i>	<i>B. goodwinii</i>	<i>B. goodwinii</i>	
	Wild 6a, b	<i>B. goodwinii</i> , <i>G. quercinecans</i> ,	<i>B. goodwinii</i>	<i>B. goodwinii</i>	<i>B. goodwinii</i>	

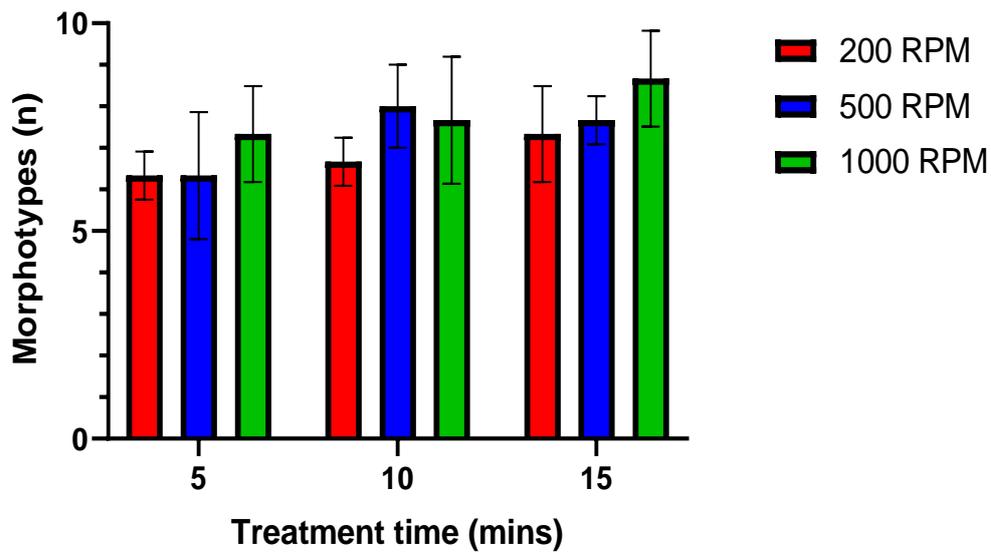
The HRM results demonstrate increased sensitivity and the benefit of not having to culture bacteria for the detection of AOD lesion pathogens, as shown by the detection of *Brenneria goodwinii* in three Shenley samples, one *Tilia* and one Wilderness sample via the TaqMan method. These samples were not detected via other methods due to limitations with the growth of cultures and PCR amplification issues with specific primers. HRM and 16S rRNA gene analysis detected *B. goodwinii* in the same samples as the TaqMan probe indicating all methods reliably identified the pathogen, however, only one AOD species was detected in each sample via those methods as opposed to three by TaqMan probe. The *Tilia* (lime tree) cultures were all tentatively identified as *E. toletana* by 16S rRNA gene sequencing. The *gyrB* PCR primers used could not amplify the *gyrB* gene, and instead, *rpoB* gene sequencing was used to confirm this identity.

### 3.3.2. Optimisation of the dispersal of bacteria from soil aggregates

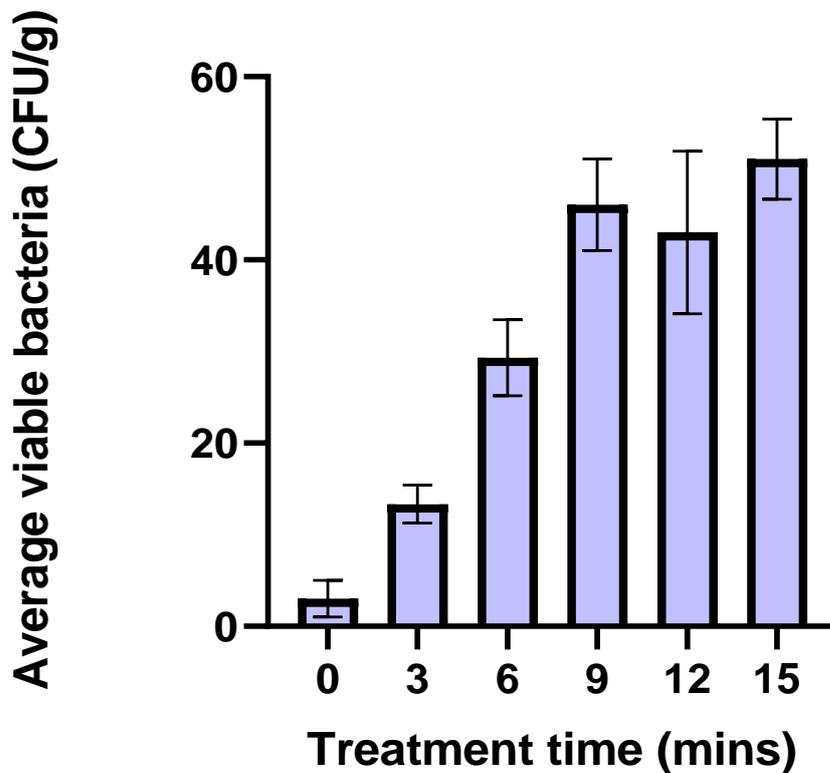
Within the literature, there is no definite published method to achieve maximum separation of bacteria from the soil matrix and as such, identification of the optimal method was required. Following the initial suspension of soil in ¼ Ringers as described in 2.1.5, the most cited procedures were assessed which included disruption in a sonicating water bath, mechanical disruption using a magnetic stirrer or disruption in a shaking incubator at 28 °C. Each disruption method was attempted at different power outputs and time frames to obtain an idea of the most efficient method for the separation of the most bacteria from the soil complex. Dilutions were made and spread on LB agar and, after incubation for two days, the number and diversity of colonies were counted, generally showing that longer and harsher treatments allow for high CFU and morphotype recovery (Figure 10 - Figure 15).



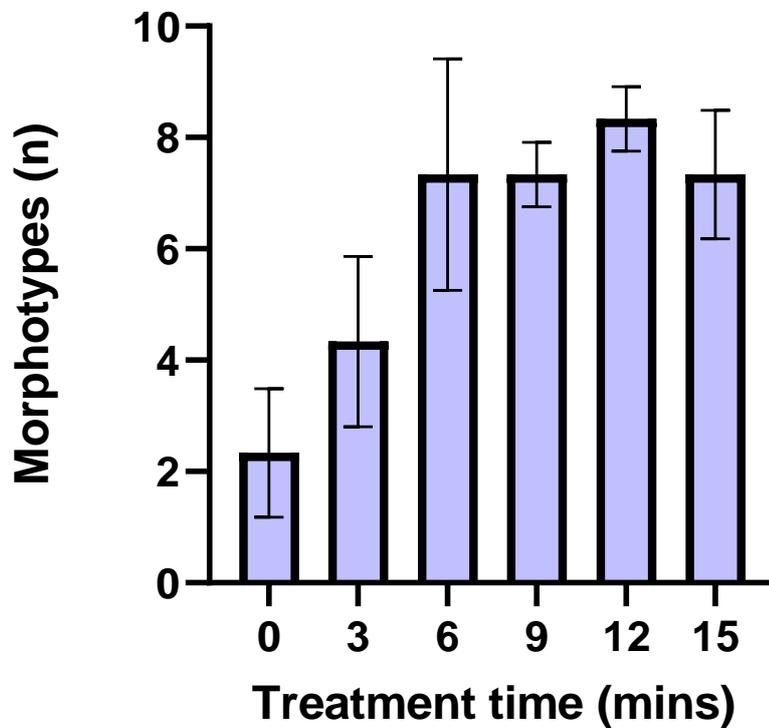
**Figure 10:** Isolation of bacteria from soil on a magnetic mixer. The soil was disrupted at three different speeds using a Teflon bar on a magnetic stirring plate, for 5, 10 and 15 minutes. The average CFU recovered on LB agar from each RPM at  $10^{-4}$  with  $n=3$  with error bars representing standard deviation.



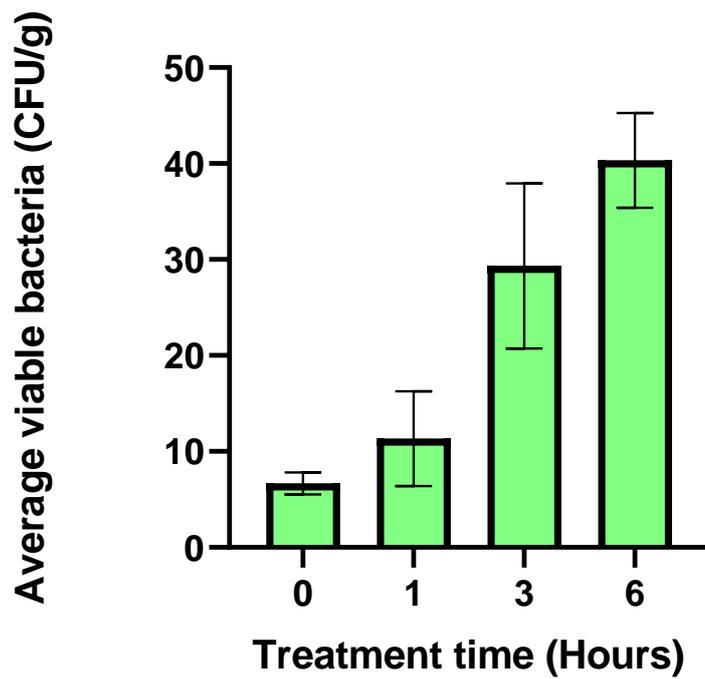
**Figure 11:** The average morphologically distinguishable colonies from soil on magnetic stirrer. The soil was disrupted at three different speeds using a Teflon bar on a magnetic stirring plate, for 5, 10 and 15 minutes. Colonies were recovered on LB agar from each RPM at  $10^{-4}$  with  $n=3$  with error bars representing standard deviation.



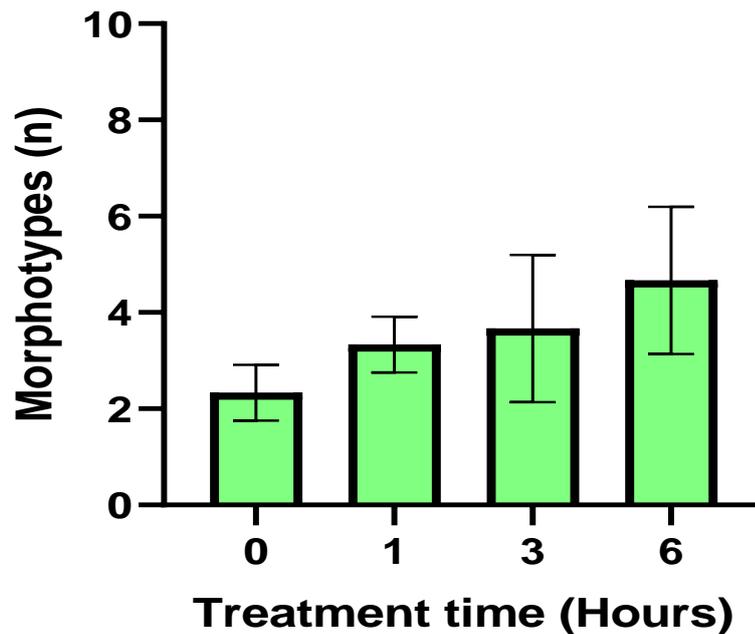
**Figure 12:** Isolation of bacteria by disruption in an ultrasonic water bath. The soil was disrupted for 15 minutes with samples being taken at three-minute intervals. The average CFU at  $10^{-4}$  with  $n=3$  and standard deviation error bars are shown.



**Figure 13:** The average morphologically distinguishable colonies from soil disrupted in an ultrasonic water bath. The soil was disrupted for 15 minutes with samples being taken at three-minute intervals. Colonies were recovered by disruption in an ultrasonic water bath at  $10^{-4}$  with  $n=3$  and standard deviation error bars shown.



**Figure 14:** Isolation of bacteria in a shaking incubator. The soil was disrupted for 6 hours with samples being taken at 0, 1 and 3-hour intervals. The average CFU recovered by disruption in a shaking incubator at 200RPM at  $10^{-3}$  with  $n=3$  and standard deviation error bars shown.



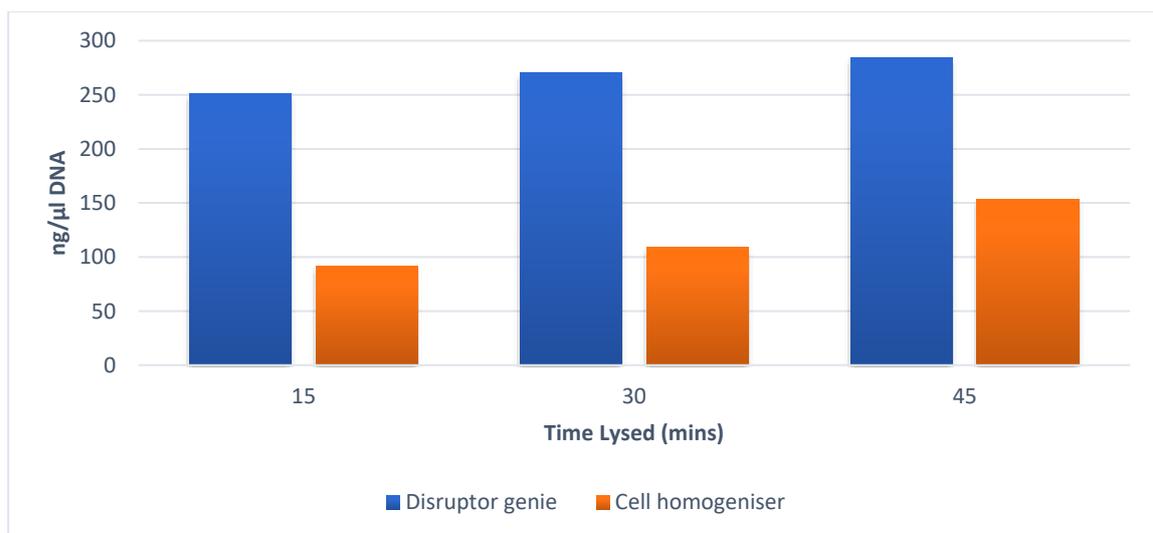
**Figure 15:** The average morphologically distinguishable colonies recovered by disruption in a shaking incubator. The soil was disrupted for 6 hours with samples being taken at 0, 1 and 3-hour intervals. Samples were incubated at 200RPM at  $10^{-3}$  with  $n=3$  and standard deviation error bars shown.

A higher number of colony-forming units (CFU) and in all cases, excluding 15-minute treatment in an ultrasonic water bath, a higher number of visually distinct morphotypes could be gained by increasing the length of treatment. For disruption using a magnetic stirrer, increasing the RPM to 1000 resulted in the highest level of morphotypes as well as CFUs as demonstrated in Figure 10 - Figure 15.

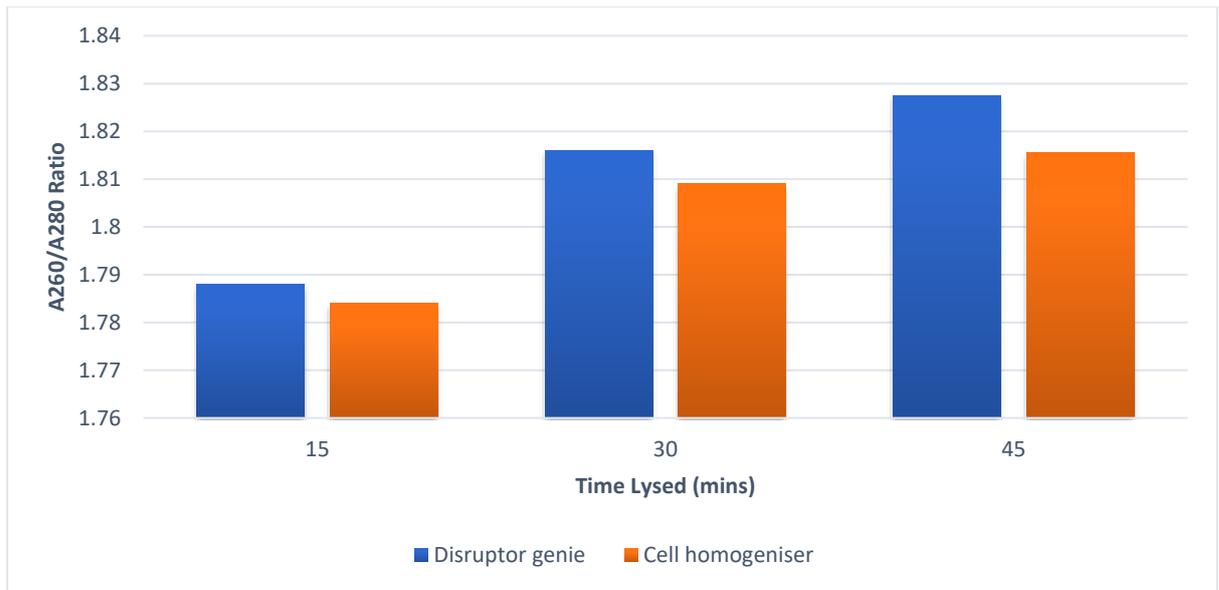
The optimal method for dispersion of cells was 10 g of soil suspended in 95 ml of Ringers followed by mechanical disruption using a magnetic stirrer at 1000 RPM for 15 minutes, then dilution to  $10^{-4}$  which on average led to the highest recovery of cells (within the 30-300 CFU range), and the widest range of morphotypes, compared to shaking incubation (no significant difference  $P > 0.05$  with ultrasonic water bath). This was tested using the numerically highest average of each variable at  $10^{-4}$  and running paired sample T-tests. The result for each pairing was  $P < 0.05$ , meaning there was a statistically significant difference in the mean CFU and the number of morphotypes recovered.

### 3.3.3. Optimisation of soil DNA extraction

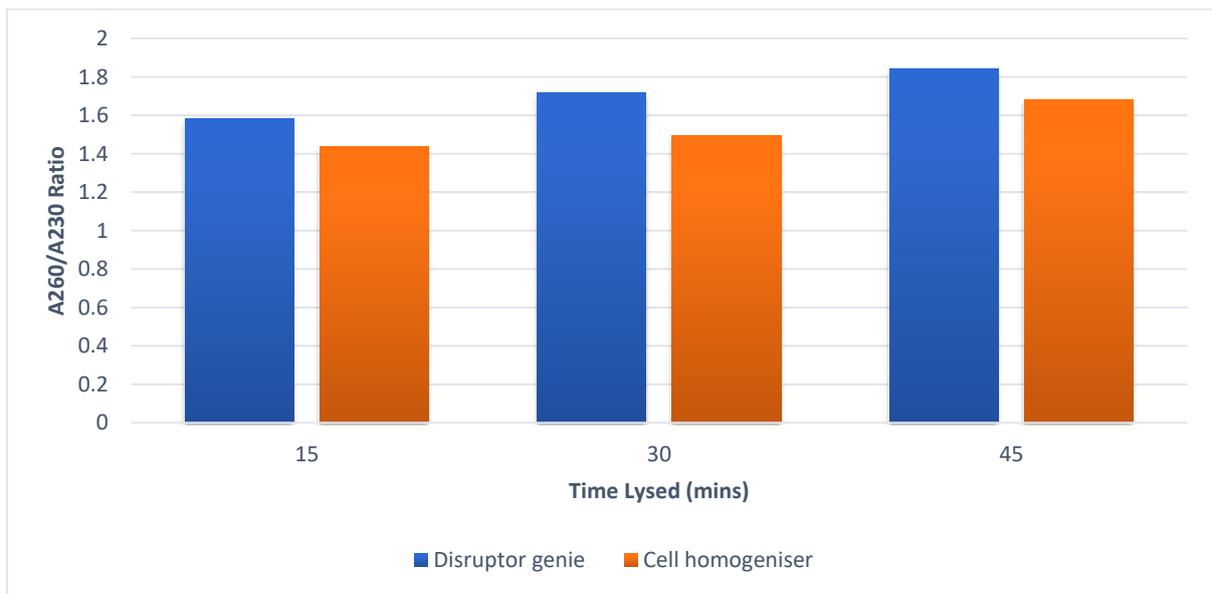
Due to the same limitations of separating bacteria from the soil matrix but with the added complexity of environmental contaminations such as humic acids further limiting recovery, extraction of high amounts of pure DNA ( $\text{ng}/\mu\text{L}$ ) is difficult. There are commercial kits that are available to recover purified DNA from soil samples, which remove contaminants that can interfere with the purification and downstream application of extracted DNA. The kits have set protocols provided by the manufacturer, however, the time that the sample is homogenised to ensure cell lysis is not specified, because it varies depending on the machine used and the composition of the soil. Both a Disruptor Genie (Scientific Industries, Inc) and a TissueLyser LT (Qiagen) were chosen as suitable equipment to test at a range of time frames. The outputs were compared using a Nanodrop DNA spectrophotometer (2.2.9) to see which method resulted in the highest amount of DNA recovery and how the length of homogenisation affected the purity of the extracted DNA (Figure 16 - Figure 18).



**Figure 16:** The amount of DNA ( $\text{ng}/\mu\text{L}$ ) recovered from the soil at both different time frames in both the Disruptor genie and TissueLyser LT  $n=1$ .



**Figure 17:** The purity of DNA from the A260/A280 ratio recovered from the soil at both different time frames in both the Disruptor genie and TissueLyser LT n=1.



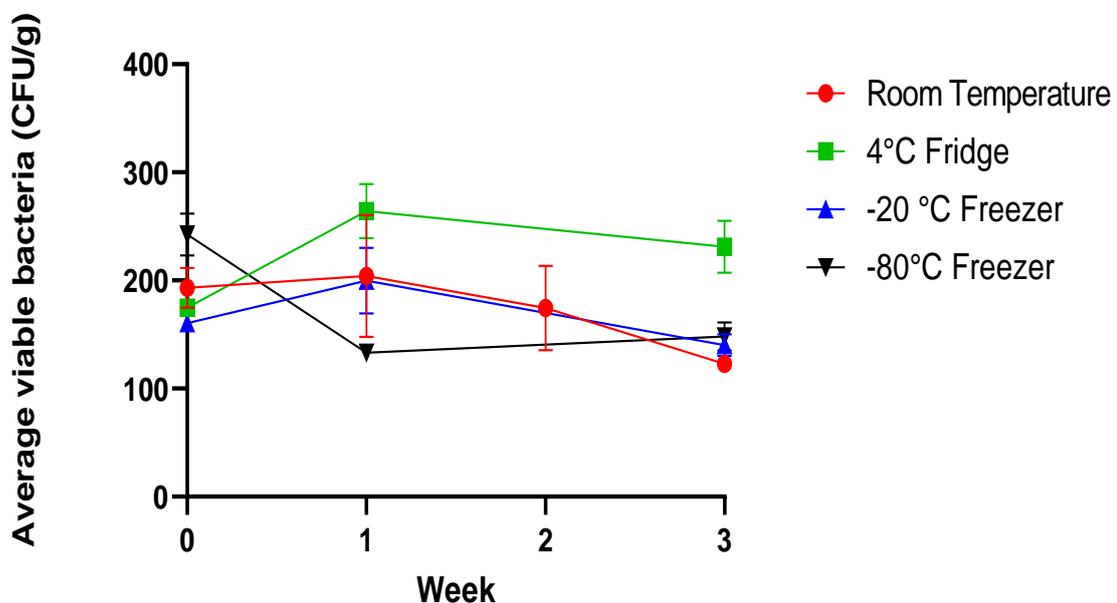
**Figure 18:** The purity of DNA from the A260/A280 ratio recovered from the soil at both different time frames in both the Disruptor genie and TissueLyser LT n=1.

The highest amount of recovered DNA was gained by using the Disruptor Genie for 45 minutes prior to extraction by the Power Soil Pro kit (Qiagen) (Figure 16). However, more interesting are the purity results for both the A260/280 and A260/230 ratio, where DNA can only be considered pure if the value exceeds 1.8. This means that the Disruptor Genie run for 45 minutes was the only run that gave 'pure' DNA (Figure 17 and Figure 18). This optimisation

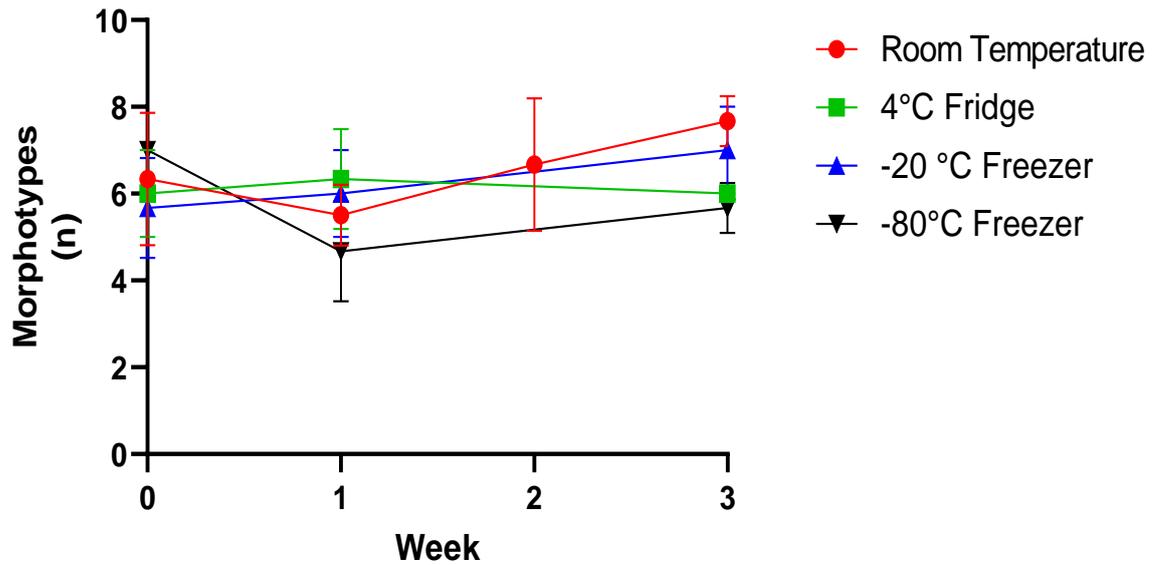
experiment was run on trial kits which limited the amount of replications that could be performed, as such the results are indicative but give no significant conclusion.

### 3.3.4. Soil storage optimisation

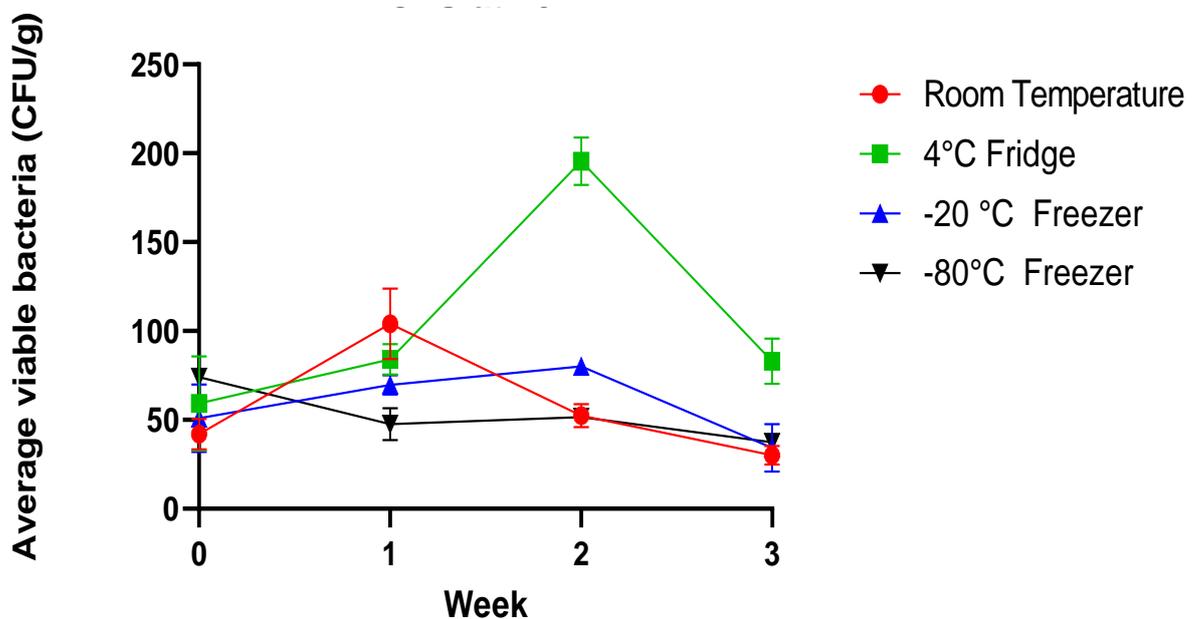
To determine the best method to ensure that the biological composition of samples remains consistent with those seen at the point of sample collection, storage conditions were tested. The currently available storage conditions were average room temperature (18-20 °C), +4 °C, -20 °C and -80 °C freezers. These storage temperatures were tested following collection of samples, by immediately extracting DNA and culturing bacteria on a solid medium (LB) for a control 0-hour result. DNA was extracted and tested weekly for purity (2.2.9) and bacteria extracted in suspension (2.2.3) and cultured to count the number of CFUs and colony morphologies recovered (Figure 19 - Figure 22). A two-sample T-test was applied to both CFUs and diversity, comparing the mean from week 0 to each subsequent week. This allowed for any statistically significant change in the mean to be detected by the generation of a P-value < 0.05. A mixed-effect analysis was used to assess if significant differences could be seen between samples. This was used due to missing values for week 2 caused by colonies being 'Too Numerous To Count' (TNTC) leading to both fixed and random errors.



**Figure 19:** Different storage conditions effect on the number of bacteria recovered from soil. The average colony forming units at  $10^{-3}$  recovered from the soil after storage in different conditions for a month and sampled weekly,  $n=3$  with standard deviation error bars.

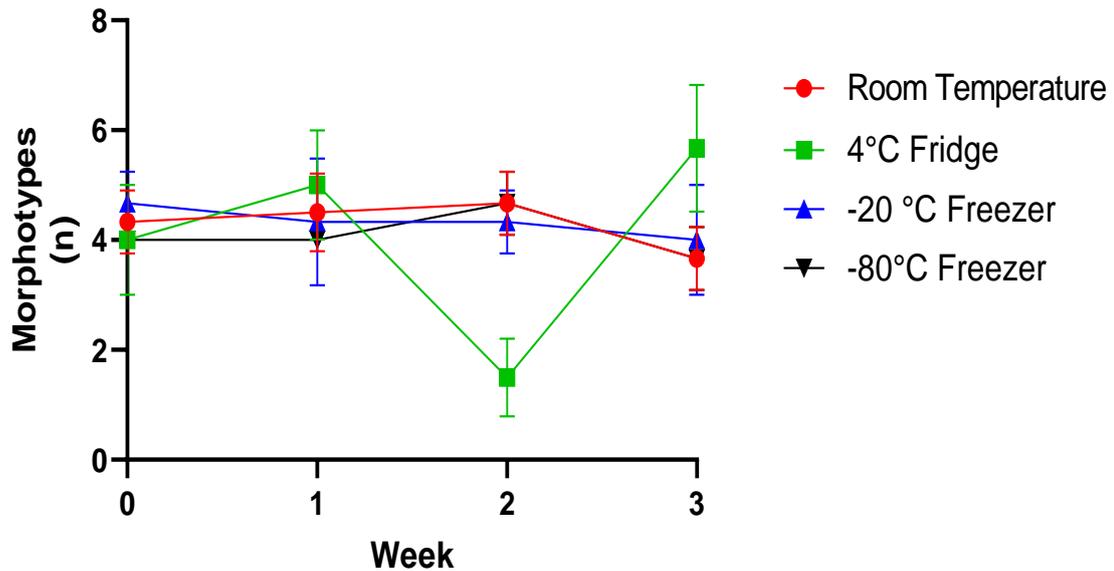


**Figure 20:** Different storage conditions effect on the types of bacteria recovered from soil. The average number of morphologically distinguishable colonies recovered from the soil after storage in different conditions for a month and sampled weekly, n=3 with standard deviation error bars.



**Figure 21:** Different storage conditions effect on the number of bacteria recovered from soil. The average number of colonies forming units at the  $10^{-4}$  dilution recovered from the soil

after being stored in different conditions for a month and sampled weekly, n=3 with standard deviation error bars.



**Figure 22:** Different storage conditions effect on the types of bacteria recovered from soil. The average number of morphologically distinguishable colonies at  $10^{-4}$  recovered from the soil after being stored in different conditions for a month and sampled weekly, n=3 with standard deviation error bars.

Both the CFUs and morphotypes were consistently recover from all storage conditions for the full three weeks of the experiment (Figure 19 - Figure 22). A spike in CFU counts and a drop in morphology can be seen during week two for fridge samples when diluted to  $10^{-4}$  (Figure 21 and Figure 22). This could possibly be attributed to high levels of respiration leading to moisture in the sample that is known to cause a sudden boom in nutrients and subsequently an increase in the species that favour these nutrients (Clein and Schimel, 1994).

Both room temperature and fridge storage show the highest level of variation in samples as demonstrated by the large error bars, which are not present for both freezer samples. The statistical analysis also showed all methods as consistent with both mixed-effects analysis and two sample T-tests, both showing no statistical significance between means ( $P > 0.05$ ). As such -20 °C appears the most consistent method as it consistently shows the lowest deviation through error bars of the course of the experiment. However, statistical analysis was only applied to  $10^{-4}$  dilution data, due to missing data in the week 2 analysis at  $10^{-3}$ .

### 3.3.5. Storage of DNA in DNA Shield

Storage of soil before DNA extraction was shown to be an issue (data not included), with pure DNA not isolated after even one day of storage under all storage conditions (room temperature, -4°C, -20°C and -80°C). Therefore, alternative options for the long-term storage of soil prior to DNA extraction were investigated. One potential option was to use DNA/RNA shield (DRS, Zymo) which inactivates enzymes and viruses to ensure stable long-term storage of DNA below -20 °C. Following soil collection, half was sieved in the field and the other half was brought back to the lab and then sieved. From both lab and field sieved samples, 250 mg of soil was placed in an Eppendorf containing no DRS, another containing 2.25 mL of DRS and finally, 1 g of soil was placed in 9 ml of DRS. DNA extraction was performed on both the field and lab sieved samples (2.2.4) in three treatments and the purity and concentration of DNA (ng/μL) was measured by using the Nanodrop (2.2.9). The remaining soil samples were then frozen at -20 °C for seven days before DNA was extracted and the purity and concentration of DNA (ng/μL) was measured by using the Nanodrop (

Table 9 and

Table 10).

**Table 9:** Purity and concentration of DNA extracted from soil (n = 3). Soil was collected and either sieved in the field or laboratory and then had immediate DNA extractions performed prior to storage.

Sample	A260/280	A260/230	ng/μL
Field	1.82 ± 0.03	1.42 ± 0.2	62.9 ± 3.7
Lab	1.85 ± 0.02	1.82 ± 0.1	130.3 ± 15.1

**Table 10:** Purity and concentration of DNA extracted from soil after 7-day storage (n = 3). Soil was collected and either sieved in the field or laboratory and then frozen at -20°C with or without shield for 7 days.

Sample	A260/280	A260/230	ng/μL
--------	----------	----------	-------

250 mg Lab sieved (no DRS)	1.77 ± 0.05	1.57 ± 0.08	96.9 ± 5.1
250 mg Field sieved (no DRS)	1.83 ± 0.03	1.82 ± 0.04	205.2 ± 20.1
Lab 250 mg in 2.25 ml DRS	2.02 ± 0.1	0.50 ± 0.2	4.1 ± 0.5
Field 250 mg in 2.25 ml DRS	1.87 ± 0.02	1.59 ± 0.01	9.8 ± 1.0
Lab 1 g in 9 ml DRS	1.85 ± 0.02	1.26 ± 0.3	3.3 ± 0.9
Field 1 g in 9 ml DRS	1.97 ± 0.1	0.65 ± 0.2	5.6 ± 1.3

The results showed that immediate sieving of soil in the lab resulted in good levels of pure DNA compared to DNA extracted from samples sieved in the field. However, after storage high-quality DNA was only recovered from field sieved samples not stored in DNA shield, with lab handled sieved not stored in DNA shield being the second highest and purest DNA recovered.

### 3.3.6. Hatchlands Park Sampling

The sampling took place over the course of two days during which oak trees were selected based on health status and the availability of other oaks to be paired with as described in Section 3.2.1. The sample numbers by location, pairing and health status can be seen in Table 11. As trees were selected, their locations were noted to be plotted on a map in ArcGIS (Figure 23). Samples were then collected in accordance with the standardised method in section 2.1.1. All 80 samples were collected and brought back to the lab for processing on the second day.

**Table 11:** Hatchlands Park sample identities. Hatchlands Park samples divided by their original location, each tree is paired to the number in the adjacent column for that location with the health status of the tree indicated by the column name.

Hatchlands Park Rhizosphere samples			
Parkland		Woodland	
AOD symptomatic	Healthy	AOD symptomatic	Healthy
H1	H2	H11	H12
H3	H4	H13	H14
H5	H6	H15	H16
H7	H8	H17	H18
H9	H10	H19	H20

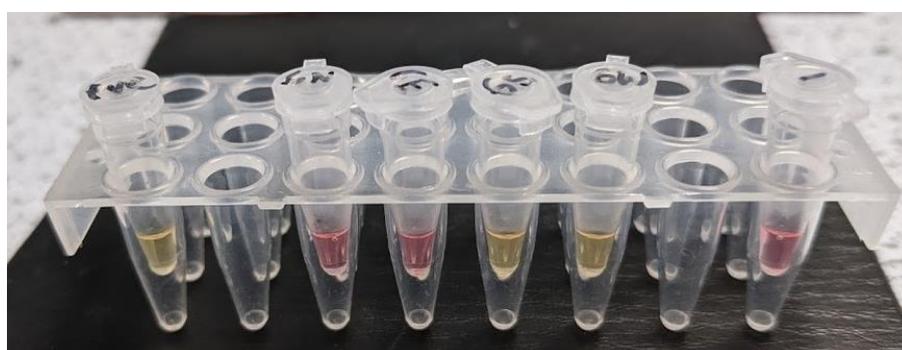


**Figure 23:** Map detailing the location of rhizosphere soil samples taken from Hatchlands Park site. Healthy and AOD symptomatic diseased trees were selected in a paired model with odds and evens being paired from 1 and 2 through to 19 and 20.

### 3.3.7. LAMP confirmation of oak roots

Loop-mediated isothermal amplification (LAMP) (3.2.4) of extracted root DNA (3.2.3) was used to confirm that rhizosphere soil was adhered to oak roots and not to other plants from the surrounding environment. Each sample was tested in batches by biological sample (4 replicates from each tree) and the resulting colour change from pink to yellow (Figure 24) was recorded as positive or negative if no colour change occurred.

**Figure 24:** Results for Hatchlands H6 samples following Loop-mediated isothermal amplification. The resulting colour change seen in a reaction for the amplification and detection of actin genes from oak species. Yellow indicates a positive reaction while negative reactions remain pink due to lack of amplification.

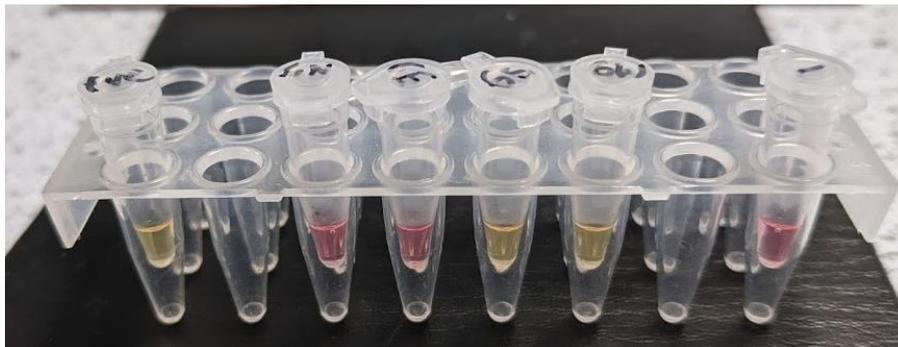


**Table 12:** Loop-mediated isothermal amplification results. Each cardinal point of each root sample separated by both their original sample location (parkland or woodland) and their health status (AOD or Healthy). Samples are positive for oak (green) if a colour change was seen in the first amplification and negative for oak (red) if no change in colour was recorded when undergoing the amplification. Samples were considered late positives (yellow) if a second DNA extraction resulted in a positive amplification as the result implied other plant roots may be present in the sample.

Hatchlands Park Rhizosphere samples			
Parkland		Woodland	
AOD symptomatic	Healthy	AOD symptomatic	Healthy
H1N	H2N	H11N	H12N
H1E	H2E	H11E	H12E
H1S	H2S	H11S	H12S
H1W	H2W	H11W	H12W
H3N	H4N	H13N	H14N
H3E	H4E	H13E	H14E

H3S	H4S	H13S	H14S
H3W	H4W	H13W	H14W
H5N	H6N	H15N	H16N
H5E	H6E	H15E	H16E
H5S	H6S	H15S	H16S
H5W	H6W	H15W	H16W
H7N	H8N	H17N	H18N
H7E	H8E	H17E	H18E
H7S	H8S	H17S	H18S
H7W	H8W	H17W	H18W
H9N	H10N	H19N	H20N
H9E	H10E	H19E	H20E
H9S	H10S	H19S	H20S
H9W	H10W	H19W	H20W

The results in



**Table 12** shows that 55 of 80 samples contained oak roots, while 27 samples were shown not to contain oak roots even after two rounds of LAMP. The 20 late positive samples shown in yellow did not amplify in the first round of LAMP and required re-extraction of DNA to obtain a positive amplification. All trees had two or more rhizosphere samples containing oak roots,

allowing for two positive samples per tree for a total of 40 samples taken forward for further work.

### 3.3.8. AOD confirmation of biological sample bleeds

All selected AOD symptomatic oak trees from both parkland and woodland had visible stem bleeds which were recorded during rhizosphere sampling (3.2.2). In cases where the stem bleeds were low enough to reach, swabs were taken for further confirmation of AOD symptoms by the presence of AOD bacteria found in lesions. As such nine swabs were processed in the lab and HRM analysis was performed (Table 13).

**Table 13:** High resolution melt results for Hatchlands lesion swabs. + indicates the presence of that species in the sample, while - indicates no detection. N/A: No cultures grew from samples 5 and 9 as such no results were obtained for these samples. \* indicates dry bleeds that were not actively weeping.

Tree	<i>Brenneria goodwinii</i>	<i>Gibbsiella quercinecans</i>	<i>Rahnella victoriana</i>	<i>Lonsdalea britannica</i>
1	+	+	-	-
3	+	+	-	-
5	N/A	N/A	N/A	N/A
7	+	-	-	-
9	N/A	N/A	N/A	N/A
11	+	-	-	-
13*	-	-	-	-
15*	-	-	-	-
17	+	-	-	-

*B. goodwinii* was found in five of seven viable samples (two samples were removed due to no growth on agar from swabs), while *G. quercinecans* was detected in two. However, neither *R. victoriana* nor *L. britannica* were found in any of the samples.

### 3.4. Discussion

Soil is a difficult medium to work with, numerous studies have explored the consequence of different sampling processes and their effect on the recovery of organisms (Brumfield *et al.*, 2020; Lauber *et al.*, 2010). In much of the literature, the soil is dried first as it provides a standardised medium to work with, in which parameters like moisture content can be controlled (Griffiths *et al.*, 2003). Likewise, sieving is used to not only remove rocks and other detritus from the soil, but to also increase homogeneity in samples by reducing the spatial heterogeneity caused by the 3D complex of pores across the soil horizon that supports different microbial communities. However, both methods cause direct reductions in respiration, cultivable bacterial counts and microbial biomass (Blaud *et al.*, 2017; Thomson *et al.*, 2010). As such, soil used for this work was not dried in an attempt to maintain a consistent microbial community structure from the point of sampling. Sieving however was necessary to process samples but was performed immediately before extracting DNA from the soil or culturing to reduce any bias that it may cause.

Storage of samples seems to be rarely discussed in the literature but was considered here due to the number of samples being collected at one time: four replicates from five biologically paired trees, giving a total of 80 samples. Long-term storage was unavoidable and therefore required assessment to ensure that that microbial community structure alteration was minimised during this time. The results of section 3.2 have shown that soil can be collected and stored with only minor effects observed on the colony-forming units and different cell morphologies recovered from samples after long periods. The optimal storage temperature for the culturing of organisms was found to be -20 °C in section 3.2.4 with very little deviation on either of the parameters assessed. One caveat of this work is that it was only performed on nutrient rich media, and subsequently use of other alternative media will have to be optimised separately. Another is that the generally heterogenous nature of soil makes it difficult to compare week to week as variation in results could simply be to lack of homogeneity found in the medium. Nevertheless, the results here can still be used to inform the storage and recovery methods, despite conclusive statements about their effects being difficult to make.

Sample storage is seen to be a critical component of the analysis of microbial communities by the 16S rRNA gene and as such the effects were investigated (Rochelle *et al.*, 1994). The results obtained here are conflicting as storage of samples for DNA extraction gave mixed results from both soil storage optimisation and storage of DNA in DNA Shield. All storage options gave rise to impure DNA after even one day of storage under all conditions, but pure DNA could be extracted from frozen soil implicating a suitable storage method. The extraction method of DNA shield was not optimised which is likely why the recovery of DNA from section 3.5 was inefficient, only yielding impure DNA, thus requiring further optimisation. Therefore, to ensure pure DNA will be recovered from samples, immediate extractions should be performed with frozen aliquots of soil with and without DNA shield being taken prior to storage as back-ups. It should be noted that storage conditions confer less variation than environmental conditions, and as such DNA extraction from soil stored at any temperature for two weeks has shown only slight variations in relative abundance of taxa (Lauber *et al.*, 2010).

For culturing of bacteria from soil, one of the key issues noted is the detachment of bacterial cells from their substratum, which is difficult to achieve at sufficient levels without the disruption or damage of cells (Buesing and Gessner, 2002). Mechanical disruption via blenders and stomachers or ultrasonic disruption with probes and water-baths are commonly employed in the literature and as such were all considered (Richter-Heitmman *et al.*, 2016; Liebeke *et al.*, 2009; Buesing and Gessner, 2002). Optimisation of the method used to extract bacteria from the soil matrix was successful in demonstrating that harsh mechanical disruption of soil in suspension is sufficient to separate cells from the micropores that they exist in. Other methods of disruption mentioned in the literature showed significant reductions in the CFUs recovered from samples, which could be attributed as either due to damaging cells or failing to separate them from the matrix. However, the limitation of this work is that it was only applied to bacteria that are easily recoverable on single standard nutrient rich media. There is potential that this method will be insufficient when applied to members of the soil microbiome that are traditionally harder to culture and these methods need to be applied across a range of media optimised for the recovery of soil bacteria.

The sample collection method used here aims to reduce biases caused by the spatial heterogeneity of soil that is known to affect the performance of plants on a site-wide to soil

particle scale level (Xue, Huang and Yu, 2016; Huang *et al.*, 2013). Three methods were employed to address this spatial heterogeneity, firstly by selecting multiple sampling points for each tree to ensure the reduction of bias on individual samples. Secondly, by reducing spatial variation between healthy and diseased trees by sampling in pairs, important individual differences between closely located trees may be observed. Thirdly, by using two different environments the effect of habitat on the sampling bias might be observed. This should allow for three levels of interaction to be observed when studying genomic data and may reveal interesting relationships that might otherwise be missed if only observing the health and disease rhizosphere without greater consideration to the effect of spatial heterogeneity. One aspect that should be noted is that while most paired samples were < 10 m apart, samples 5-6 and 9-10 are almost 80 m apart. These were the closest healthy and diseased paired trees that could be obtained in the parkland site where trees were mainly found in small coppices (Figure 23).

The LAMP and HRM results from sections 3.2.7 and 3.2.8 demonstrate that AOD lesion bacteria were present in bleeds in both parkland and woodland trees, proving that AOD was present across the site. While some swabs from bleeds did not reveal the presence of AOD bacteria this could be due to bleeds being dry at the time of swabbing, which led to no bacteria being recovered on agar. The LAMP results showed that despite the same sampling method being used for each sample, oak roots were not always recovered. This demonstrates that it is important to check root samples do originate from the species of interest for the validity of results. Initially, it was hoped that this would be done in the field so that if oak roots were not recovered then samples could be taken again. However, the method was not field-tested at the date of sampling and as such had to be performed when back in the laboratory meaning samples that did not contain oak had to be excluded for further analysis. This led to issues with the samples taken as 27 were not positive for the oak genes. To ensure even sampling numbers for each biological sample, only two positive samples were taken forward for further analysis (Chapter 3 and 4) following the LAMP results, with other samples being excluded from further analysis.

The work presented in this chapter demonstrates an optimised method for the collection of soil samples, their storage and handling for the isolation of both physical cultures and DNA from samples. It highlights the importance of considering the effect of each parameter on the

outcome gained when investigating such a complex sample. Finally, this work shows the need for sample validation when working with plants in a natural environment to ensure that subsequent results are of relevance to their field, which is not currently practised in most studies.

Chapter 4. Environmental  
reservoirs of the AOD  
bacteria.

## 4.2. Introduction

With no current point of origin currently confirmed for the four bacteria associated with AOD lesions, the two most likely options are that the bacteria originate from the soil, or exist as endophytes (Compant *et al.*, 2021; Chaparro *et al.*, 2012). *A. biguttatus* could also be acting as a vector, but its inconsistent presence in AOD makes this less likely (Denman *et al.*, 2018; Reed *et al.*, 2018). The majority of soil-borne diseases are caused by fungal species due to their ability to survive in the soil as saprophytes on residual plant matter or in a dormant state as spores and sclerotia, until a chance to infect arises (Sharma, Sharma and Kardile, 2021; Delgado-Baquerizo *et al.*, 2020). However, there are a limited number of plant pathogenic bacteria which are capable of surviving in soil for long periods that have environmental and economic impacts due to their damage to woody plants and food crops (Raaijmakers *et al.*, 2009).

For example, an unclassified *Rahnella* species, closely related to *Rahnella aquatilis* causes mild decay in onion bulbs and has been recovered from over 25 % of onion bulbs which cannot be sold (Asselin *et al.*, 2019). *Agrobacterium tumefaciens* invades plants through wounds in the rhizosphere where they bind to cell walls and transfer tumour-inducing plasmids into the plants' cells to trigger cell proliferation and the production of a gall (Finer, Fox and Finer, 2016). Another example is *Pectobacterium carotovorum* which spreads from seed tubers into soil, moving to new points of infection in a wide variety of food crops (Kushalappa and Zulfiqar, 2001). However, *Erwinia amylovora*, which shares many pathogenicity traits with *P. carotovorum*, is unable to survive for longer than five weeks in the soil complex (Zhao, Blumer and Sundin, 2005; Hildebrand, Tebbe and Geider, 2001). Nonetheless, *E. amylovora* in water continues to be pathogenic and soil irrigated with this water can still cause fireblight disease in pear trees (Santander *et al.*, 2020). Could the AOD bacteria be well-adapted to the soil complex, allowing it to survive for long periods? Failing this, can AOD bacteria survive the soil complex for long enough to use roots as the point of entry to cause infection in oaks, in a manner analogous to *E. amylovora*.

It has been discussed at length in previous chapters how bacteria are not randomly distributed. One of the original tenets of microbiology arises from Baas Becking who originated the statement 'everything is everywhere, but the environment selects' is no longer

considered true (De Wit and Bouvier, 2006). The reality is that everything is not everywhere, and the environment does not always select. Global distribution, environment and other factors play key roles in the distribution of microorganisms (van der Gast, 2015). As such it cannot be said that the bacteria that cause AOD are going to be present in the soil. Instead, individual assessment based on their genus characteristics is the only way to predict their ability to survive in soil.

All four of the AOD lesion bacteria used to belong to the family *Enterobacteriaceae* due to the previous classification on classical phenotypic descriptions (Paradis *et al.*, 2005). However, more modern genomic-based approaches have led to the division of the family (Adeolu *et al.*, 2016b). The new division leaves 33 of the original 68 genera in the family *Enterobacteriaceae*, with none of the AOD pathogens remaining. *Brenneria* and *Lonsdalea* moved to the *Pectobacteriaceae* and *Gibbsiella* and *Rahnella* to the *Yersiniaceae* (Janda and Abbott, 2021; Soutar and Stavrinides, 2020). For their species-level groupings, the *Pectobacteriaceae* are predominantly associated with vegetation and insects, but not with environmental samples. *Rahnella* and the *Yersiniaceae* appear to be ubiquitous organisms not confined to a singular niche, being found commonly in environmental samples including soil, water and food (Brady *et al.*, 2022; Asselin *et al.*, 2019; Kämpfer, 2015b). *Gibbsiella*, is not seen to be associated with any one niche at the family level, but again the limited genus isolates have come from vegetation such as apple, pear and oak trees, insects and animal samples (Geider *et al.*, 2015; Brady *et al.*, 2012; Saito, Shinozaki-Kuwahara and Takada, 2012). As such it seems reasonable to expect the detection of *Gibbsiella* and *Rahnella* from soil samples, but less certain for the other members of the AOD bacteria.

The possibility of the recovery of *Brenneria goodwinii* and *Lonsdalea britannica* from soil seems less feasible when looking at the soft rot bacteria, of which the two most studied *Pectobacteriaceae* (*Dickeya* and *Pectobacterium* species) are close relatives of *Brenneria* and *Lonsdalea*. Environmental survival of these bacteria is noted as poor in optimum conditions, and although survival in the rhizosphere is favourable, the antagonistic effects of other microorganisms are well-noted in preventing the detection of low numbers of these bacteria seen in environmental samples (Czajkowski *et al.*, 2015; Perombelon and Hyman, 1989). Further doubt is cast due to a more recent paper in which the detection of *Brenneria* and *Gibbsiella* in both rainwater and soil was tested over 84 days. *Gibbsiella* was reisolated from

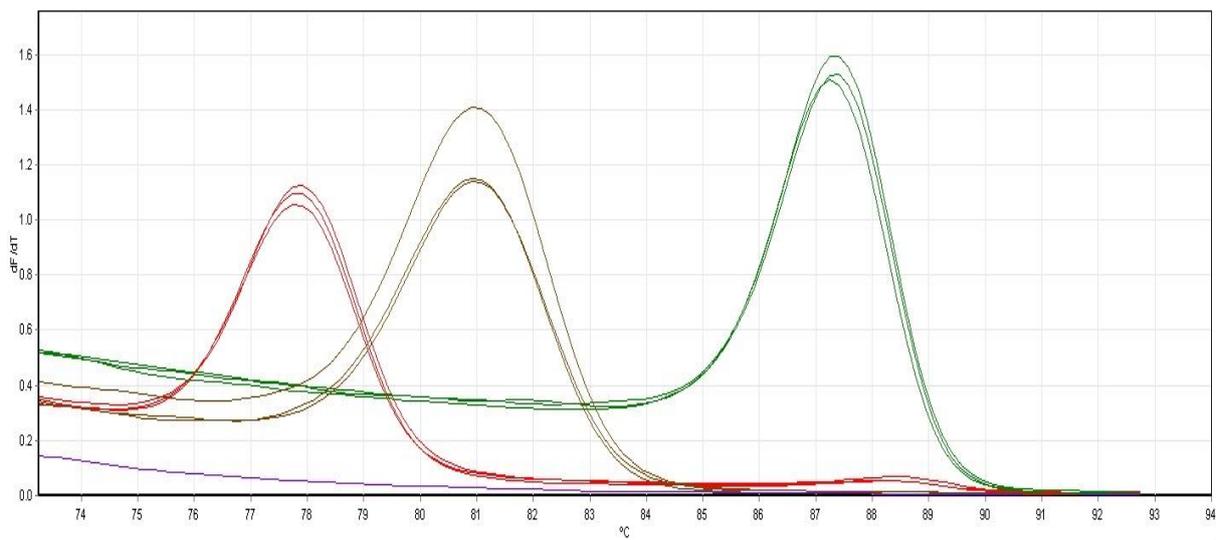
rainwater up to 84 days after inoculation and 28 days after inoculation from the soil. *Brenneria* however, was deemed non-viable in soil and rainwater immediately following inoculation and only detectable with qPCR methods for 28 days, leading to the conclusion that while *G. quercinecans* was viable in a range of ecological niches, *B. goodwinii* was unable to survive outside of the host (Pettifor *et al.*, 2020).

Further understanding of the potential of the ability of AOD lesion bacteria to survive in the soil is required to see if it could function as a reservoir of infection. The literature raises several clear questions, can the AOD bacteria survive in soil, can a specific recovery method be used to isolate these bacteria at a higher rate, due to their lack of finesse in soil and can they be detected in any wild oak rhizosphere samples collected from a site currently suffering from AOD? If they are not present in soil, can the same method be used to detect them from other environmental niches?

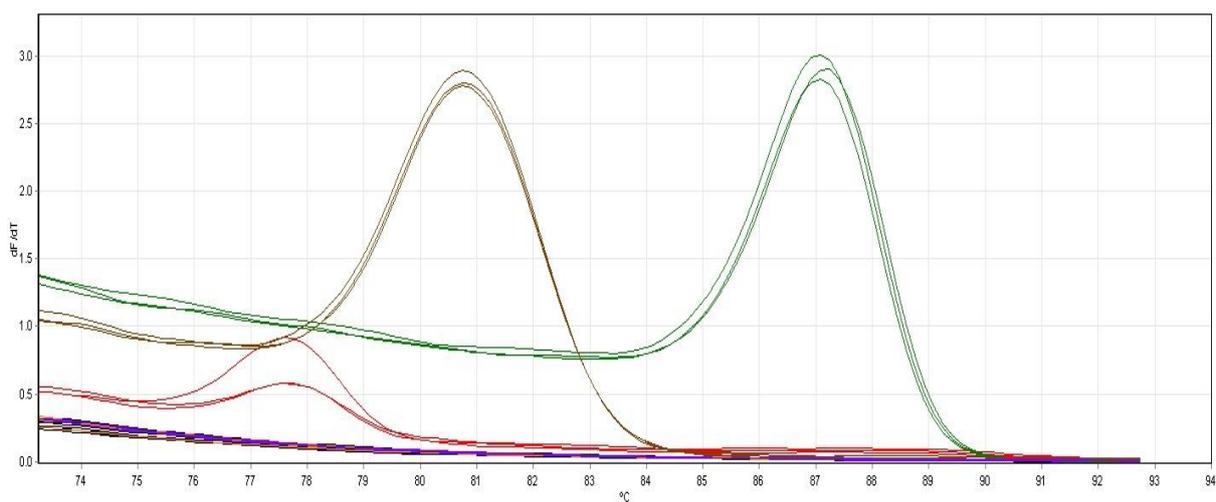
### 4.3. Results

#### 4.3.1. Detection of AOD bacterial DNA in soil

The quality and quantity of DNA isolated from pure cultures of AOD bacteria (Method 2.4) were assessed by Nanodrop (method 2.9). This DNA was then used in a multiplex HRM analysis (2.12) to ensure that positive identification of each bacterium could be obtained (Figure 25). 50 µL of pure DNA was spiked into 250 mg microcosms of soil which was then used for DNA extraction. The eluted soil DNA was analysed via the multiplex HRM method with the pure DNA run alongside as a positive control. The resulting HRM curves show that positive controls amplified while eluted DNA extract from soil did not (Figure 26).



**Figure 25:** Positive identification of *Brenneria goodwinii* (red), *Rahnella victoriana* (brown) and *Gibbsiella quercinecans* (green), using DNA extracted from pure cultures in a multiplex HRM analysis. Negative control included (purple).



**Figure 26:** Positive amplification of DNA from pure cultures of *Brenneria goodwinii* (red), *Rahnella victoriana* (brown) and *Gibbsiella quercinecans* (green), but negative amplification of the same respective species after spiking DNA into the soil and then performing DNA extraction on the whole soil sample.

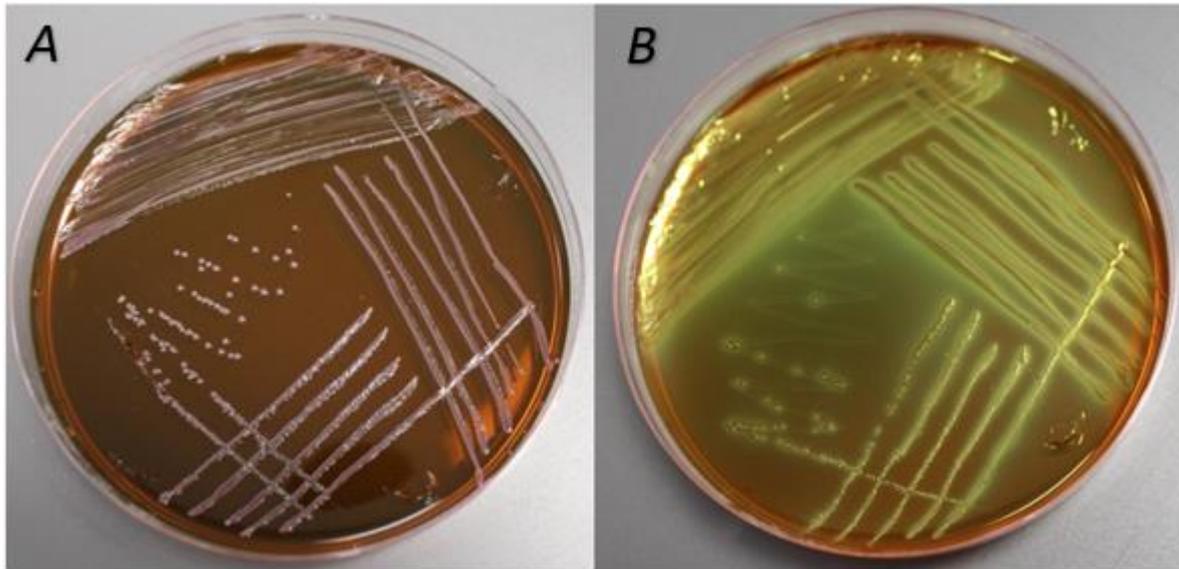
#### 4.3.2. EMB colour testing

While working with different *B. goodwinii* strains many appeared as white colonies on EMB, despite being frequently identified by their metallic green sheen on the agar. As such seven strains were assessed to see under which conditions *B. goodwinii* becomes metallic green on EMB. Strains were streaked out on EMB in triplicate and then incubated for 48 hours under both anaerobic and aerobic conditions. Table 14 shows that anaerobic conditions are required

for all *B. goodwinii* strains to turn green, while only two exhibited pigmentation under aerobic conditions. Example comparisons of the type strain have been provided showing how the conditions change the morphology of species (Figure 27).

**Table 14:** List of the strains of *Brenneria goodwinii* streaked on EMB agar and whether metallic green colonies were obtained after 48-hour incubation under aerobic and anaerobic conditions.

<b>Presence of metallic green colonies on EMB</b>		
<b>Strain</b>	<b>Aerobic incubation</b>	<b>Anaerobic incubation</b>
FRB 141 <sup>T</sup>	NO	YES
FRB 171	NO	YES
BH 4/25a	YES	YES
FRB 186	NO	YES
FRB 183	NO	YES
BH 1/28a	NO	YES
BH 1/28b	YES	YES

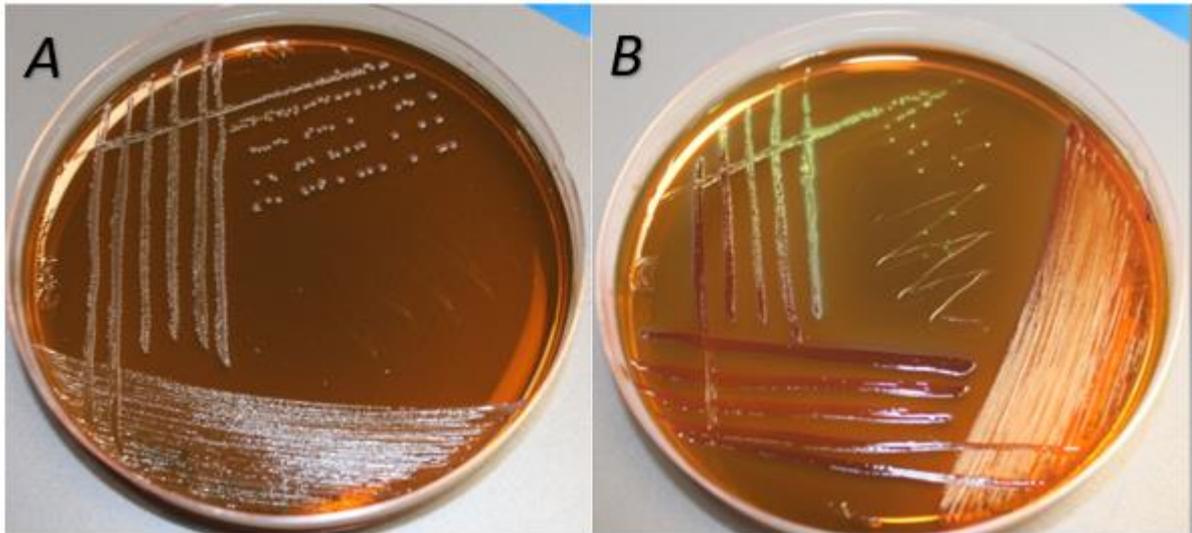


**Figure 27:** Images of *Brenneria goodwinii* type strain FRB 141<sup>T</sup> on EMB. Plate A was incubated under aerobic conditions and did not show any metallic green colonies, while plate B was incubated under anaerobic conditions and shows complete pigmentation to metallic green.

Due to the lack of pigmentation observed for the majority of *Brenneria* strains grown aerobically on EMB agar, the type strain for *Gibbsiella quercinecans* was tested. The type strain also did not appear metallic green so the same test was applied for all *G. quercinecans* strains. While three strains of *G. quercinecans* could aerobically turn metallic green on EMB, the majority also required anaerobic conditions to do so (Table 15). An example of the difference between colony appearance is provided below, showing the metallic green colour obtained in anaerobic conditions (Figure 28).

**Table 15:** List of the strains of *Gibbsiella quercinecans* streaked on EMB agar and whether metallic green colonies were obtained after 48-hour incubations under aerobic and anaerobic conditions.

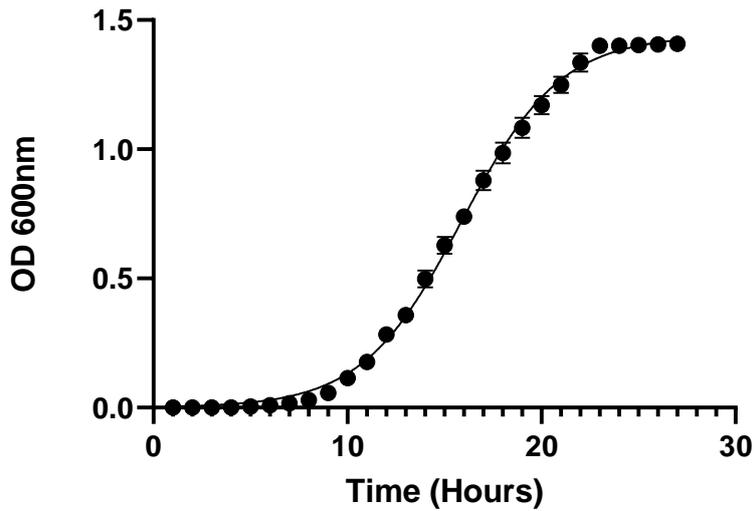
<b>Presence of Green colonies on EMB</b>		
<b>Strain</b>	<b>Aerobic incubation</b>	<b>Anaerobic incubation</b>
FRB 97 <sup>T</sup>	NO	YES
N79	NO	YES
BH 1/65b	YES	YES
FOD 9.25	NO	YES
Gq5	YES	YES
Gq6	NO	YES
Gq1	NO	YES
R-14	NO	YES
R-28	NO	YES
R-126	NO	YES



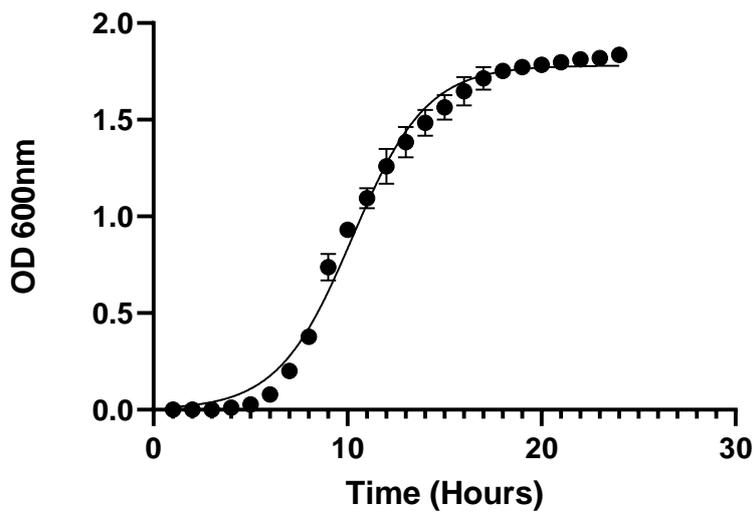
**Figure 28:** Images of *Gibbsiella quercinecans* strain Gq1 on EMB. The left plate was incubated under aerobic conditions and did not show any metallic green colonies, while the right plate was incubated under anaerobic conditions and shows complete pigmentation to metallic green.

#### 4.3.3. Growth curves for soil spiking experiments

Before spiking soil with known concentrations of bacteria it was important to know how long it takes to grow bacteria up to an OD of 0.5. This optical density was chosen as it guarantees bacteria to be in the exponential stage of growth and at their most metabolically active and hopefully most adaptable to a new medium, in this case, soil. Method 2.1.8 was followed for each bacterium (*B. goodwinii*, *G. quercinecans*, *R. victoriana* and *Pseudomonas laurylsulfatorans*). The growth curves are presented in Figure 29 - Figure 32.

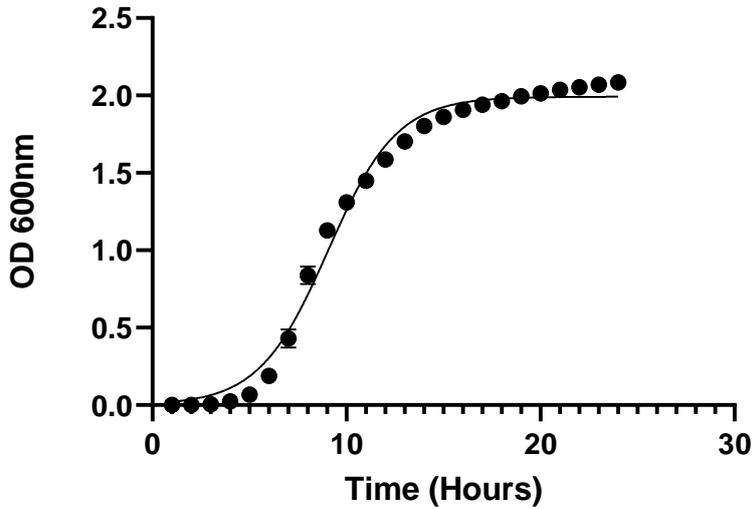


**Figure 29:** Growth of *Brenneria goodwinii* in LB batch culture over 26 hours. Optical density at 600 nm was measured at hourly intervals. Mean and SD are presented for two biological replicates.



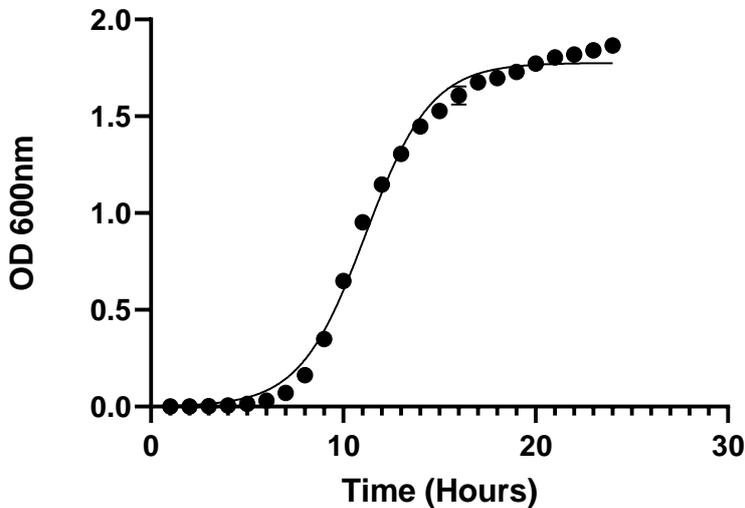
**Figure 30:** Growth of *Gibbsiella quercinecans* in LB batch culture over 26 hours. Optical density at 600 nm was measured at hourly intervals. Mean and SD are presented for two biological replicates.

### *Rahnella victoriana* Growth Curve



**Figure 31:** Growth of *Rahnella victoriana* in LB batch culture over 26 hours. Optical density at 600 nm was measured at hourly intervals. Mean and SD are presented for two biological replicates.

### *Pseudomonas laurylsulfatorans* Growth Curve



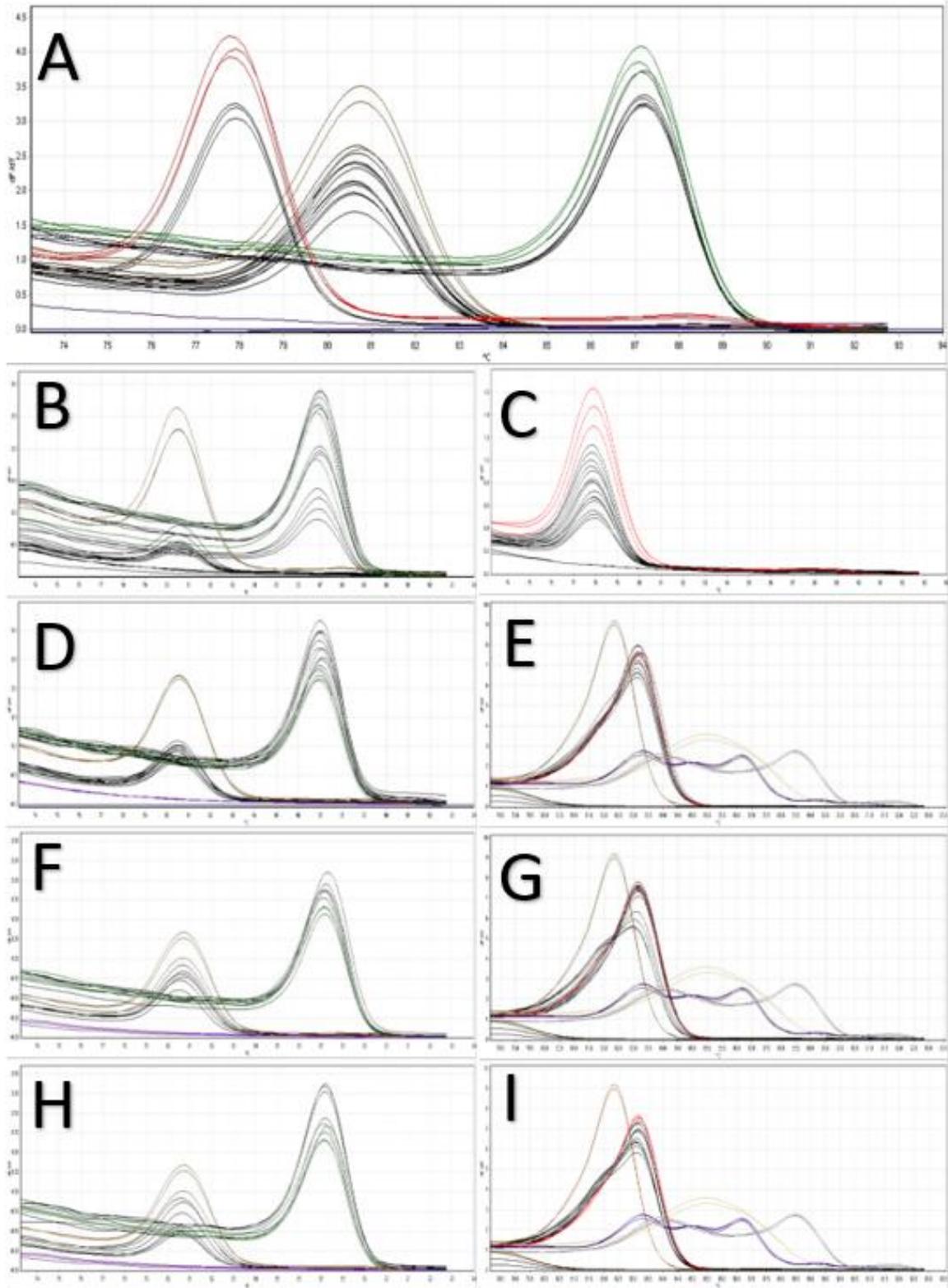
**Figure 32:** Growth of *Pseudomonas laurylsulfatorans* in LB batch culture over 26 hours. Optical density at 600 nm was measured at hourly intervals. Mean and SD are presented for two biological replicates.

All species excluding *B. goodwinii* reached the stationary phase within 24 hours, with the majority doubling their optical density every hour after the initial lag phase. The time taken to reach an optical density of 0.5 at 600 nm differed for each bacterium with *B. goodwinii*

being the slowest at nearing 15 hours. *R. victoriana*, *G. quercinecans* and *P. laurylsulfatorans* reached an OD of 0.5 required for spiking in 9, 7 and 10 hours, respectively.

#### 4.3.4. Survivability of AOD bacteria in soil

Recent publications have provided findings that bacteria isolated from the lesions of AOD symptomatic trees are not able to survive in the soil matrix (Pettifor *et al.*, 2020). To assess this *Brenneria goodwinii*, *Gibbsiella quercinecans*, *Rahnella victoriana* and *Pseudomonas laurylsulfatorans* were spiked into the soil environment with method 2.1.10. Initially, alkaline lysis was attempted on cultures plated directly from the undiluted Ringers soil suspension to isolate mixed genomic DNA for HRM detection of AOD pathogens from mixed cultures. Identification using this method was not possible due to the resulting unpredictable spikes from nonspecific amplification during HRM analysis. This meant that detection of AOD pathogens from mixed whole plates would not be possible, limiting the application of detection in high numbers of samples. Instead, single colonies would have to be re-streaked on both EMB and LB agar after which they were re-incubated for 48 hours at 28 °C. Finally, the identity of the single isolates was confirmed using the multiplex HRM method (2.2.13) following alkaline lysis (Figure 33).



**Figure 33:** A) High-resolution melt curve for *Brenneria goodwinii* (red) *Rahnella victoriana* (brown), and *Gibbsiella quercinecans* (green) after the soil was spiked at 0 hours using method 2.4 multiplex protocol 1 that identifies all four key species associated with AOD using specific primers for each species: spiked samples (black) show positive amplification of to each reference curve. B, D, F, H. Protocol 1: High-resolution melt curves for *R. victoriana* and *G.*

*quercinecans* for weeks 1, 2, 3 and 4 following soil spike at 0 hours. C) Protocol 1: High-resolution melt curve for *B. goodwinii* for week 1 following soil spike at 0 hours. E, G, I) *Brenneria* Protocol 3 that specifically identified the different species of *Brenneria*: High-resolution melt curves for *B. goodwinii* for weeks 2, 3 and 4 following soil spikes at 0 hours. All samples were run with a negative (purple) from sterile soil. All samples and controls were performed as replicates of n = 3.

The high-resolution melt curves generated weekly here show that all three species of AOD bacteria are capable of surviving in the soil matrix over a period of four weeks. Protocol 3 was used after week 2 as the multiplex identification for *B. goodwinii* stopped working, including amplification of the positive control. All negative controls remained negative throughout the testing period, demonstrating that none of the AOD bacteria were observed in non-sterile soil. No growth was recorded on any of the negative control sterile soil plates, showing that autoclaving had produced sterile samples.

#### 4.3.5. Survivability of AOD bacteria in EE broth

To ensure that specific enrichment could be used to increase the chance of recovery of AOD bacteria from soil, *B. goodwinii*, *G. quercinecans*, *R. victoriana* and *L. britannica* were inoculated into 15 ml of EE broth and left to grow for 48 hours in a 28 °C shaking incubator before plating out on EMB. All four bacteria were capable of growth in EE broth followed by culturing of EMB media.

#### 4.3.6. Survival and recovery of AOD bacteria oak-related niches

Due to the limitations of detecting the AOD bacteria in mixed cultures from soil identified in 4.2.4, an alternative method involving the enrichment of enteric bacteria was trialled. This method aimed to reduce the number of bacteria isolated, reducing competition and therefore increasing the potential for identification of the AOD lesion bacteria. Due to the lack of AOD bacteria identified from soil collected from Hatchlands park using the method used for the artificial spiking experiment (4.3.4), acorns and leaves were also tested as alternative (endophytic) sources of the AOD bacteria.

To assess if the EE recovery method was suitable for the identification of the members of the AOD lesion microbiome from different oak-related materials, the inoculation method described in method 1.8 was used to spike six replicates of 10 g of soil, acorns and leaves. Miles and Misra triplicate drop plates were made for each bacterium to identify the CFU being spiked into each microcosm (Table 16). Six of each niches were spiked for *B. goodwinii* (Bg),

*G. quercinecans* (Gq), *R. victoriana* (Rv) and *L. britannica* (Lb), as well as six of each niche being withheld from inoculation to act as a negative control. Each week one acorn, leaf and soil sample were suspended in EE broth as described in 2.1.7 and incubated for 48 hours at 28 °C while shaking at 250 RPM. Following 48-hour enrichment, undiluted samples were plated on EMB, GAM and R2A media under both anaerobic and aerobic conditions. DNA extraction was performed by alkalic lysis on a loopful of culture from undiluted samples on all plates. The resulting DNA was then used in a multiplex HRM analysis (2.5), and high-resolution melt curves were generated to identify AOD bacteria that had been successfully re-isolated from each niche. Table 17 details the bacteria that were re-isolated from each niche per week.

**Table 16:** The colony forming units calculated from Miles & Misra plates (n=3) for each species used in the spiking experiments.

Bacterium	Average CFU/mL from OD 0.5
<i>B. goodwinii</i>	$1.3 \times 10^9$ CFU ml <sup>-1</sup> ± $0.3 \times 10^9$
<i>G. quercinecans</i>	$8.6 \times 10^8$ CFU ml <sup>-1</sup> ± $1.2 \times 10^8$
<i>R. victoriana</i>	$2.3 \times 10^9$ CFU ml <sup>-1</sup> ± $0.4 \times 10^9$
<i>L. britannica</i>	$5.1 \times 10^9$ CFU ml <sup>-1</sup> ± $1.3 \times 10^9$

**Table 17:** The identities of AOD bacteria recovered from the spiked soil, acorns and leaves. Bacteria were spiked into the three niches and sampled at weekly intervals over the course of six weeks before being identified by HRM in triplicate. \* = pre-enrichment of samples in BPW. - = no detection of AOD pathogens in samples. In many columns more than the artificially spiked bacteria have been identified in samples indicating that these AOD bacteria were present prior to spiking. This is reinforced by the identification of AOD bacteria in large numbers of the negative control samples.

Week one	<i>Brenneria goodwinii</i>	<i>Gibbsiella quercinecans</i>	<i>Rahnella victoriana</i>	<i>Lonsdalea britannica</i>	Negative control
Soil	Bg*, Rv	Gq	Rv	Lb*, Rv	Rv
Acorn	Bg, Lb, Gq	Lb, Gq	Rv	Lb	Lb

Leaf	Bg, Rv	Gq	Rv, Gq, Bg	Lb	Rv, Lb, Gq
<b>Week 2</b>					
Soil	Bg*	Gq, Rv	Rv	Lb*	Rv
Acorn	Bg, Rv	Gq	Rv	Lb	-
Leaf	Bg	Gq, Rv, Bg	Rv, Gq, Bg	Lb, Rv	-
<b>Week 3</b>					
Soil	Bg*	Gq	Rv	Lb*	Rv
Acorn	Bg, Lb	Gq, Bg	Rv, Lb	Lb	Lb
Leaf	Bg, Rv	Gq, Rv	Rv, Gq, Bg	Lb, Rv	Lb
<b>Week 4</b>					
Soil	Bg*	Gq	Rv	Lb*	Rv
Acorn	Bg	-	Rv	Lb	
Leaf	Bg	Gq	Rv, Gq, Bg	Lb, Gq, Rv, Bg	Lb
<b>Week 5</b>					
Soil	Bg*,	Gq	Rv, Gq	Lb*	Rv
Acorn	Bg	Gq	Rv	Rv	Gq
Leaf	Bg, Gq	Gq	Bg, Gq	Lb	Lb
<b>Week 6</b>					
Soil	Bg*	Gq, Rv	Rv	Lb*	Rv

Acorn	Gq	Gq	Bg, Gq	Lb	-
Leaf	Bg	Gq, Rv	Bg, Gq	Lb	-

While the use of EE is suitable for the recovery of all the key AOD lesion bacteria from both acorns and leaves, both *B. goodwinii* and *L. britannica* were initially not recovered from the soil after the 0-hour time point. However, including a pre-enrichment step in buffered peptone water (BPW), allowed recovery of *B. goodwinii* and *L. britannica* for the full six weeks. Each spike was successful for the recovery of the target organism excluding the *B. goodwinii* spiked acorn tested in week 6, *G. quercinecnans* acorn spike in week 4, *R. victoriana* leaf spikes in weeks 5 and 6, *R. victoriana* acorn spike in week 6 and *L. britannica* acorn spike week 5. It appears *B. goodwinii* and *R. victoriana* were lost from their acorn spike niches after weeks 5 and 4 respectively, while *R. victoriana* was also lost from the leaf spike niche following week 5. The remaining negative results indicate a different issue, as the bacteria were isolated from the same niche in the following weeks.

The results also show more than just the original spiked AOD bacteria were recovered from many niches. *R. victoriana* was recovered from nine leaf and two acorn samples which it was not spiked into, as well as being consistently recovered from the negative control soil samples. *G. quercinecnans* was recovered from one soil sample, four acorn samples and nine leaf samples which it was not spiked into. *L. britannica* was recovered from six extra acorn samples and four-leaf samples. Finally, *B. goodwinii* was identified in four leaf samples and one acorn which it was not spiked into.

#### 4.3.7. Comparison of single isolate identities from both recovery methods

Following recovery in both Ringer's and EE broth, 16S rRNA and *gyrB* gene sequencing was performed on single colony isolates from two Hatchlands Park samples. The aim was to identify which method could be used in unison with HRM screening for the classification of potentially interesting novel isolates that showed a preference for healthy or diseased rhizosphere soil. Suspension of samples 7E, 7N, 8N, 8W, 13N, 13W, 14E and 14W led to the 103 Ringer's isolates and 96 EE isolates. A table detailing the EzBiocloud and BlastN results can be seen in Suppl. Table 1. The Ringer's suspension method led to the isolation of a

consistent set of bacteria regardless of the sample. 48 % of these bacteria belonged to the genera *Pseudomonas*, *Bacillus* and *Stenotrophomonas*, with the remaining 52 % distributed over other genera such as the *Streptomyces* and *Paenibacillus*. EE isolations, however, showed more division by the tree sample with H7 being dominated by *Serratia*, H8 containing more *Citrobacter*, and uncultured bacteria, H13 *Rhodococcus*, *Buttiauxella* and *Escherichia*, and H14 *Rahnella* and *Serratia*.

#### 4.3.8. HRM analysis of bacteria from Hatchlands Park samples

From the results obtained through sections 4.2.4-4.2.6, an appropriate testing method for the detection of AOD pathogens in soil, leaves and acorns was decided upon. The method described in 1.6 in which samples were manually disrupted pre-enriched in BPW, enriched using EE broth and plated on EMB was followed by alkalic lysis of a loopful of culture from undiluted samples. Extracted DNA underwent multiplex HRM in triplicate, Table 18 presents the identification for each sample.

**Table 18:** The AOD bacteria identified in environmental samples from Hatchlands Park, Guildford, UK. Bg = *Brenneria goodwinii*, Gq = *Gibbsiella quercinecans*, Lb = *Lonsdalea britannica* and Rv = *Rahnella victoriana*. Samples that could not be collected are marked by N/A and samples without no AOD bacteria identified are marked by -. Tree identity is broken up by tree number (H#) and P = parkland, W = woodland, H = healthy and D = diseased.

Tree identity	Soil	Leaves	Acorn
H1 PD	Rv	N/A	N/A
H2 PH	Rv	-	Gq
H3 PD	Rv	-	Lb, Gq
H4 PH	Rv	-	-
H5 PD	Rv	-	-
H6 PH	Rv	-	-
H7 PD	Rv	Rv	Rv

H8 PH	Rv	Rv	Rv
H9 PD	Rv, Gq	Rv	Rv
H10 PH	Rv	-	Rv
H11 WD	-	Gq, Rv	Gq, Rv
H12 WH	Rv	-	N/A
H13 WD	Rv	-	Gq
H14 WH	Rv, Gq	-	Rv Gq
H15 WD	Rv	N/A	Lb, Gq
H16 WH	-	-	Gq
H17 WD	Rv, Gq	-	N/A
H18 WH	Rv	-	Lb, Rv
H19 WD	Rv	Rv	N/A
H20 WH	Rv	Rv	Bg, Lb, Gq

The results from the Hatchlands Park samples demonstrate similar patterns to the results from the spiking experiment. *R. victoriana* can be isolated from all three niches, being present in 18 of 20 soil samples, and the most commonly isolated AOD bacteria from leaves and acorns. *G. quercinecans* also appears to be environmentally widespread being isolated in soil, though far less frequently. However, it is more frequently isolated as an acorn endophyte than as a member of the phyllosphere. *L. britannica* and *B. goodwinii* are the least frequently isolated species from oak-related material within this study, only being isolated from acorns from both healthy and diseased trees.

#### 4.3.9. *gyrB* identification of EE isolates

While performing the HRM analysis of oak isolates, rhizosphere EE enrichments were serially diluted to produce mixed colony plates which were then repeatedly restreaked for the isolation of single colonies. Glycerol stocks were made for the single colonies and subsequently re-cultured onto LB agar. Overall, 505 isolates were obtained from the 40 enriched rhizosphere soil samples. *gyrB* sequencing was performed on all isolates, which could be split into six groups following identification.

1. *Buttiauxella* and *Cedecea*
2. *Citrobacter*, *Klebsiella*, *Raoultella*, *Kluyvera* and *Lelliottia*
3. *Escherichia*
4. *Leclercia* and *Scandinavium*
5. *Rahnella*
6. *Serratia*

From the 1000 bootstrap maximum likelihood trees generated for each group, the isolates shown below in Table 19 were chosen for MLSA to identify novel species which were taxonomically described in later work.

**Table 19:** The Genus group and the number of isolates in each group. Isolates are further separated by their BlastN identity based on the *gyrB* gene and the percentage identity from the Blast result.

Genus group	Blast result	Percentage identity (%)
<b><i>Buttiauxella</i>: 11</b>		
H11S14	Uncultured bacterium 3AFRM03	95.96
H13N10	Uncultured bacterium 3AFRM03	95.96
H7N4	Uncultured bacterium 3AFRM03	95.28
H19S16	<i>Rhodococcus erythropolis</i> / Uncultured bacterium 3AFRM03	96.50/95.96

H6S16	<i>Rhodococcus erythropolis/</i> bacterium 3AFRM03	Uncultured	96.50/95.97
H4E8	<i>Rhodococcus erythropolis/</i> bacterium 3AFRM03	Uncultured	96.36/95.82
H15E14	<i>Rhodococcus erythropolis/</i> bacterium 3AFRM03	Uncultured	96.36/95.83
H2N9	<i>Rhodococcus erythropolis/ Rahnella aquatilis</i>		99.60/99.29
H3N1(2)	<i>Rhodococcus erythropolis/ Rahnella aquatilis</i>		99.46/99.06
H8N7	<i>Rhodococcus erythropolis/ Rahnella aquatilis</i>		99.73/99.53
H5W8	<i>Rhodococcus erythropolis/ Rahnella aquatilis</i>		99.46/99.06
<b>Cedecea: 13</b>			
H11S18	3AFRM03/ <i>Cedecea neteri</i>		89.22/89.08
H18E4	3AFRM03/ <i>Cedecea neteri</i>		89.95/88.81
H9W11	3AFRM03/ <i>Cedecea neteri</i>		89.22/89.08
H6S10	3AFRM03/ <i>Cedecea neteri</i>		89.22/89.08
H16N7	3AFRM03/ <i>Cedecea neteri</i>		89.22/89.08
H18W14	3AFRM03/ <i>cedecea neteri</i>		89.49/89.23
H19S8	3AFRM03/ <i>cedecea neteri</i>		89.35/89.23
H6W4	3AFRM03/ <i>cedecea neteri</i>		89.35/89.23
H20N1	3AFRM03/ <i>cedecea neteri</i>		89.41/89.16
H1W6	3AFRM03/ <i>cedecea neteri</i>		89.49/89.37
H14W9	3AFRM03/ <i>cedecea neteri</i>		89.49/89.37
H15E8	3AFRM03/ <i>cedecea neteri</i>		89.49/89.37
H20N9	3AFRM03/ <i>cedecea neteri</i>		89.49/89.37
<b>Rahnella: 10</b>			

H16E12	<i>Rahnella</i> sp. CB-2021b/ <i>R. variigena</i>	98.11/97.71
H11S4	<i>Rahnella</i> sp. CB-2021b/ <i>R. variigena</i>	97.94/97.63
H9E2	<i>Rahnella</i> sp. CB-2021b/ <i>R. variigena</i>	98.25/97.57
H11S5	<i>Rahnella</i> sp. CB-2021b/ <i>R. variigena</i>	98.10/97.47
H16E13	<i>Rahnella</i> sp. CB-2021b/ <i>R. variigena</i>	98.11/97.44
H18E1	<i>Rahnella</i> sp. CB-2021b/ <i>R. variigena</i>	98.11/97.44
H13N2	<i>Rahnella aquatilis</i> / <i>R. variigena</i>	92.45/92.23
H20N7	<i>Rahnella aquatilis</i> / <i>R. variigena</i>	92.18/92.18
H7N8a	<i>Rahnella aquatilis</i> / <i>R. variigena</i>	92.18/92.18
H12E4	<i>Rahnella aquatilis</i> / <i>R. variigena</i>	91.91/91.91
<b>Serratia: 6</b>		
H13W7	<i>Serratia quinivorans</i>	97.04
H12W9	<i>Serratia quinivorans</i>	97.04
H1S10	<i>Serratia inhibens</i> / <i>S. plymuthica</i>	99.73/99.06
H2N10	<i>Serratia inhibens</i> / <i>S. plymuthica</i>	99.73/99.06
H6S14	<i>Serratia inhibens</i> / <i>S. plymuthica</i>	100/99.06
H3E7	<i>Serratia inhibens</i> / <i>S. plymuthica</i>	99.60/98.65
<b>Leclercia: 13</b>		
H10E8	<i>Leclercia adecarboxylata</i> / <i>Leclercia</i> sp. LSNIH3	99.46/99.46
H9E1a	<i>Leclercia adecarboxylata</i> / <i>Leclercia</i> sp. LSNIH4	99.46/99.46
H10E4	<i>Leclercia adecarboxylata</i> / <i>Leclercia</i> sp. LSNIH5	99.46/99.46
H6S9	<i>Leclercia</i> sp. 29361/ <i>Leclercia adecarboxylata</i>	99.46/96.63

H6W5	<i>Leclercia</i> sp. 29361/ <i>Leclercia adecarboxylata</i>	99.46/96.63
H6S3	<i>Leclercia</i> sp. 29361/ <i>Leclercia adecarboxylata</i>	96.50/95.15
H20N5	<i>Leclercia</i> sp. 29361/ <i>Leclercia adecarboxylata</i>	96.63/95.28
H6W6a	<i>Leclercia</i> sp. 29361/ <i>Leclercia adecarboxylata</i>	96.36/95.01
H6W8	<i>Leclercia</i> sp. 29361/ <i>Leclercia adecarboxylata</i>	96.63/95.28
H4N4	<i>Leclercia</i> sp. 29361/ <i>Leclercia</i> sp. 119287	93.13/92.99
H18E8	<i>Leclercia</i> sp. 119287/ <i>L. adecarboxylata</i>	93.26/93.26
H19S6	<i>Leclercia</i> sp. 119287/ <i>L. adecarboxylata</i>	93.40/93.40
H17S15	<i>Scandinavium goeteborgense</i> / <i>Kluyvera intermedia</i>	92.60/90.19
H11S7	<i>Scandinavium goeteborgense</i> / <i>Leclercia</i> sp. 119287	93.53/90.04
<b>Scandinavium: 4</b>		
H4N3	<i>Scandinavium goeteborgense</i> / <i>Kluyvera ascorbata</i>	97.98/89.99
H5W4	<i>Scandinavium goeteborgense</i> / <i>Kluyvera ascorbata</i>	97.84/89.85
H4E14	<i>Scandinavium goeteborgense</i> / <i>Kluyvera ascorbata</i>	97.71/90.26
H5W7	<i>Scandinavium goeteborgense</i> / <i>Kluyvera ascorbata</i>	97.71/89.99
H5W5	<i>Scandinavium goeteborgense</i> / <i>Kluyvera ascorbata</i>	97.84/89.95

#### 4.4. Discussion

Bacteria have well-noted limitations on their range of habitats due to spatial environmental heterogeneity (Walters and Martiny, 2020). The first section of the work presented in this chapter looks at how bacteria can be detected from different environmental sources. DNA extraction followed by HRM detection was not sufficient for the detection of AOD bacteria in soil, independent of their direct presence or the presence of their pure DNA. Detection of bacteria using differential media such as EMB was not suitable, even for preliminary identification due to the inconsistency in the colour change of colonies from red/black/white to metallic green. The majority of *B. goodwinii* and *G. quercinecans* strains required anaerobic conditions for the fermentation of lactose to occur, leading to the development of a metallic green colour on EMB. As such, AOD bacterial colonies could not be visually distinguished from others for further confirmation using HRM. Instead, plates made from undiluted EE suspension would be the most viable way to detect AOD bacteria from soil and other samples.

Recently it was indicated that *B. goodwinii* could not be recovered from soil following inoculation, suggesting it enters a viable but non-culturable state and as such, shows a limited range of environments outside of its host (Pettifor *et al.*, 2020). Here conclusive evidence is provided that this is not the case. *B. goodwinii*, alongside *G. quercinecans*, *L. britannica* and *R. victoriana* are all capable of survival following inoculation into soil and remain recoverable for four and then six weeks in two separate experiments. The initial four-week experiment required nothing more than disrupting soil in Ringer's solution, followed by both selective and non-selective plating on solid media. However, the method ran into severe limitations due to the inability to use undiluted Ringer's suspensions for the detection of the AOD bacteria. Melt curves generated showed spikes differing from those of the reference strains so that individual specific melts for the AOD bacteria could not be observed. This was presumably due to non-target amplification from the variety and number of taxa isolated through non-specific isolation. As such single colonies had to be isolated from plates made from serially diluted Ringer's suspensions, increasing the time for the workflow and labour required to identify AOD bacteria in samples. During the experiment these issues were minor, however, when applying the workflow to the rhizosphere samples from Hatchlands Park it would be unfeasible to screen every isolate. Although the method was sufficient to identify the *B. goodwinii*, *G. quercinecans* and *R. victoriana* from the soil over the course of four weeks, a

separate method, which would allow for the identification of all four AOD bacteria from environmental samples, without the need for dilutions and single colony isolations would be favourable.

The former family *Enterobacteriaceae* have often been the focus of investigation due to their pathogenic potential in humans. As such enrichment methods for their isolation from food are well documented, with a focus on the isolation of bacteria that have been left in viable but non-culturable states (Clarke, 2004). The low nutrient content of soil in some ways is comparable to this, in that bacteria can be shocked upon being put into higher nutrient environments, thus limiting their recovery (Aslam *et al.*, 2010). As such the use of recovery methods for the *Enterobacteriaceae* from food has a reasonable basis. The first step here was to observe the growth and detection of the AOD bacteria in EE broth, which was successful. The aim was then to use EE broth to specifically select the *Enterobacteriaceae*, increasing the chances of recovering the AOD bacteria while reducing the noise generated in the melt curve from non-specific binding in other taxa. Due to the lack of AOD bacteria detection in soil from the results in 4.2.4, it was also decided to include other oak-related material in the second spiking test to see if the method could be applied to other samples. Leaves and acorns were chosen, as aside from invasion through the rhizosphere, the phyllosphere and inherited seed endophytes are where endophytic bacteria originate from (Compant *et al.*, 2021; Rahman *et al.*, 2018). *B. goodwinii* and *L. britannica* were not recoverable in EE broth following the 0-hour time point unless suspended in BPW first. Whether this is due to the low nutrient environment of soil, leading them to enter the viable but non-culturable state as suggested by Pettifor *et al.*, 2020, or due to nutrient shock from the composition of EE broth (Aslam *et al.*, 2010), or a weakened state leaving them to be outcompeted in broth (Vos *et al.*, 2013), can only be speculated. Nutrient shock, followed by out competition seems most plausible, due to the ability of prior BPW suspension to enable the recovery of these species. Aside from this issue, the method appeared to be sufficient for the identification of all four AOD-related bacteria from each of the three niches. The method also did not suffer from the same pitfalls as those encountered with Ringer's suspension, with plates made from undiluted EE suspensions producing clear melts for the AOD bacteria in their sample. The most obvious benefit from this was the detection of multiple bacteria within samples, with species that were not spiked into samples being identified. This led to the identification of *R. victoriana* in

most sample types, *G. quercinecans* in many leaf samples and *L. britannica* in acorns, giving support to the testing of leaves and acorns alongside the Hatchlands Park rhizosphere samples.

While the EE recovery method had appeared to circumvent the issue of having to dilute the suspensions prior to plating and identifying single colonies for HRM analysis, the drawback of *B. goodwinii* and *L. britannica* showing reduced ability to be recovered from soil did raise some issues. A comparison of both methods for the isolation of novel bacteria for the taxonomic classification of unknown members of the oak rhizosphere in the next chapter was also of consideration. With only 1 % of bacteria cultivable under normal laboratory conditions, it is well-recorded that the use of enrichment methods for specific groups of bacteria greatly increases the recovery of traditionally hard-to-culture members of the soil microbiome (Nguyen *et al.*, 2018). The results provided a further basis to use EE broth, as a distinct isolation pattern arose with specific genera being identified in each of the representative trees. Moreover, the *gyrB* percentage identities from BlastN for the *Buttiauxella/Rhodococcus* isolates were low, averaging around 95 % indicating the potential for novel species. Finally, sample H14W and H14E showed a large number of *Rahnella* isolates, namely *R. variigena* and *R. victoriana* indicating the potential for the isolation of AOD-relevant species. Overall, despite the limitation identified through the spiking experiment, *Enterobacteriaceae* Enrichment was chosen as the most suitable method for the screening of Hatchlands Park samples for the AOD bacteria, while simultaneously being used to isolate species for further investigation.

Using the EE enrichment method, both *R. victoriana* and *G. quercinecans* were identified in oak rhizosphere soil, though *G. quercinecans* was only identified in three of the 20 samples screened and the presence of *R. victoriana* in 14 of 20 samples suggest a widespread distribution. Leaf endophytes gave lower numbers of recoveries with *R. victoriana* being isolated from six of 20 samples and *G. quercinecans* being isolated from only one. This is of less surprise as the phyllosphere is seen to be a hostile environment to microbial life with a thick physical barrier to internal entry and inhospitable conditions due to environmental fluctuations (Bashir *et al.*, 2022). Acorns, however, proved a consistent reservoir of the AOD bacteria with each member being identified at least once. *R. victoriana* and *G. quercinecans* were consistently isolated, while *L. britannica* was isolated four times in triplicate from 60

samples, twice from acorns from both healthy and diseased oaks. *B. goodwinii* was only clearly isolated once from the acorn of a healthy oak, though given the results from section 4.2.6 this may be more of a limitation of the recovery method, than an indication of their full distribution. While the study was small, being limited to 20 different trees from one site, the consistency in results allow us to hypothesize the origin of the AOD bacteria. While all bacteria appear to have the ability to survive in a range of environmental niches, *B. goodwinii* and *L. britannica* lack the competitive finesse to survive in the competitive, nutrient-deficient environment that soil offers. *G. quercinecans* and *R. victoriana* are competitive, flexible microbes which appear to have a wide environmental dispersal. However, all AOD bacteria appear to have the ability to be endosymbiotic in acorns. Further investigation, coupled with refinements to the enrichment method which increase the recovery of *B. goodwinii* may show a further relation of these bacteria to seed stock in oak. If they were to be inherited endosymbionts from acorns, their role in AOD would be purely a consequence of the environmental predisposition that leads to the tree no longer able to keep their pathogenic ability at bay, making them opportunistic pathogens. They may be, as many endosymbionts are, playing the role of latent pathogens which cause disease because the conditions allow for it (Turner, James and Poole, 2013).

The isolates from enrichment with EE obtained from the rhizosphere soil samples embodied a large proportion of this work. Sequencing of *gyrB* genes allowed 505 isolates to be identified, of which 58 appeared to belong to novel species with interesting associations with either healthy or diseased trees. This work was then further carried on in chapter 5 with the complete set of taxonomic classifications of the *Cedecea*, *Leclercia* and *Scandinavium*-like isolates.

Overall, this chapter presents a framework for the identification of the AOD bacteria from different environmental samples, by selecting the relevant bacteria at multiple different levels, their isolation and chance of detection are increased. Though not performed here, single colonies could also be isolated with this method. This methodology has led to the identification of AOD bacteria in several different samples and has begun to reveal their possible environmental reservoirs. However, the method is not perfect, with *B. goodwinii* seemingly lacking the competitive edge to consistently lead to isolation from real-world samples. Further improvements with agars, such as crystal violet pectate medium could be

used to increase the chance of identification. Alternatively, the use of a more sensitive identification method such as the Taqman probe could also increase the chance of identification. However, one major drawback of having to enrich the bacteria is that determination of cell density from samples is no longer possible reducing the analysis to qualitative rather than quantitative assessment.

Chapter 5. The Taxonomic  
Classification of Oak  
Rhizosphere Novel Species

The work in this chapter represents the taxonomic classification of the novel species from the oak rhizosphere and have been published in the following papers:

Maddock D, Arnold D, Denman S, Brady C. Description of a novel species of *Leclercia*, *Leclercia tamurae* sp. nov. and proposal of a novel genus *Silvania* gen. nov. containing two novel species *Silvania hatchlandensis* sp. nov. and *Silvania confinis* sp. nov. isolated from the rhizosphere of oak. BMC Microbiology. 2022;22:1–18.

Maddock D, Kile H, Denman S, Arnold D, Brady C. Description of three novel species of *Scandinavium*: *Scandinavium hiltneri* sp . nov ., *Scandinavium manionii* sp . nov . and *Scandinavium tedordense* sp . nov ., isolated from the oak rhizosphere and bleeding cankers of broadleaf hosts. Frontiers in microbiology. 2022; October:1–14.

Maddock D, Brady C, Denman S, Arnold D. Description of *Dryocola* gen. nov. and two novel species, *Dryocola boscaweniae* sp. nov. and *Dryocola clanedunensis* sp. nov. isolated from the rhizosphere of native British oaks. Systematic Applied Microbiology. 2023;46:12639

## 5.1. Introduction

Taxonomy has been a key feature of the journey to understanding the bacteria associated with Acute Oak Decline. All isolates identified from bleeding lesions, namely *Brenneria goodwinii*, *Gibbsiella quercinecans* and *Rahnella victoriana* underwent taxonomic classification using a polyphasic approach as discussed fully in Chapter 1 (section 1.6).

Bacterial taxonomy utilises a polyphasic approach to supply enough information on the identity and unique features of a species, that differentiate it from others. As such the current approach for taxonomic classification requires not only genotypic information but also, chemotaxonomic and phenotypic data (Raina *et al.*, 2019). These features have been utilised since the first models of systematic bacterial classification which were based on morphology, culturing conditions and biochemical and pathogenic characteristics to identify bacteria (Hugenholtz *et al.*, 2021). They remain essential, not only for fully defining a species, but for providing clarity in situations where genome based species delineation has caused taxonomic confusion (Palmer *et al.*, 2020). For example, when describing several new *Bradyrhizobium* species, the average nucleotide identity (ANI) values showed no clear differentiation between the novel species, nor members of the same species. Meanwhile, clear phenotypic traits (alongside other phylogenetic analysis) supported the formation of these new species (Ramírez-Bahena *et al.*, 2009). As such many taxonomists maintain that due to different evolutionary rates (a fundamental concept of evolutionary pressure), combined with the different computational procedures, genomic data must be used in conjunction with phylogenetic and phenotypic data to give a holistic, polyphasic taxonomic definition of new species (Rosselló-Móra and Amann, 2015). The combination of these features with genotypic data to best categorise biological organisms and their high level of variability, was fully realised as a polyphasic taxonomy later into the 1990s when 16S rRNA gene analysis became instrumental (Vandamme *et al.*, 1996).

Phylogenetic studies using T<sub>1</sub> RNase digestion of the 16S rRNA gene were first utilised to classify 10 methanogenic bacteria in 1977 (Fox *et al.*, 1977). The 16S rRNA genes fulfil many of the essential features of a phylogenetic marker and as such has been considered the gold standard for taxonomists for the last 30 years (Clarridge, 2004). All microorganisms contain a copy, averaging 1,500 bp, evolving slowly due to its important highly conserved nature as it

plays a critical role in translation, while showing a range of conserved, variable and hypervariable regions (Church *et al.*, 2020). This is further aided by widely available universal primers for the gene and growing amount of 16S rRNA gene data available through online repositories such as GenBank (Benson *et al.*, 2013; Coenye *et al.*, 1999). However, the highly conserved nature of the gene is also a limitation on the differentiation of certain families of bacteria, including the Enterobacterales. A meta-study using data from GenBank found that the phylogenetic trees generated were inconsistent in the taxonomic position of the species included. Consistent poor bootstrap values and Bayesian support were also observed (Naum, Brown and Mason-Gamer, 2008). Using the best representative sequence for each species available on GenBank, Church *et al.* performed a meta-study aligning and identifying the differences in the 16S rRNA gene which can be seen in Table 20. The results highlight the poor taxonomic resolution of the 16S rRNA gene within the Enterobacterales, where 10 % cannot be identified to species level and only a single mismatch is seen between the *Escherichia-Shigella-Pantoea-Klebsiella-Raoultella-Cronobacter* group of genera (Church *et al.*, 2020; Paradis *et al.*, 2005).

**Table 20:** Summary of ability of the 16S rRNA gene to identify clinically relevant Enterobacterales (formerly *Enterobacteriaceae*). Taken from Church *et al.*, 2020.

Genus	No. of sequences in the genus multisequence alignment	Total no. of tested positions	No. of identical positions	No. of divergent positions	% Identity
<i>Escherichia</i>	4	1,463	1,435	28	98.09
<i>Shigella</i>	4	1,539	1,530	9	99.42
<i>Pantoea</i>	13	1,424	1,332	92	93.54
<i>Klebsiella</i>	7	1,379	1,322	57	95.87
<i>Raoultella</i>	4	1,453	1,426	27	98.14

<i>Cronobacter</i>	7	1,548	1,499	49	96.83
<i>Enterobacter</i>	11	1,428	1,340	88	93.84
<i>Proteus</i>	5	1,466	1,448	18	98.77
<i>Citrobacter</i>	13	1,456	1,376	80	94.51
<i>Salmonella</i>	2	1,505	1,480	25	98.34
<i>Providencia</i>	9	1,436	1,370	66	95.40
<i>Cedecea</i>	3	1,466	1,446	20	98.64
<i>Edwardsiella</i>	5	1,549	1,537	12	99.23
<i>Hafnia</i>	3	1,415	1,371	44	96.89
<i>Serratia</i>	18	1,379	1,265	114	91.73
<i>Yersinia</i>	18	1,449	1,395	54	96.27

As such when investigating the order Enterobacterales a polyphasic approach which incorporates phylogenetics and phylogenomics is seen as the favourable option. Multilocus sequence analysis (MLSA) is a phylogenetic approach which uses partial sequences of conserved housekeeping genes to generate higher resolution phylogenetic trees, to illuminate evolutionary distances based on short allelic mismatches (Glaeser and Kämpfer, 2015b). Essential housekeeping genes are known to display a negative correlation between their level of expression and their evolutionary rate, a feature shown in all branches of life (Vieira-Silva *et al.*, 2011; Drummond and Wilke, 2008). However, to overcome the bias single gene sequences cause in phylogenies due to the distorting effects of recombination at a single locus, multiple genes must be used. In an ideal situation a set of genes would apply to prokaryotes allowing for recombination. However, this is not possible as genes in distantly

related taxa would be uninformative if present at all, and if they were conserved enough for one set of primers, would be unlikely to evolve at a sufficient speed for classification. Instead, it is more appropriate to use single copy genes that do not show selective advantages due to recombination or linkage, for distinguishing species within a family (Gevers *et al.*, 2005).

In recent years with reductions in price driving availability, high throughput whole genome sequencing has improved the understanding of the evolutionary relationship of bacteria, allowing for more powerful taxonomic delineation of species. Experimental evidence has long suggested that computational comparison of genomes is to be the new gold standard of prokaryotic taxonomy (Chun *et al.*, 2018). ANI and *in silico* DNA-DNA hybridisation (*isDDH*) are the most widely used, with both software and online tools available (Jain *et al.*, 2018). These tools have been used to identify *Candidatus* species, which cover novel bacteria not yet isolated, but identified based on genomic sequencing information alone (Overmann *et al.*, 2019). *Candidatus* species currently represent over 50 % of the microbial phyla in taxonomy databases, based on genomic information. However, they remain informally named as the International Code of Nomenclature of Prokaryotes requires a pure-culture type strain (Konstantinidis, Rosselló-Móra and Amann, 2017). In 2020 the International Committee on Systematics of Prokaryotes (ICNP) rejected the proposal to incorporate DNA sequences as nomenclature into the ICNP infrastructure (Sutcliffe, Dijkshoorn and Whitman, 2020). As such a new initiative called SeqCode, a new code of nomenclature based on DNA sequences as the type information has been launched (Hedlund *et al.*, 2022). However, this new code of nomenclature has not been fully implemented and instead the holistic approach that polyphasic taxonomy gives to identifying new species is still required (among other criteria) for valid recognition of any new species names by the International Journal of Systematic and Evolutionary Microbiology (IJSEM).

These are essential features to understand in the face of modern taxonomy which is rapidly changing in response to both sequencing and bioinformatic developments, and the order Enterobacterales is a perfect example of this. The single family Enterobacteriaceae in the order Enterobacterales was initially created in 1937 to contain a single genus *Enterobacter* and 112 species; many of which are now genera including *Escherichia*, *Shigella* and *Proteus* (Janda and Abbott, 2021; Rahn, 1937). While the genera in the family changed much over the following 40 years it was plagued by nomenclature issues, limited phenotypic data and non-

standardised testing. Much of this was remedied in the 1970s when the Centre for Disease Control and Prevention classified a number of known and novel taxa using a standardised polyphasic approach, involving ~200 phenotypic properties with DNA-DNA hybridisation and molecular G + C content comparisons (Farmer *et al.*, 1985). The most recent development in the taxonomy of the Enterobacterales involved the separation of the previously single-family order into seven new families: *Erwiniaceae* fam. nov., *Pectobacteriaceae* fam. nov., *Yersiniaceae* fam. nov., *Hafniaceae* fam. nov., *Morganellaceae* fam. nov., and *Budviciaceae* fam. nov. along with the family *Enterobacteriaceae*. This was based on the comparison of a number of phylogenetic and genomic features, namely phylogenetic construction based on 1548 core proteins, 53 ribosomal proteins and four MLSA genes combined with whole genome comparisons (Adeolu *et al.*, 2016a).

Aside from the environmental and clinical significance of this bacterial group, from a purely taxonomic standpoint the Enterobacterales represent a fascinating set of organisms. As such the work in this chapter presents a polyphasic taxonomic approach to classify several novel species and genera, isolated from rhizosphere samples of both healthy and AOD symptomatic oak. The aim of these taxonomic classifications was to further understand the role of the Enterobacterales in relation to native British oaks while further elucidating the evolutionary and environmental function of a rapidly changing and important group of bacteria.

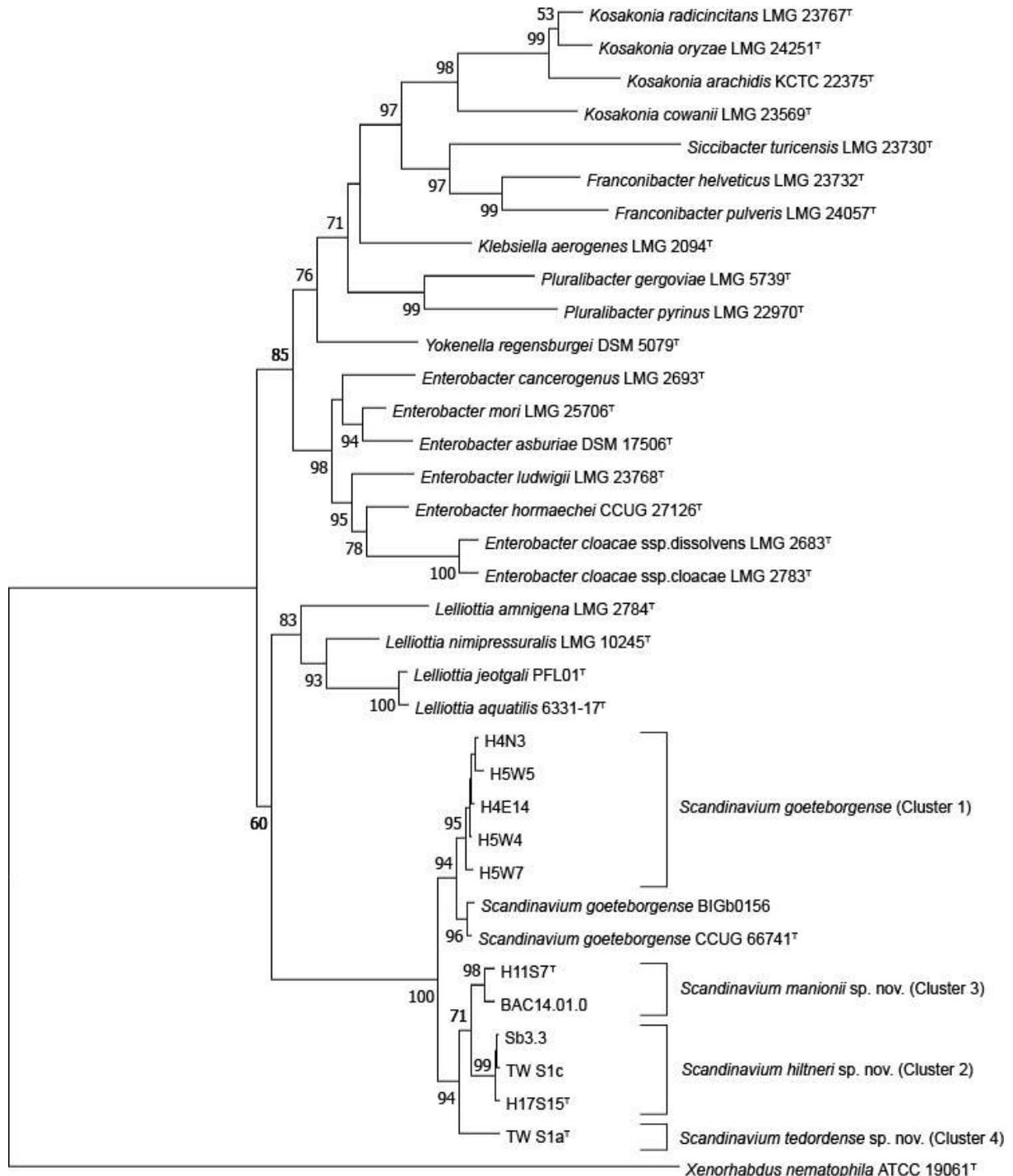
## 5.2. Results:

### 5.2.1. Genotypic characterisation:

#### *Scandinavium*

The concatenated MLSA maximum likelihood phylogenetic tree (Figure 34) shows all novel strains isolated in 4.2.9 included in the study were separated into four clusters in a single well-supported clade containing the type strain of *S. goeteborgense*. Cluster 1 constituted five strains from *Q. robur* rhizosphere soil, the type strain of *S. goeteborgense* (CCUG 66741<sup>T</sup>) and BIGb0156, isolated from rotting apple in France (Samuel *et al.*, 2016). The high bootstrap support and minor sequence variation within this cluster suggests these strains belong to *S. goeteborgense*. Cluster 2 contained one *Q. robur* rhizosphere strain, H17S15<sup>T</sup>, and two strains

isolated from *Tilia* spp. lesions, SB 3.3 and TWS1c. One *Q. robur* rhizosphere strain, H11S7<sup>T</sup>, and one strain isolated from *Q. rubra*, BAC 14–01-01, formed Cluster 3, while Cluster 4 constituted a single strain from a *Tilia x europaea* lesion, TWS1a<sup>T</sup>. The clear division from the *S. goeteborgense* cluster and high bootstrap support of Clusters 2 to 4 suggested that the strains belong to three potential novel *Scandinavium* species.

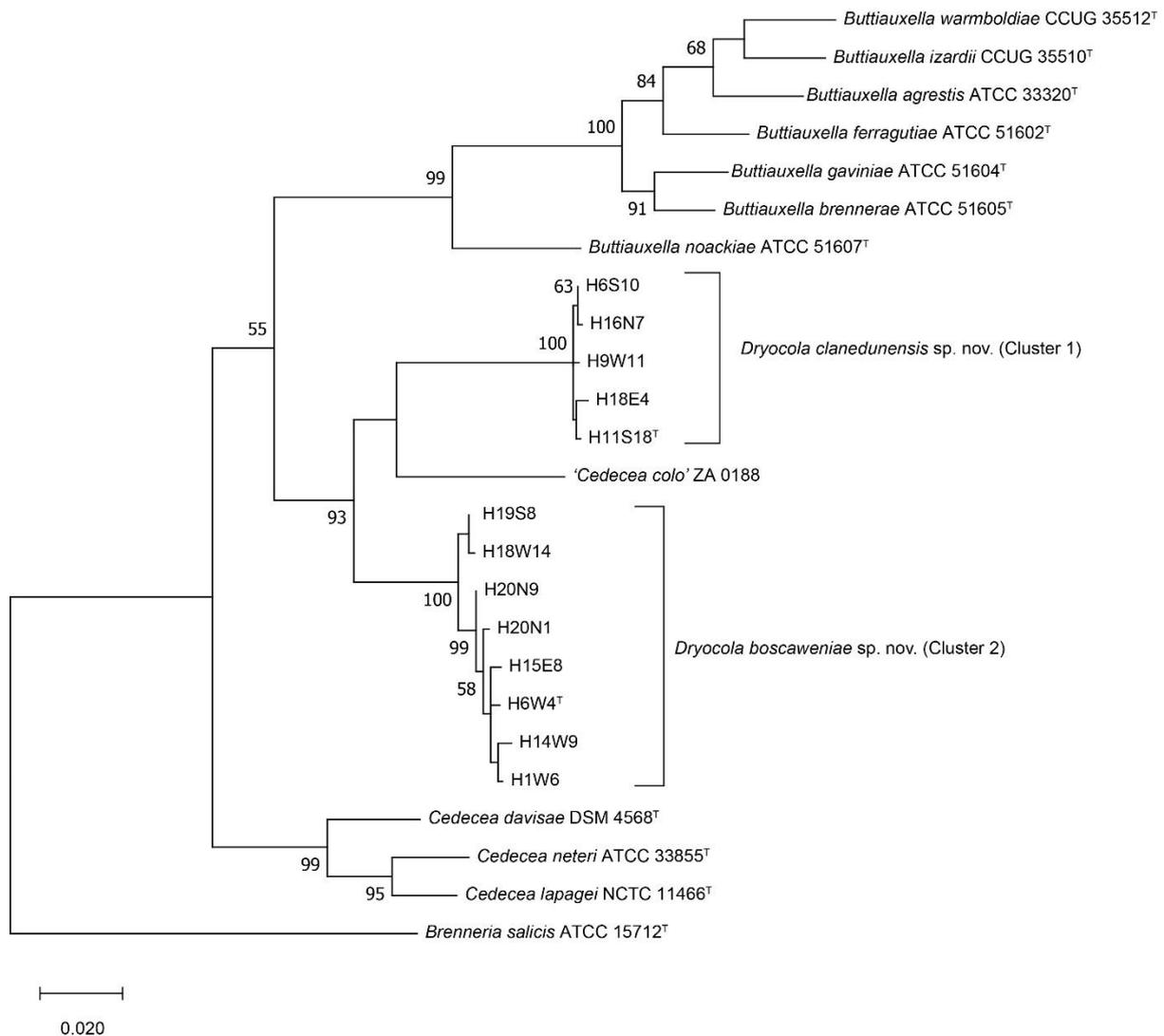


**Figure 34:** Maximum likelihood tree based on the concatenated partial gene sequences of *infB*, *atpD*, and *gyrB*. Sequences included are for the genus *Scandinavium*, the three proposed novel species, *Scandinavium hiltneri* sp. nov., *Scandinavium manionii* sp. nov. and *Scandinavium tedordense* sp. nov., as well as close phylogenetic neighbours. *Xenorhabdus nematophila* (ATTCC 190601<sup>T</sup>) was included as the outgroup. Percentages for bootstrap values exceeding 50 % following 1000 replicates are shown. The scale bar indicates the number of substitutions per site. Type strains are denoted via <sup>T</sup>.

The 16S rRNA gene comparisons demonstrated high pairwise similarity percentages of all three proposed type strains to *S. goeteborgense*, with 100 % completeness. TWS1a<sup>T</sup> was lowest at 99.03 % similar, while H17S15<sup>T</sup> showed 99.80 % similarity and H11S7<sup>T</sup> showed 99.86 % sequence similarity. Although the 16S rRNA gene comparisons indicated that the strains belong to *Scandinavium*, differentiation between any of the strains from each other and the type strain was not possible based on their 16S rRNA gene sequences. This is of little surprise as the 16S rRNA gene is poor at differentiating between species of *Enterobacteriaceae* due to its high homogeneity (Glaeser and Kämpfer, 2015b; Naum, Brown and Mason-Gamer, 2008). As such the 16S rRNA gene phylogenetic tree shown in Suppl. Fig. S1 does not accurately represent the position of the novel species, with low bootstrap values and loose clustering of the proposed novel type strains with the type strain of *S. goeteborgense*.

#### *Dryocola*

In the maximum likelihood phylogenetic tree based on the concatenated MLSA data (Figure 35), the strains isolated from oak rhizosphere soil (Chapter 4.2.9) were divided into two strongly supported clusters descended from the same node, with no validly published species present in the clade. The first cluster (Cluster 1) contained five isolates that showed little sequence variation, while the second cluster (Cluster 2) contained eight isolates with minor sequence variation, but both with 100 % bootstrap support suggesting two potential novel species. A recently described, but not yet validly published species, '*Cedecea colo*' (Boath *et al.*, 2020) was included in the clade on a separate branch on the border of Cluster 1, but with low bootstrap support. The position of both potential novel species clusters, in relation to their closest phylogenetic neighbours, suggests they belong to a novel genus. The inclusion of '*Cedecea colo*' in the potential novel genus clade suggests that this species may not belong to *Cedecea*, as the three validly published *Cedecea* species are contained in a separate monophyletic clade with 99 % bootstrap support.

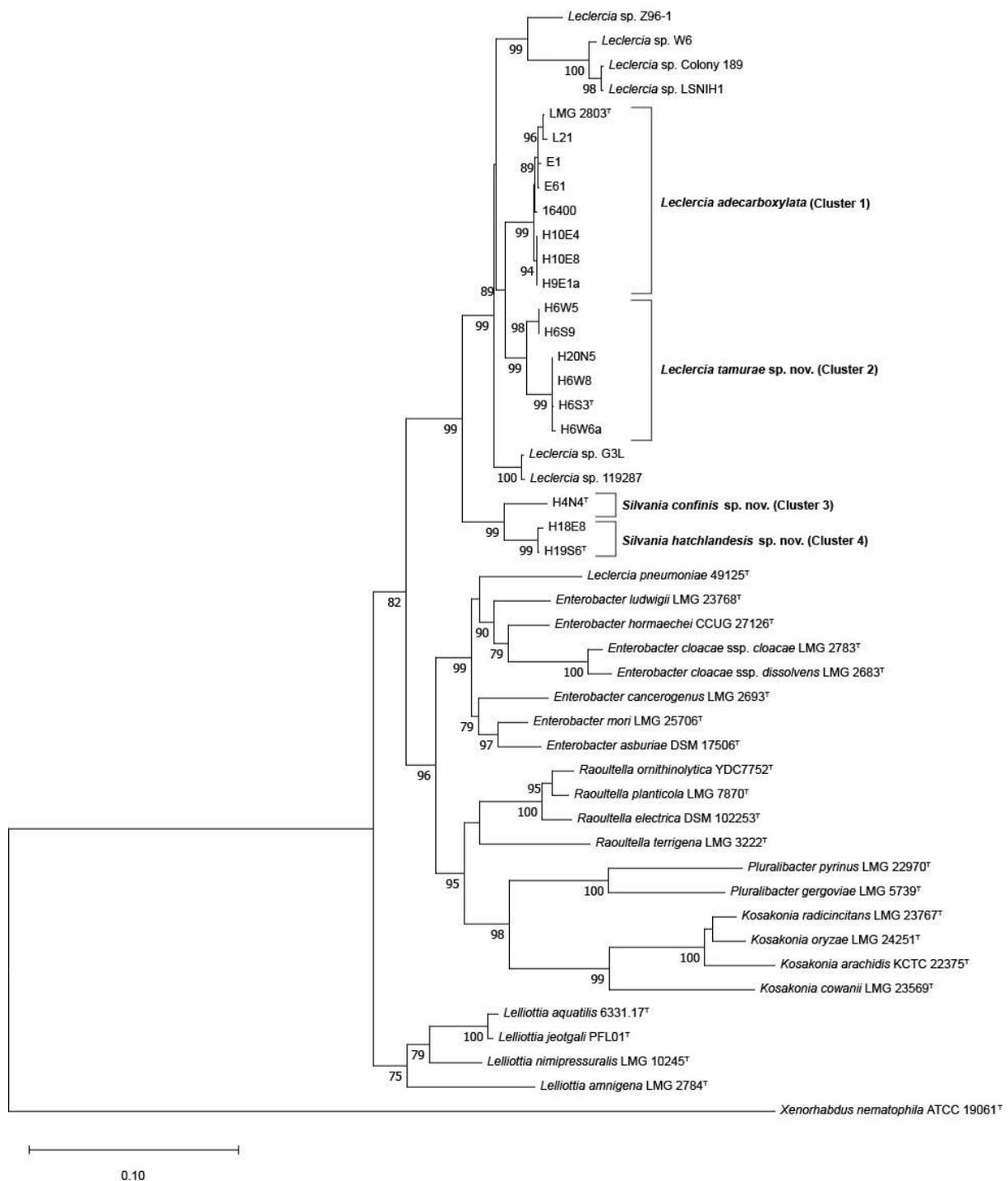


**Figure 35:** Maximum likelihood tree based on the concatenated partial gene sequences of *fusA*, *leuS*, *pyrG* and *rpoB* from species of the proposed genus *Dryocola* gen. nov., and its closest phylogenetic neighbours *Cedecece* and *Buttiauxella*. *Brenneria salicis* (ATCC 15712<sup>T</sup>) was included as the outgroup. Percentages for bootstrap values exceeding 50 % following 1000 replicates are shown. The scale bar represents the number of substitutions per site. Type strains are denoted via <sup>T</sup>.

H11S18<sup>T</sup> (Cluster 1) demonstrated highest 16S rRNA pairwise sequence similarity to *Cedecece neteri* (98.7 %), while H6W4<sup>T</sup> (Cluster 2) exhibited highest similarity to *Buttiauxella izardii* (98.6 %). Again, the taxonomic position of the proposed novel genus and species was not reliably represented by either the neighbour joining or maximum likelihood 16S rRNA gene phylogenetic trees (Suppl. Fig. S2 and Suppl. Fig. S3). In both trees the isolates clustered loosely with species of *Buttiauxella* and *Cedecece*, but with no, or very low, bootstrap support.

### *Leclercia and Sylvania*

In the maximum likelihood phylogenetic tree based on the concatenated MLSA sequences (Figure 36), the 12 strains were separated into four clusters. Cluster 1 contained three strains isolated from one healthy and one diseased oak tree, the type strain of *L. adecarboxylata* (LMG 2650<sup>T</sup>) and four strains identified as belonging to *Leclercia* whose genome sequences were downloaded from GenBank. Due to the lack of sequence variation and phylogenetic distance between these strains and the type strain of *L. adecarboxylata*, it was concluded that they belong to this species. Cluster 2, situated proximal to the *L. adecarboxylata* cluster, contained strains isolated from three cardinal points around two healthy oaks, one in the parkland and another in the woodland, and was strongly supported by a bootstrap value of 99 %, suggesting the strains belong to a novel species. Clusters 3 and 4 were contained in a clade with 99 % bootstrap support and consisted of one and two strains, respectively, isolated from both healthy parkland oak and diseased woodland oak rhizosphere soil. This clade was situated on a separate lineage on the border of the *Leclercia* clade with a greater phylogenetic distance, suggesting the strains could belong to a potential novel genus with two novel species. An additional six strains, identified as *Leclercia* sp. in GenBank, clustered on three separate lineages in the *Leclercia* clade (G3L and 119287; Z96-1 and W6; and Colony 189 and LSNIH1), suggesting they belong to several further potential novel *Leclercia* species. Of the six strains, Z96-1 has been incorrectly assigned to *L. adecarboxylata*, strain W6 was suggested as a novel species based on the computational analysis of its whole genome and the remaining four have yet to be classified at the species level. Additionally, based on MLSA, the taxonomic status of *Leclercia pneumoniae* 49125<sup>T</sup> was unclear, as it clustered on the border of the *Enterobacter* clade, far removed from *Leclercia*.



**Figure 36:** Maximum likelihood tree based on the concatenated partial gene sequences of *atpD*, *infB*, *gyrB* and *rpoB* from species of the proposed genus *Silvania* gen. nov., the novel species *Leclercia tamurae* sp. nov. and their closest phylogenetic neighbours. *Xenorhabdus nematophila* (ATCC 190601<sup>T</sup>) was included as the outgroup. Percentages for bootstrap values exceeding 50 % following 1000 replicates are shown. The scale bar represents the number of substitutions per site. Type strains are denoted via <sup>T</sup>.

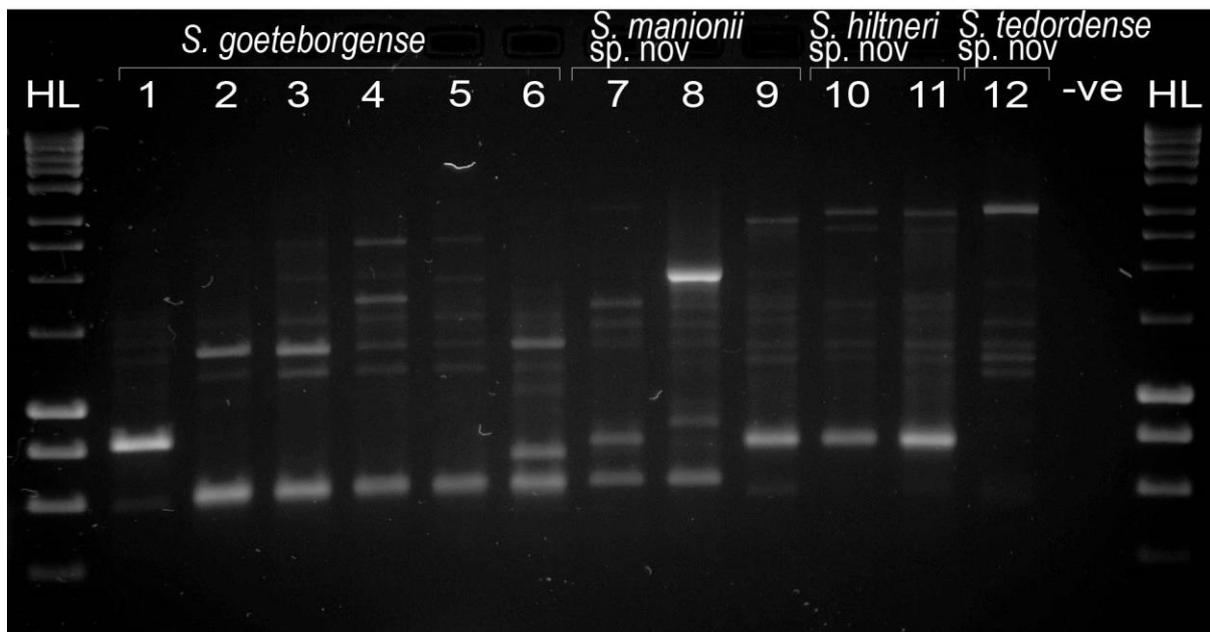
Strains H6S3<sup>T</sup> and H6W5 (Cluster 2) showed 99.40 – 99.55 % 16S rRNA gene pairwise sequence similarity to several *Enterobacter* species including *E. huaxiensis*, *E. cancerogenus*, *E.*

*sichuanensis* and *E. chengduensis* as well as 99.33 % to *L. adecarboxylata*. The strains suggested as belonging to a potential novel genus by MLSA, H4N4<sup>T</sup> (Cluster 3) and H19S6<sup>T</sup> (Cluster 4), displayed highest pairwise similarity to *Lelliottia jeotgali* with 99.48 % and 99.45 % to *L. adecarboxylata*, respectively and a generally high similarity to *Lelliottia* and *Enterobacter* species. This is reflected in the 16S rRNA gene maximum likelihood phylogenetic tree (Suppl. Fig. S4) where strains from the potential novel *Leclercia* species cluster within the *Enterobacter* clade, and species of the potential novel genus are situated on separate lineages in proximity to the *Lelliottia* clade.

### 5.2.2. Box PCR:

#### *Scandinavium*

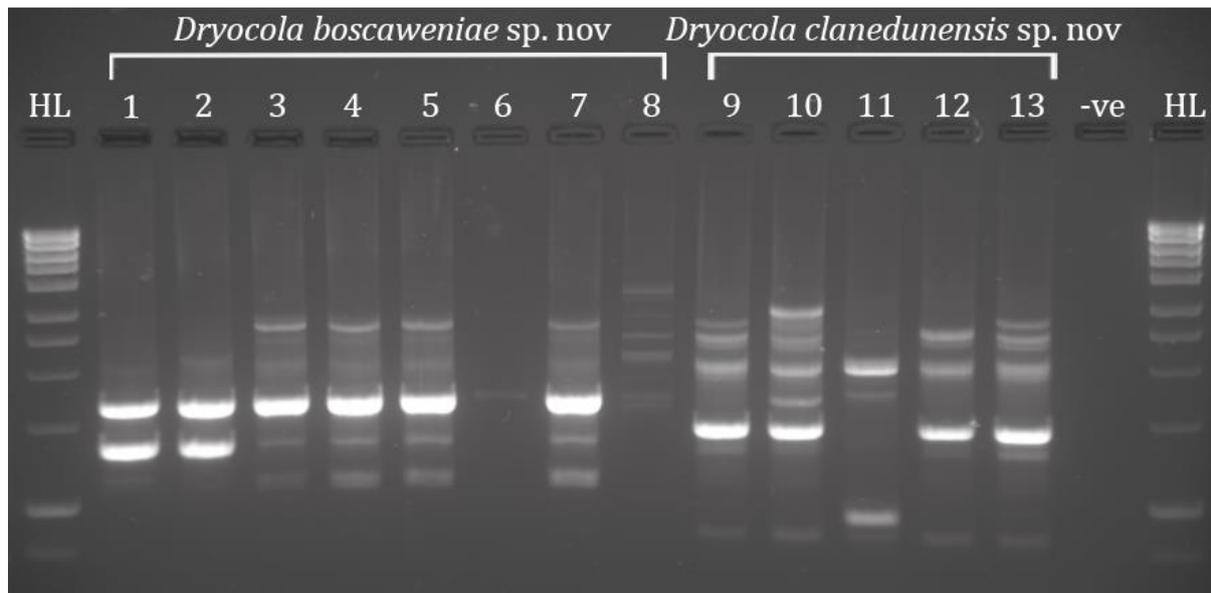
The BOX PCR results demonstrated notable levels of intra-species diversity within the samples as seen in Figure 37. All strains belonging to one of the proposed novel species could be differentiated from each other based on their banding patterns, indicating that the novel isolates were not clonal.



**Figure 37:** BOX PCR fingerprinting patterns generated from strains of the proposed novel species of *Scandinavium*. (1) *Scandinavium goeteborgense* CCUG 66741<sup>T</sup>, (2) *Scandinavium goeteborgense* H5W7, (3) *Scandinavium goeteborgense* H4E14, (4) *Scandinavium goeteborgense* H5W4, (5) *Scandinavium goeteborgense* H4N3, (6) *Scandinavium goeteborgense* H5W5, (7) *Scandinavium manionii* H17S15<sup>T</sup>, (8) *Scandinavium manionii* TWS1c (19) *Scandinavium manionii* SB 3.3, (10) *Scandinavium hiltneri* H11S7<sup>T</sup>, (11) *Scandinavium hiltneri* BAC 14-01-01 (12) *Scandinavium tedordense* TWS1a<sup>T</sup>. -ve is negative control. A 1Kb Hyperladder (Bioline) was run as a size marker in both the first and last wells.

## Dryocola

The results from the BOX PCR demonstrated that although some isolates were clonal regardless of sampling location, a degree of genetic diversity exists within both proposed novel species (Figure 38).

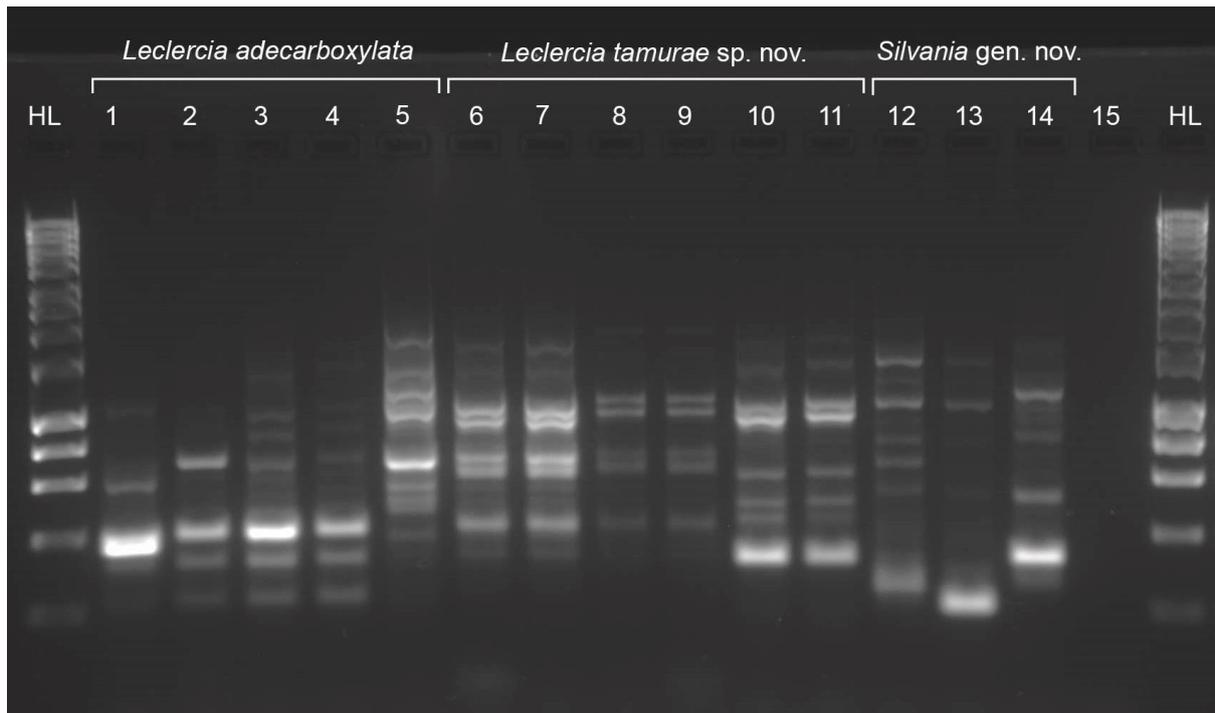


**Figure 38:** BOX PCR fingerprinting patterns generated from strains of the proposed novel genus *Dryocola*. (1) *Dryocola boscaaweniae* H18W14, (2) *Dryocola boscaaweniae* H19S8, (3) *Dryocola boscaaweniae* H14W9, (4) *Dryocola boscaaweniae* H6W4<sup>T</sup>, (5) *Dryocola boscaaweniae* H1W6, (6) *Dryocola boscaaweniae* H15E8, (7) *Dryocola boscaaweniae* H20N1, (8) *Dryocola boscaaweniae* H20N9, (9) *Dryocola clanedunensis* H16N7, (10) *Dryocola clanedunensis* H6S10, (11) *Dryocola clanedunensis* H9W11, (12) *Dryocola clanedunensis* H18E4, (13) *Dryocola clanedunensis* H11S18<sup>T</sup>, (14) negative control. A 1Kb Hyperladder (Bioline) was run as a size marker in both the first and last wells.

### 5.2.3. ERIC PCR:

#### *Leclercia and Silvania*

ERIC PCR was performed on all 12 strains. The results from the ERIC PCR (Figure 39) showed that clonal isolates were present, especially in isolates H6S3<sup>T</sup>, H6W8 and H6W6a from tree six, however genetic diversity was observed between and within all of the clusters identified in the MLSA phylogenetic tree.



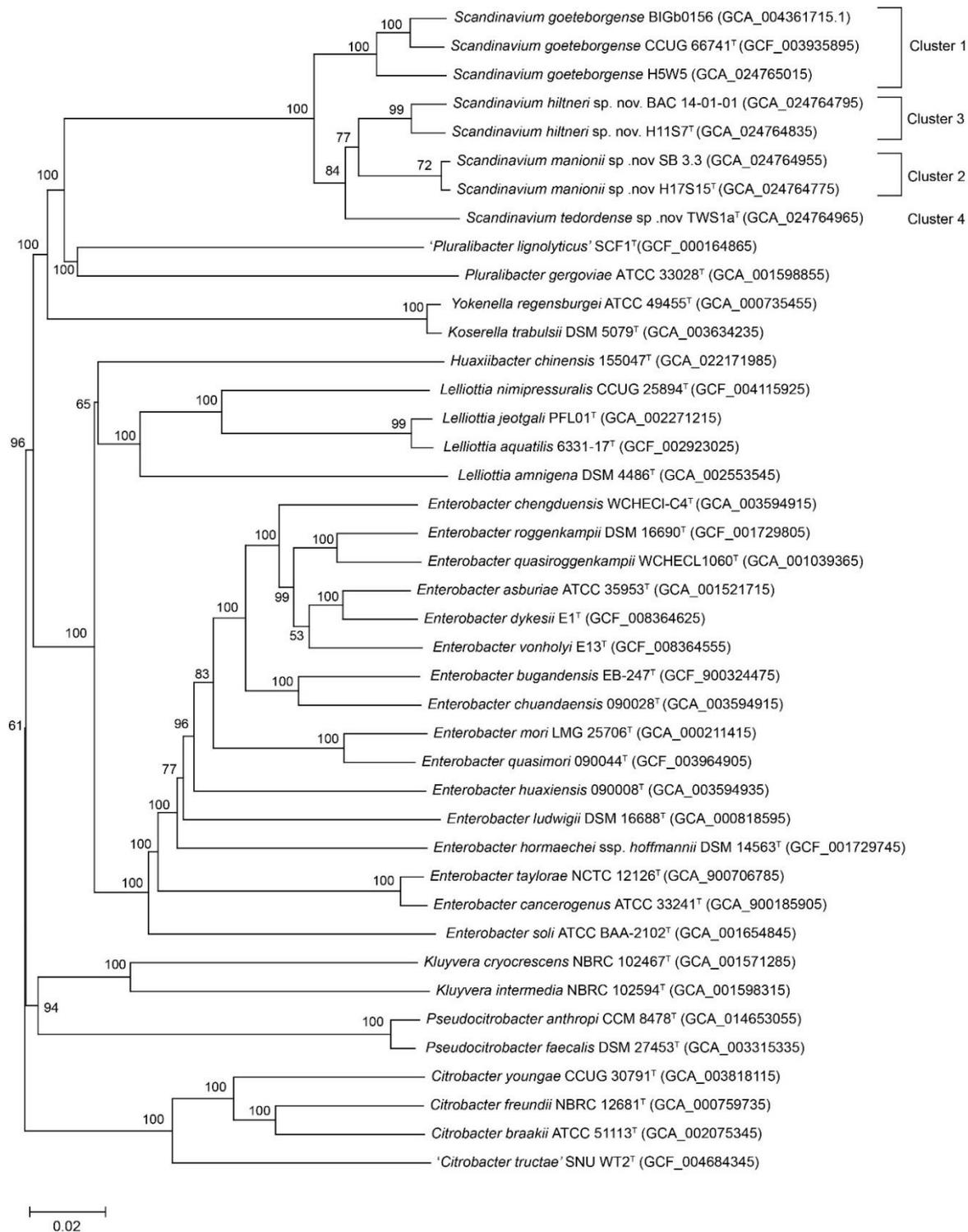
**Figure 39:** ERIC PCR fingerprinting patterns generated from strains of *Leclercia adecarboxylata*, *Leclercia tamurae* sp. nov. and the novel genus *Sylvania* gen. nov. (1) *Leclercia adecarboxylata* LMG 2803<sup>T</sup>, (2) *Leclercia adecarboxylata* LMG 2650, (3) *Leclercia adecarboxylata* H10E4, (4) *Leclercia adecarboxylata* H9E1a, (5) *Leclercia adecarboxylata* H10E8, (6) *Leclercia tamurae* H6S3<sup>T</sup>, (7) *Leclercia tamurae* H6W8, (8) *Leclercia tamurae* H6W6a, (9) *Leclercia tamurae* H20N5, (10) *Leclercia tamurae* H6W5, (11) *Leclercia tamurae* H6S9, (12) *Sylvania hatchlandensis* H19S6<sup>T</sup>, (13) *Sylvania hatchlandensis* H18E8, (14) *Sylvania confinis* H4N4<sup>T</sup>, (15) negative control. A 1Kb Hyperladder (Bioline) was run as a size marker in both the first and last wells.

#### 5.2.4. Genomic characterisation:

##### *Scandinavium*

Genomes sizes varied slightly by MLSA Cluster, with strains exhibiting an average size of 4.67 Mbp. The DNA G + C content reported for the genus remains consistent, ranging from 53.9 to 54.5 %. The genome features and assembly accessions are listed in Suppl. Table. S2. The resulting phylogenomic tree (Figure 40) generated from these sequences demonstrated, similarly to the MLSA maximum likelihood tree, that Clusters 2–4 constitute potentially novel species within the genus *Scandinavium*. Cluster 2 (H17S15<sup>T</sup> and SB 3.3), Cluster 3 (H11S7<sup>T</sup> and BAC 14–01-01) and Cluster 4 (TWS1a<sup>T</sup>) are distinct from each other, group with no known species and have between 80 – 100% bootstrap support for each branch.

The *isDDH* and *ANIb* values, presented in Table 21, confirm that Clusters 2–4 constitute three novel taxa. The strains within Clusters 2 and 3 exhibited *isDDH* and *ANIb* values exceeding the suggested cut-off values on 70 and 95 % (Goris *et al.*, 2007; Konstantinidis and Tiedje, 2005a), respectively, while the values between each Cluster were below these cut-off values. H5W5, a representative strain from Cluster 1, shared 69.1 and 70.2 % *isDDH* similarity with the type strain of *S. goeteborgense*, CCUG 66741<sup>T</sup>, and with BIGb0156, respectively. These values could be considered borderline, however as the *ANIb* similarity between H5W5 and CCUG 66741<sup>T</sup> and BIGb0156 is 95.9 %, there is support for its classification as *S. goeteborgense*. The Average Amino Identity (AAI) values were less informative with all strains included in the comparisons showing 94–100 % similarity with each other. However, despite the conserved nature of the proteins analysed, isolates assigned to the same species showed 98–100 % similarity while different species could be identified by AAI values ranging from 94 to 97 %, the top end of which only slightly exceeds the 96 % cut-off for species delimitation (Konstantinidis and Tiedje, 2005b)



**Figure 40:** Phylogenomic tree for the genus *Scandinavium* including the three proposed novel species, *Scandinavium hiltneri* sp. nov., *Scandinavium manionii* sp. nov. and *Scandinavium tedordense* sp. nov., and close phylogenetic neighbours. 100 replicate GBDP pseudo-bootstrap support percentages are shown (> 50 %).  $d_5$  GBDP distance formula was used to scale branch lengths, and the tree was rooted at the midpoint. Type strains denoted via <sup>T</sup>.

**Table 21:** *in silico* DNA - DNA Hybridisation d5 matrix (*isDDH* – top right) and Average Nucleotide Identity based on BLAST (ANiB – bottom left) percentage values for novel *Scandinavium* species and *S. goeteborgense*. Shaded boxes represent values that exceed the cut off point for species delimitation (>70 % *isDDH* and >95 % for ANI).

<i>isDDH</i> ANiB	1	2	3	4	5	6	7	8
1	100	69.1	84.1	52	51.5	52.2	52	48.6
2	95.9	100	70.2	53.3	53.3	53.7	53.5	50.4
3	97.71	95.9	100	52.1	51.5	52.3	52.1	48.6
4	92.5	92.6	92.4	100	84.5	63.5	62.8	59
5	92.6	92.7	92.6	97.75	100	62.7	62.3	58.5
6	92.8	92.9	92.7	94.8	95	100	96.7	56
7	92.6	92.7	92.5	94.5	94.8	99.4	100	55.3
8	92	92	91.9	92.3	94.1	93.7	93.4	100

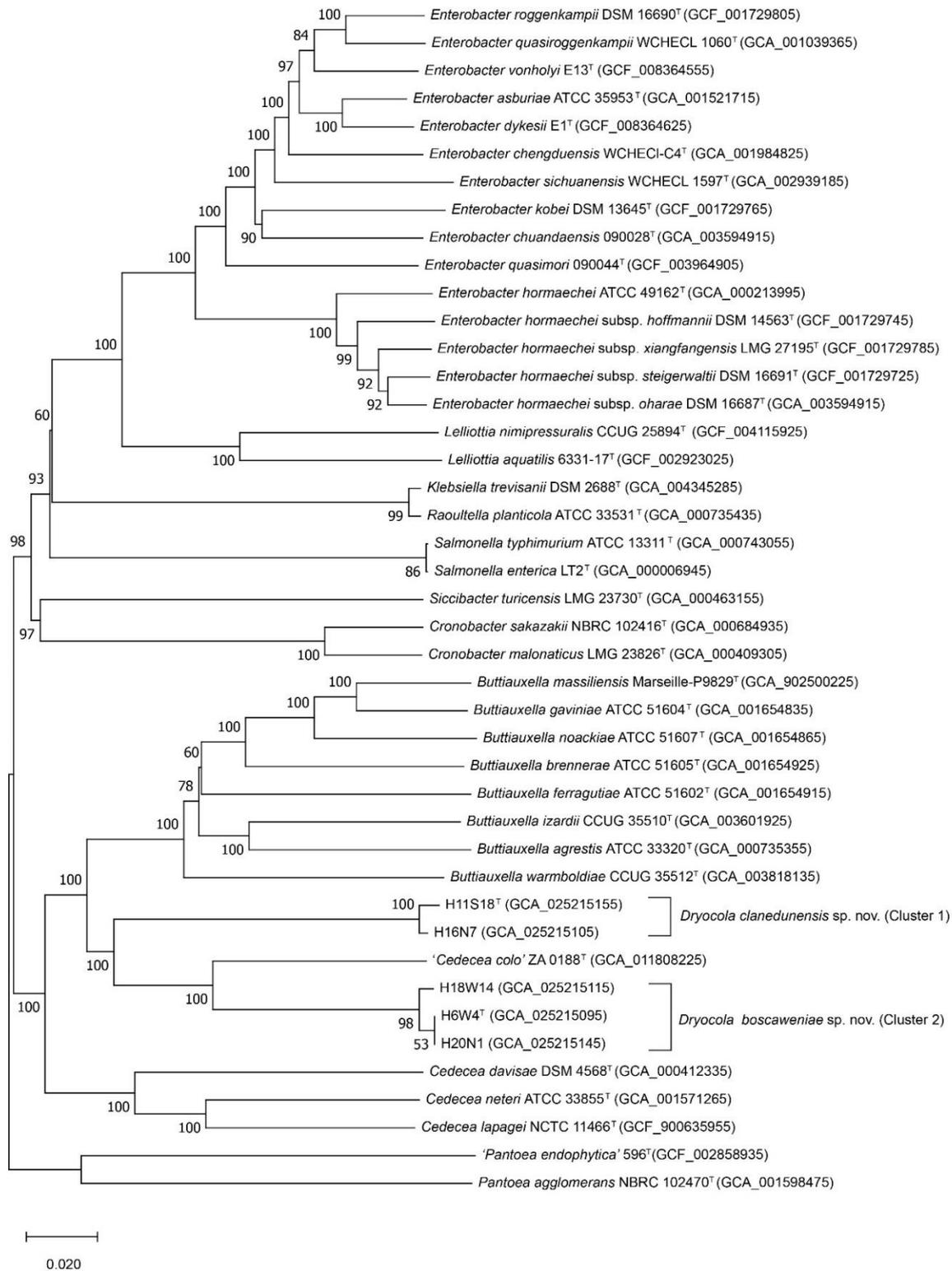
(1) *Scandinavium goeteborgense* CCUG 66741<sup>T</sup> (GCA\_003935895), (2) *Scandinavium goeteborgense* H5W5 (GCA\_024765015), (3) *Scandinavium goeteborgense* BIGb0156 (GCA\_004361715), (4) *Scandinavium hiltneri* H11S7<sup>T</sup> (GCA\_024764835), (5) *Scandinavium hiltneri* BAC 14-01-01 (GCA\_024764795), (6) *Scandinavium manionii* H17S15<sup>T</sup> (GCA\_024764775), (7) *Scandinavium manionii* SB 3.3 (GCA\_024764955), (8) *Scandinavium tedordense* TWS1a<sup>T</sup> (GCA\_024764965).

## *Dryocola*

Genome sizes of the two potential novel species varied, with strains of Cluster 1 ranging from 4.82 – 5.23 Mbp and strains of Cluster 2 being 4.4 – 4.45 Mbp. The DNA G + C content was 53.0 - 53.9 mol % across strains from both potential novel species. Genomes sequences were submitted to GenBank under the BioProject number PRJNA814476, and genome features and assembly accessions are listed in Suppl. Table. S3.

The resulting phylogenomic tree (Figure 41) from the TYGS analysis demonstrated both potential novel species clusters formed a robust clade with '*Cedecea colo*', with 100 % bootstrap support. The clade is positioned between the *Buttiauxella* and *Cedecea* genus clades also with 100 % support, reflecting the position observed in the MLSA phylogenetic tree (Figure 35) and further supporting the description of a novel genus, two novel species and the possible transfer of '*Cedecea colo*' to the novel genus.

The genome comparison values demonstrated that strains from Cluster 1 shared ANI values of 99.1 % and AAI values of 99 % to each other, and 81.7 – 82.4 % ANI and 82 – 86 % AAI to species of *Buttiauxella* and *Cedecea*. Cluster 2 strains demonstrated 99.1 - 100 % similarity through ANI values and 99.0 % AAI values with each other, and 81.9 – 82.8 % ANI and 83 – 87 % AAI to species of *Buttiauxella* and *Cedecea*. The intra-species values exceed the suggested 95 % similarity required to delimit species for ANI and 96 % for AAI (Konstantinidis and Tiedje, 2005b), confirming that both clusters constituted single taxa (Figure 41). The conclusions drawn from the ANI analysis were confirmed by *in silico* DNA-DNA hybridisation (*isDDH*) with both clusters exceeding the cut-off value of >70 % indicating a different species while showing high similarity to each other (Goris *et al.*, 2007). Both the ANI values and *isDDH* values for each cluster can be seen in Table 22.



**Figure 41:** Phylogenomic tree of the proposed genus *Dryocola* gen. nov., and its closest phylogenetic neighbours. GBDP pseudo-bootstrap support values from 100 replicates exceeding 50 % are shown at the nodes, with an average branch support of 94.4 %. Branch lengths are scaled from the  $d_5$  GBDP distance formula and the tree is rooted at the midpoint. <sup>T</sup> signifies the type strain and GenBank assembly numbers are shown in parentheses.

**Table 22:** In silico DNA-DNA Hybridisation (*isDDH* – top right) and Average Nucleotide Identities matrix (*fastANI* – bottom left) percentage values for *Dryocola boscaweniae* sp. nov. and *Dryocola clanedunensis* sp. nov. and all species of *Cedecea* and *Buttiauxella*, the closest phylogenetic neighbours. Shaded boxes represent values that exceeded the cut off point for species delimitation (>70 % *isDDH* and >95 % for ANI).

<i>isDDH</i> <i>fastANI</i>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
<b>1</b>	<b>100</b>	94	94	26.4	26.4	35.7	22.8	22.7	22.8	23.7	23.8	23.6	23.6	24.2	23.3	24.1
<b>2</b>	99.1	<b>100</b>	100	26.5	26.5	35.8	22.8	22.7	22.6	23.9	23.7	23.6	23.6	24.2	23.3	24.1
<b>3</b>	100	99.1	<b>100</b>	26.4	26.4	35.8	22.8	22.7	22.6	23.7	23.7	23.6	23.5	24.2	23.3	24.1
<b>4</b>	84.2	84.2	84.2	<b>100</b>	89.9	26.4	22.7	22.7	22.6	22.9	23	22.9	22.9	23.3	22.7	23.7
<b>5</b>	84.1	84.1	84.1	99.1	<b>100</b>	26.5	22.7	22.8	22.7	23.9	24.5	23.7	26.4	23.9	24.5	24.2
<b>6</b>	88.5	88.6	88.6	83.9	88.7	<b>100</b>	22.9	22.8	22.8	23.6	23.5	23.5	23.3	24.1	23.2	24.1
<b>7</b>	81.9	81.8	81.9	81.7	81.8	81.9	<b>100</b>	28.9	28.9	21.6	21.5	21.5	21.3	22	21.4	22.3
<b>8</b>	81.6	81.6	81.6	81.8	81.8	81.6	85.9	<b>100</b>	28.9	21.8	21.4	21.5	21.3	21.9	21.3	22.2
<b>9</b>	81.7	81.9	81.7	81.7	81.8	81.7	86.1	89.1	<b>100</b>	21.6	21.6	21.6	21.3	21.9	21.5	22.3
<b>10</b>	82.4	82.5	82.4	82.1	82.6	82.3	81	81.1	81.1	<b>100</b>	31.3	30.8	31	35.8	30.3	31.5

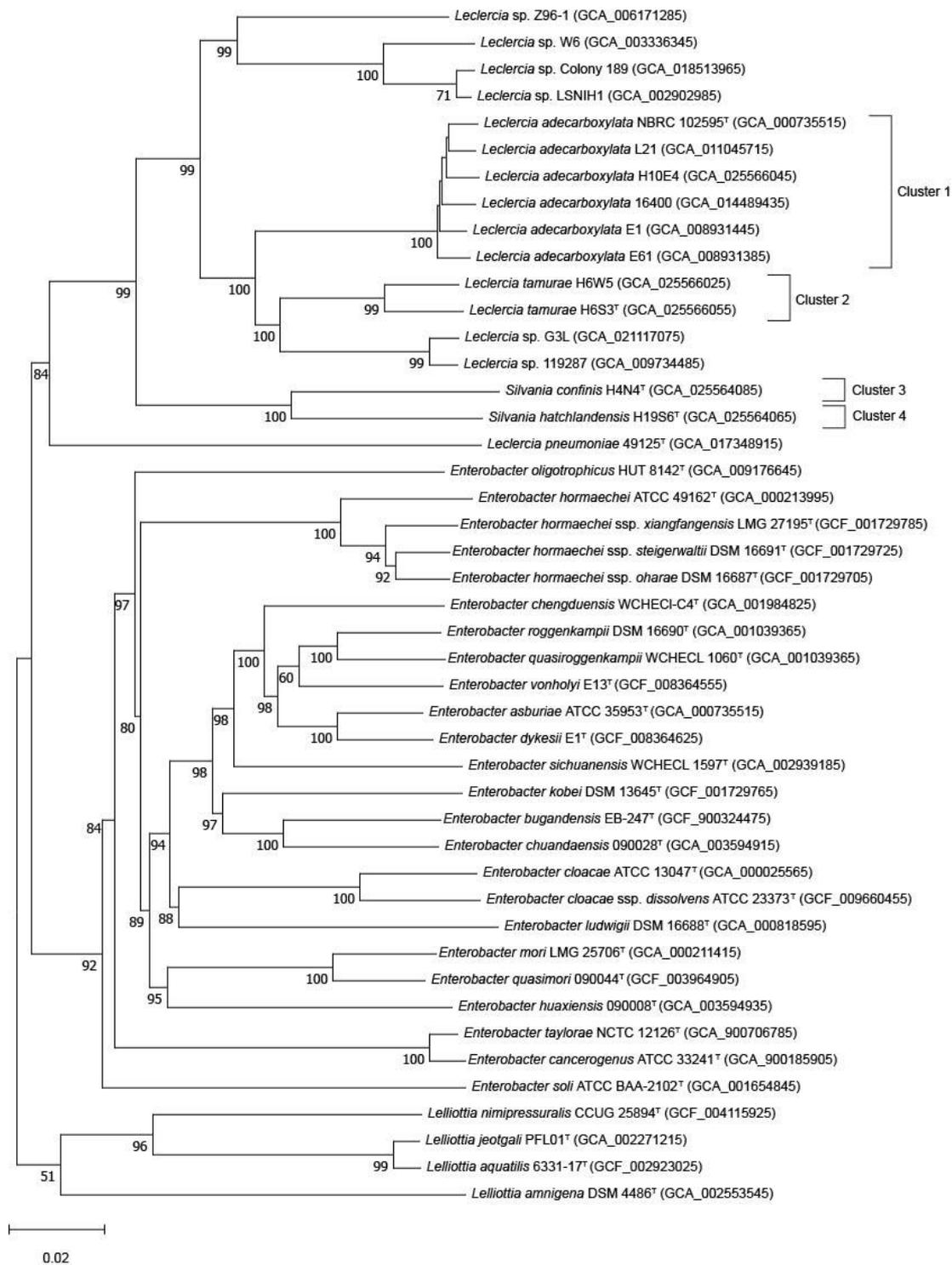
<b>11</b>	82.3	82.3	82.3	81.7	82.9	81.9	80	80.4	80.7	87.5	<b>100</b>	30.8	39.1	30.1	32.7	29.3
<b>12</b>	82.2	82.1	82.2	81.8	82.2	82	81	80.7	80.9	86.6	86.8	<b>100</b>	30.2	29.8	29.8	28.9
<b>13</b>	82.2	82.2	82.2	81.7	84.1	81.8	80.9	80.6	80.7	87.3	91	86.3	<b>100</b>	29.8	44.6	29.4
<b>14</b>	82.6	82.6	82.6	82.1	82.6	82.3	81.3	80.9	81.2	89.2	86.7	86.3	86.5	<b>100</b>	29.6	31.3
<b>15</b>	82.1	82.9	82.1	81.5	82.9	81.9	80.9	80.7	80.9	86.7	87.9	86	91.9	86.2	<b>100</b>	29.5
<b>16</b>	82.9	82.8	82.8	82.4	82.8	82.6	81.5	81.2	81.2	87	85.6	85.3	86.7	86.7	85.5	<b>100</b>

(1) *Dryocola boscaweniae* H6W4<sup>T</sup> (GCA\_025215095), (2) *Dryocola boscaweniae* H18W14 (GCA\_025215115), (3) *Dryocola boscaweniae* H20N1 (GCA\_025215145), (4) *Dryocola claudunensis* H11S18<sup>T</sup> (GCA\_025215155), (5) *Dryocola claudunensis* H16N7 (GCA\_025215105), (6) '*Cedecea colo*' ZA 0188<sup>T</sup> (GCA\_011808225), (7) *Cedecea davisae* DSM 4562<sup>T</sup> (GCA\_000412335), (8) *Cedecea lapagei* ATCC 33855<sup>T</sup> (GCA\_001571265), (9) *Cedecea neteri* NCTC 11466<sup>T</sup> (GCF\_900635955), (10) *Buttiauxella agrestis* ATCC 33320<sup>T</sup> (GCA\_000735355), (11) *Buttiauxella brennerae* ATCC 51605<sup>T</sup> (GCA\_001654925), (12) *Buttiauxella ferragutiae* ATCC 51602<sup>T</sup> (GCA\_001654915), (13) *Buttiauxella gaviniae* ATCC 51604<sup>T</sup> (GCA\_001654835), (14) *Buttiauxella izardii* CCUG 35510<sup>T</sup> (GCA\_003601925), (15) *Buttiauxella noackiae* ATCC 51607<sup>T</sup> (GCA\_001654865), (16) *Buttiauxella warmboldiae* CCUG 35512<sup>T</sup> (GCA\_003818135).

### *Leclercia and Silvania*

Whole genome sequencing was performed on five novel isolates from the four MLSA clusters (H10E4 – Cluster 1, H6S3<sup>T</sup> and H6W5 – Cluster 2, H4N4<sup>T</sup> – Cluster 3, and H19S6<sup>T</sup> – Cluster 4). The genomes showed little variation, with the size and G + C DNA content ranging from 4.71 – 4.87 Mbp and 55.6 – 56.4 mol %, respectively. The genomes were submitted to GenBank under the BioProject numbers PRNJA837588 and PRNJA837589, and the genome features and accession numbers are listed in Suppl. Table. S4. All sequenced genomes were found to be free of contamination following alignment and comparison of the 16S rRNA gene sequences obtained from both the whole genomes and Sanger sequencing.

The phylogenomic tree (Figure 42), based on whole genome comparisons, supported the phylogeny demonstrated in the MLSA tree, with H10E4 confirmed as belonging to *L. adecarboxylata* along with other strains identified as *Leclercia* sp. in the MLSA tree. H6S3<sup>T</sup> and H6W5 formed a well-supported cluster in the *Leclercia* clade, along with strains GL3 and 119287 in a separate cluster which could constitute another novel species as in the MLSA tree. Strains Z96-1, W6, Colony 189 and LSNIH1 appear further removed from the *Leclercia* clade, suggesting that they could constitute another novel genus, with three novel species. The two strains from Clusters 3 and 4 formed a clade with 100 % bootstrap support, clearly distant from the *Leclercia* clade and did not contain any validly published type strain or reference strain confirming these strains constitute a novel genus. Finally, *Leclercia pneumoniae* 49125<sup>T</sup> was furthest removed from the *Leclercia* clade on a separate lineage and did not cluster with any known type strain or reference strain.



**Figure 42:** Phylogenomic tree of the proposed genus *Sylvania* gen. nov., the novel species *Leclercia tamurae* sp. nov. and their closest phylogenetic neighbours. GBDP pseudo-bootstrap support values from 100 replicates (> 50 %) are shown at the nodes, with the average branch support of 94.4 %. Branch lengths are scaled from the  $d_5$  GBDP distance formula and the tree is rooted at the midpoint. <sup>T</sup> denotes type strain, and GenBank assembly numbers are shown in parentheses.

To complement the phylogenomic comparison, a DNA similarity matrix was created through ANI, AAI and *is*DDH comparisons. The ANI and *is*DDH values are presented in Table 23. H10E4 displayed *is*DDH values between 87.4 – 89.0 %, ANI values between 98.4 – 98.6 % and AAI values of 99 % to the type strain of *L. adecarboxylata* LMG 2803<sup>T</sup> and other strains identified as *L. adecarboxylata*, far exceeding the 70 %, 95 % and 96 % similarity values used to delimit species for *is*DDH, ANI and AAI (Goris *et al.*, 2007; Konstantinidis and Tiedje, 2005b). Likewise, strains from Cluster 2 (H6S3<sup>T</sup> and H6W5) demonstrated 89.9 % *is*DDH, 98.6 % ANI and 98 % AAI values to each other but < 70 % *is*DDH, < 95 % ANI and 95 % AAI values to *L. adecarboxylata*, confirming they belong to a single novel taxon. Finally, H4N4<sup>T</sup> (Cluster 3) and H19S6<sup>T</sup> (Cluster 4) were 45.6 % similar based on *is*DDH, and 92.1 % and 95 % similar based on ANI and AAI, respectively. Both strains demonstrated lower values of < 35 % *is*DDH, < 88 % ANI and 90 – 91 % AAI to strains of *Leclercia*, providing further support for their classification of a novel genus. Therefore, *Leclercia tamurae* sp. nov. for strains in Cluster 2, and *Silvania* gen. nov. with *Silvania hatchlandensis* sp. nov. and *Silvania confinis* sp. nov. for strains in Clusters 3 and 4 are proposed.

The *is*DDH, ANI and AAI values for the additional *Leclercia* strains support the phylogenies of the MLSA and phylogenomic trees. Strains GL3 and 119287 demonstrated similarity values indicating they belong to a novel species closely related to *L. adecarboxylata* and *L. tamurae* sp. nov. Strains Z96-1, W6, Colony 189 and LSNIH1 are less related to species of *Leclercia* based on DNA similarity values, suggesting these strains most likely belonging to another novel genus, although further work would be required to fully understand their taxonomic position. *L. pneumoniae* 49125<sup>T</sup> was least related to all strains of *Leclercia* species displaying the lowest *is*DDH, ANI and AAI values, indicating that it is not a true member of the genus and should be transferred to a novel genus.

**Table 23:** Genome comparison values for *in silico* DNA - DNA Hybridisation (*isDDH* – top right) and Average Nucleotide Identity (*fastANI* – bottom left) for *Leclercia* and *Silvania* species. Shaded boxes represent values that exceed the cut off point for species delimitation (>70 % *isDDH* and >95 % for ANI).

<i>isDDH</i> <i>fastANI</i>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
<b>1</b>	<b>100</b>	90.1	89.4	87.6	88.6	87.4	26.2	44.9	44.6	45.0	45.1	36.5	36.6	36.8	37.6	31.6	31.2
<b>2</b>	98.6	<b>100</b>	88.4	88.0	88.6	87.5	26.5	44.9	44.8	45.4	45.3	36.7	36.7	37.0	37.9	31.7	31.2
<b>3</b>	98.6	98.5	<b>100</b>	87.4	89.0	87.5	26.3	44.7	44.7	45.0	44.9	36.5	36.7	36.7	37.8	31.7	31.2
<b>4</b>	98.3	98.4	98.4	<b>100</b>	88.2	86.6	26.2	44.8	45.0	44.9	45.2	36.4	37.1	37.0	39.3	31.6	31.0
<b>5</b>	98.5	98.5	98.6	98.4	<b>100</b>	89.2	26.4	45.0	45.2	45.4	45.4	36.8	37.5	37.2	40.5	31.7	31.2
<b>6</b>	98.3	98.4	98.4	98.2	98.6	<b>100</b>	26.5	44.6	44.8	45.2	45.2	36.8	37.3	36.9	39.4	31.7	31.1
<b>7</b>	84.2	84.4	84.3	84.3	84.4	84.3	<b>100</b>	26.6	26.6	27.1	27.0	26.7	26.7	26.9	27.0	26.0	25.8
<b>8</b>	91.8	91.7	91.6	91.7	91.7	91.6	84.4	<b>100</b>	73.2	49.9	49.7	37.8	38.0	38.4	39.3	32.9	32.1
<b>9</b>	91.6	91.6	91.6	91.7	91.7	91.6	84.4	96.8	<b>100</b>	49.8	49.6	38.0	38.1	38.4	39.1	32.8	32.0

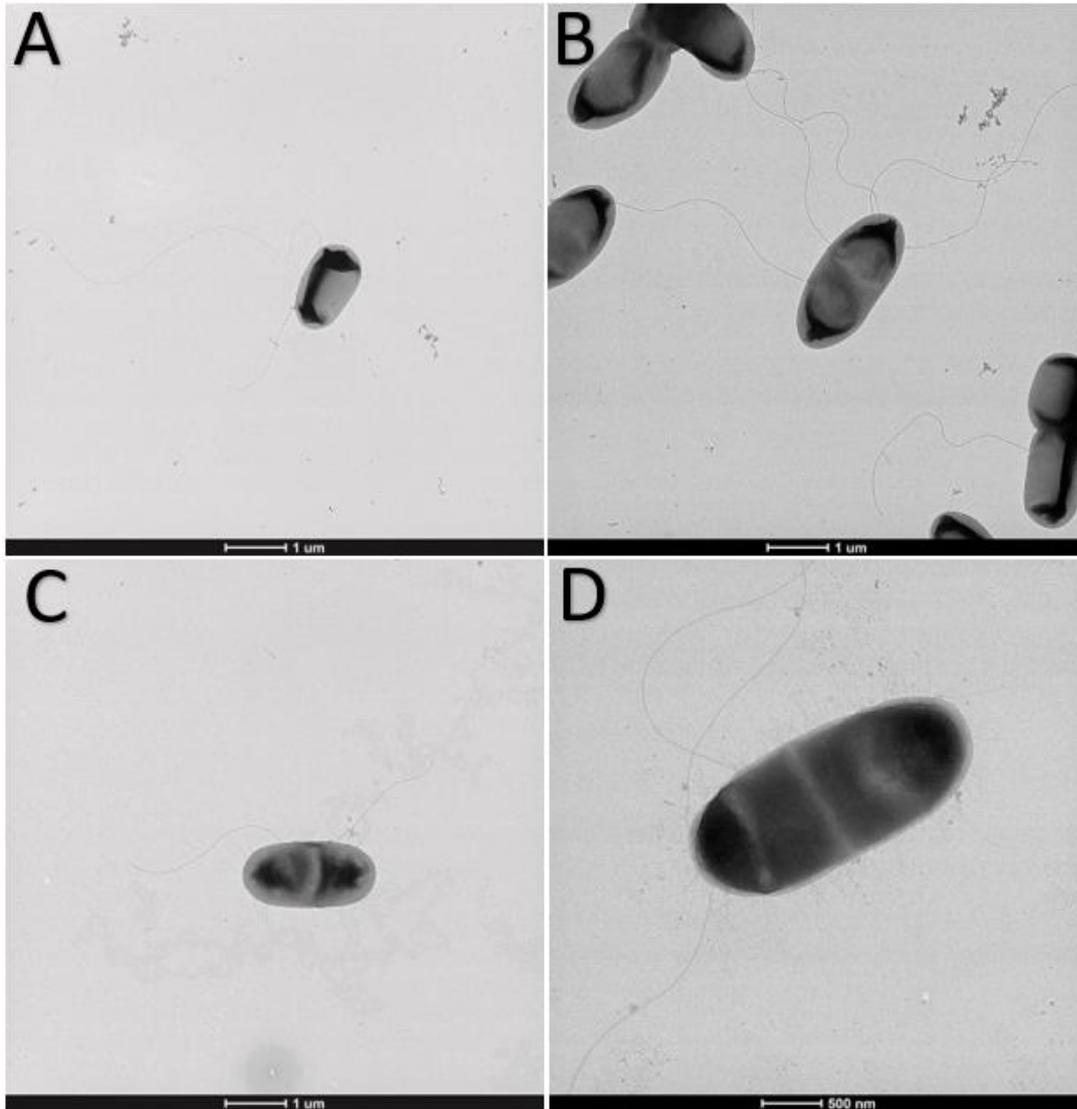
<b>10</b>	91.6	91.7	91.6	91.6	91.7	91.7	84.7	92.9	92.9	<b>100</b>	89.9	39.0	39.3	39.7	40.4	32.9	32.2
<b>11</b>	91.7	91.7	91.7	91.8	91.8	91.7	84.6	93.0	93.0	98.6	<b>100</b>	39.1	39.0	39.7	40.2	33.0	32.2
<b>12</b>	89.0	89.1	89.1	89.0	89.2	89.2	84.5	89.8	89.8	90.0	90.0	<b>100</b>	94.6	69.6	43.1	31.3	30.6
<b>13</b>	89.0	89.1	89.1	89.2	89.4	89.3	84.5	89.8	89.8	90.1	90.0	99.2	<b>100</b>	69.3	43.5	31.2	30.5
<b>14</b>	89.2	89.2	89.1	89.4	89.5	89.4	84.5	89.9	89.9	90.2	90.4	96.4	96.4	<b>100</b>	43.1	31.4	30.8
<b>15</b>	89.4	89.6	89.6	90.0	90.4	90.0	84.7	90.3	90.1	90.5	90.5	91.3	91.3	91.2	<b>100</b>	32.1	31.4
<b>16</b>	87.2	87.1	87.1	87.2	87.1	87.1	84.0	87.9	87.8	87.6	87.7	86.9	86.9	87.0	87.5	<b>100</b>	46.5
<b>17</b>	86.7	86.8	86.8	86.8	86.8	86.6	83.8	87.4	87.3	87.3	87.3	86.6	86.5	86.6	87.0	92.1	<b>100</b>

Strains which exceed the cut of values used for species delimitation are shown in shaded boxes (>70 % *is*DDH or >95 % ANI). 1 = *Leclercia adecarboxylata* NBRC 102595<sup>T</sup> (GCA\_001515505), 2 = *Leclercia adecarboxylata* L21 (GCA\_011045715), 3 = *Leclercia adecarboxylata* H10E4 (GCA\_025566045), 4 = *Leclercia adecarboxylata* 16400 (GCA\_014489435), 5 = *Leclercia adecarboxylata* E1 (GCA\_008931445), 6 = *Leclercia adecarboxylata* E61 (GCA\_008931385), 7 = *Leclercia pneumoniae* 49125<sup>T</sup> (GCA\_018987305), 8 = *Leclercia tamurae* H6S3<sup>T</sup> (GCA\_025566055), 9 = *Leclercia tamurae* H6W5 (GCA\_025566025), 10 = *Leclercia* sp. G3L (GCA\_021117075), 11 = *Leclercia* sp. 119287 (GCA\_009734485), 12 = *Leclercia* Colony 189 (GCA\_018513965), 13 = *Leclercia* sp. LSNIH1 (GCA\_002902985), 14 = *Leclercia* sp. W6 (GCA\_003336345), 15 = *Leclercia* sp. Z96-1 (GCA\_006171285), 16 = *Silvania hatchlandensis* H19S6<sup>T</sup> (GCA\_025564065), 17 = *Silvania confinis* H4N4<sup>T</sup> (GCA\_025564085).

5.2.5. Cell imaging:

*Scandinavium*

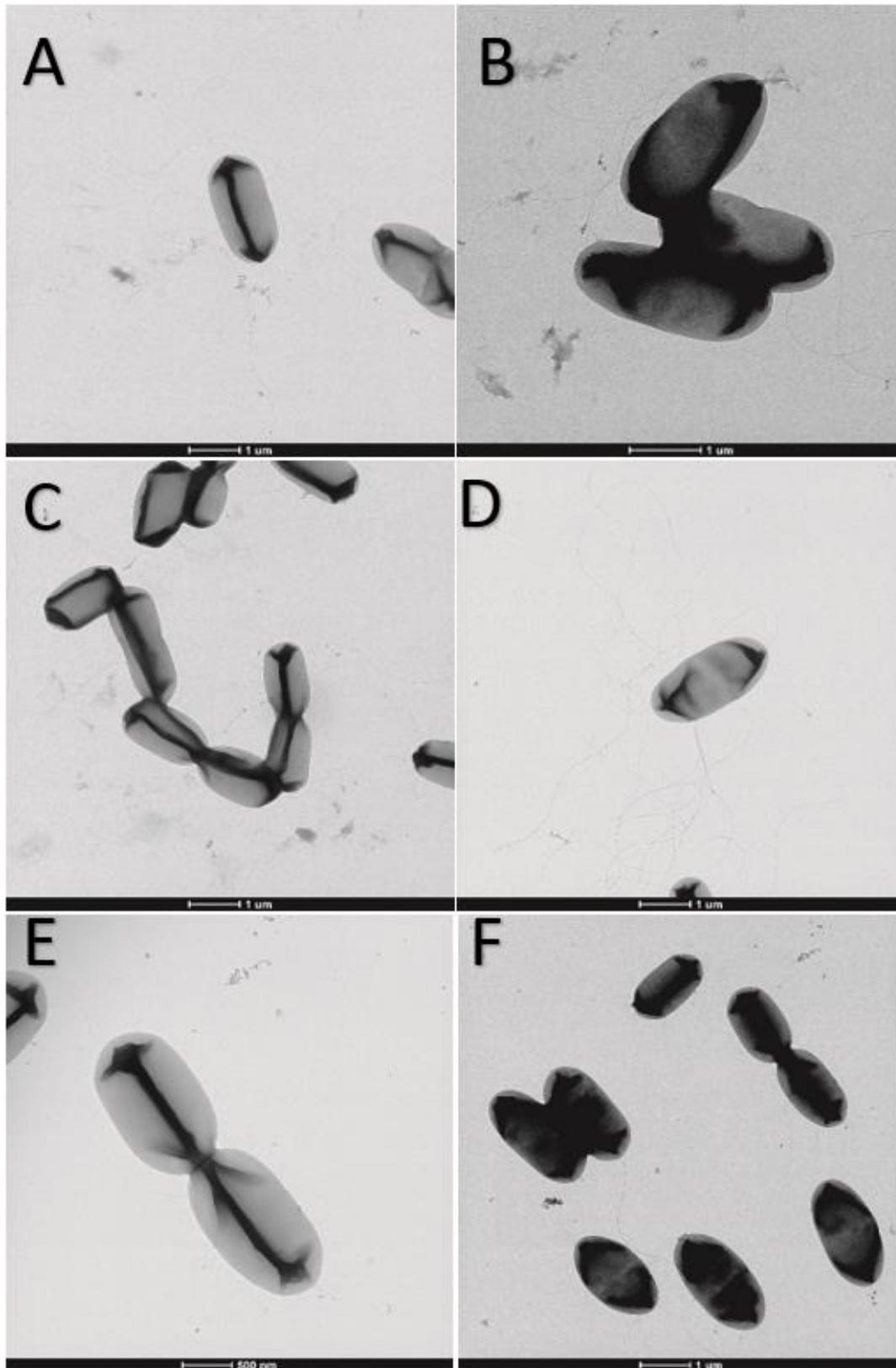
All strains were observed to be straight rods averaging  $1.14 \times 2.27 \mu\text{m}$  with peritrichous flagella and fimbriae. TEM images showing this are presented in Figure 43. These features are consistent with those seen for *Scandinavium* as originally described by Marathe et al. (2019).



**Figure 43:** Transmission electron microscopy of A) *Scandinavium goeteborgense* CCUG 66741<sup>T</sup>, B) *Scandinavium manionii* sp. nov. H17S15<sup>T</sup>, C) *Scandinavium hiltneri* sp. nov. H11S7<sup>T</sup> and D) *Scandinavium tedordense* sp. nov. TWS1a<sup>T</sup> displaying their peritrichous flagella arrangement. Scale bar, 1  $\mu\text{m}$  and 500  $\mu\text{m}$ .

## *Dryocola*

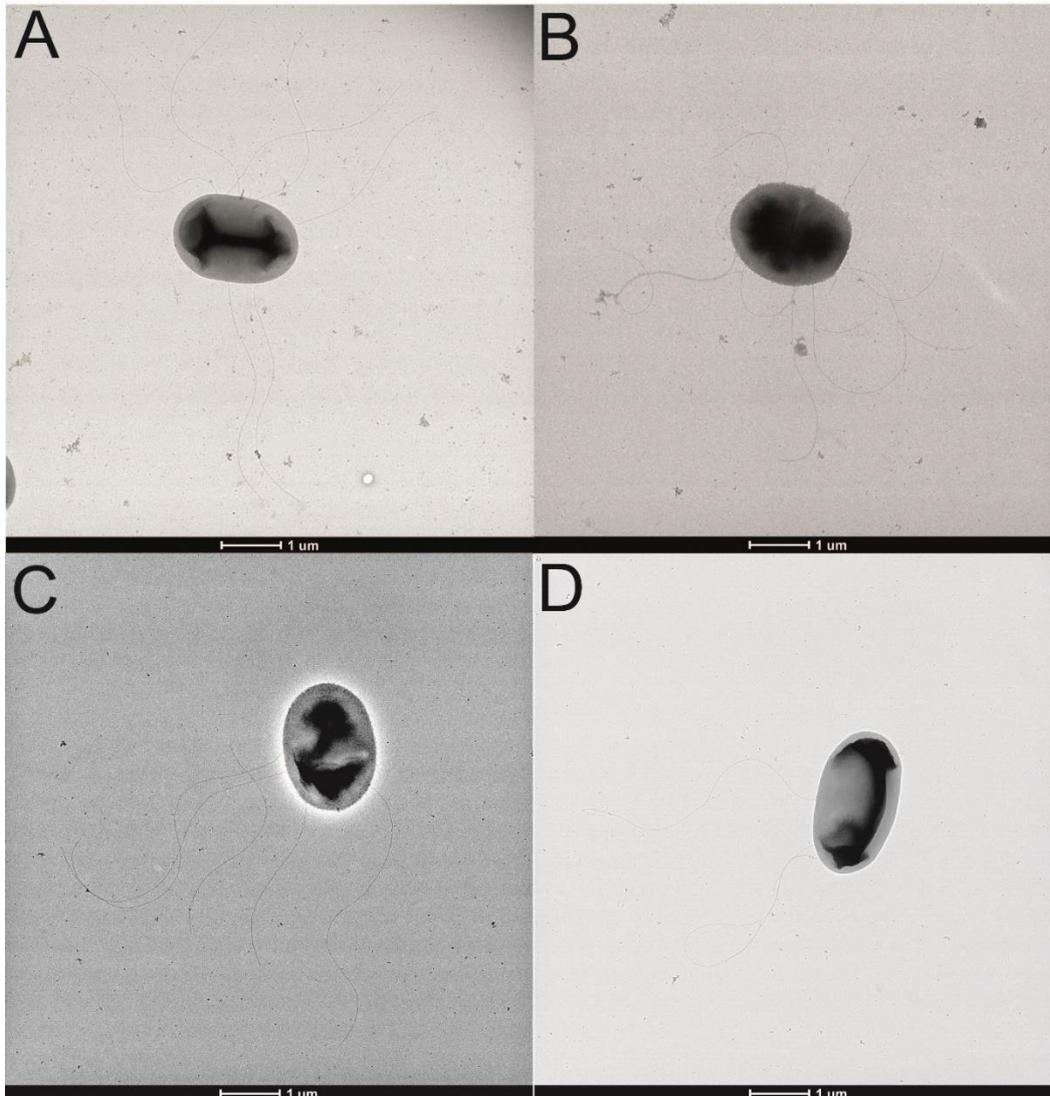
All strains were straight rods with an average size of 1.17 x 2.28  $\mu\text{m}$ . Cells appeared as single units, in groups of two or three and in chains. All strains were motile by peritrichous flagella (Figure 44 a-f). This is in keeping with the genome annotation of the sequenced strains, which revealed the presence of 55 genes that control the synthesis and motility of flagella in prokaryotes. Both new species also show the presence of fimbriae.



**Figure 44:** Transmission electron microscopy of *Dryocola boscaweniae* sp. nov. H6W4<sup>T</sup> (A, B, C) and *Dryocola clanedunesis* sp. nov. (D, E, F) displaying the peritrichous flagella, fimbriae and cell arrangement. Scale bar, 1 μm.

*Leclercia* and *Silvania*

All strains identified as *Leclercia* were straight rods averaging 1.38 x 2.26  $\mu\text{m}$ , while *Silvania* gen. nov. strains were short straight rods averaging 1.31 x 1.81  $\mu\text{m}$ . Cells are motile with peritrichous flagella and appear singly or in pairs (Figure 45).



**Figure 45:** Transmission electron microscopy images of A) *Leclercia tamurae* H6S3<sup>T</sup> B) *Leclercia tamurae* H6W5 C) *Silvania hatchlandensis* H19S6<sup>T</sup> D) *Silvania confinis* H4N4<sup>T</sup> displaying their peritrichous flagella arrangement. Scale bar, 1  $\mu\text{m}$ .

#### 5.2.6. Cell physiology:

##### *Scandinavium*

Strains appeared as moist circular colonies with a smooth, white centre and clear rim on CBA, with colonies ranging from 1–3 mm in size. Growth was observed from 4–37°C for all strains, with no growth observed at 41°C. Growth was observed in TSB at a pH range of 6–8, while salt was tolerated in concentrations from 1–8 % in NaCl-supplemented saline-free broth. All strains in this study formed agglomerated masses of growth when the salt or pH range was exceeded. All strains are oxidase negative, catalase positive and facultatively anaerobic.

##### *Dryocola*

The appearance of colonies on TSA after 48 h incubation at 28 °C were cream coloured, with a darker convex centre, uneven margins and 1-2 mm in diameter. Strains from Cluster 1 grew at 4, 10, 25, 28, 30 and 37 °C, but not 41 °C, while strains from Cluster 2 grew at 4, 10, 25, 28, 30, 37 and 41 °C. All strains grew across a pH range of 6 – 8, although strains from Cluster 2 were also capable of weak growth at pH 9. Growth for all strains was also observed in TSB supplemented with up to 6 % NaCl, with some strains from both clusters exhibited weak growth at 7 and 8 %.

##### *Leclercia and Silvania*

On TSA all strains tested appeared as circular, cream-coloured, convex colonies between 2 - 4 mm in diameter with entire, slightly undulate margins. All strains were observed changing from cream to yellow pigmented which is a known feature associated with *Leclercia*, although the time and conditions required for the pigment to form were not consistent (De Baere *et al.*, 2001). *Leclercia* species grew from 10 – 41 °C, while species of *Silvania* gen. nov. grew at 4 - 37 °C but not at 41 °C. The pH range at which growth was observed showed no difference between strains from both genera, with consistent growth seen from pH 6 – 9. All strains from both genera grew in a supplemented salt range of 1 – 7 % w/v, with the exception of *L. adecarboxylata* LMG 2803<sup>T</sup>, *L. tamurae* sp. nov. H6W6a and H6W8, and *S. confinis* sp. nov. H4N4<sup>T</sup> which could not grow at 7 %. All strains were recorded as negative for oxidase and positive for catalase production, which are key descriptive factors of the family *Enterobacteriaceae*.

#### 5.2.7. Antibiotic resistance:

##### *Scandinavium*

Antibiotic resistance was recorded for penicillin V, penicillin G, ampicillin, and cefotaxime while susceptibility to ciprofloxacin and tetracycline was recorded for all strains excluding H17S15T, which showed low level susceptibility to ampicillin.

##### *Dryocola*

Antibiotic resistance to penicillin G and V was recorded, with susceptibility to tetracycline, ampicillin, chloramphenicol and colistin sulphate observed for all strains of the novel species as well as the type strains of *Cedecea* species.

##### *Leclercia and Sylvania*

Antibiotic resistance for all strains from both genera was recorded for penicillin V and G, while susceptibility was recorded for tetracycline, ampicillin, chloramphenicol, colistin sulphate, streptomycin, cefotaxime, ciprofloxacin, cefepime, gentamycin and kanamycin.

#### 5.2.8. Phenotypic characterisation:

##### *Scandinavium*

Differentiation of *Scandinavium* species based on phenotypic properties is quite limited with only a few discriminating features between strains isolated in this study. The most important phenotypic features for differentiation of *Scandinavium* species can be seen in Table 24. The full results for reactions to each biochemical test can be found in the protologues in the supplementary material. The novel strains belonging to *S. goeteborgense* were phenotypically indistinguishable from CCUG 66741<sup>T</sup> apart from their ability to produce acid from potassium 2-ketogluconate. *S. hiltneri* sp. nov. is phenotypically unique in its ability to utilise citrate and produce acid from sorbitol and D-raffinose. *S. manionii* sp. nov. can be phenotypically differentiated from other *Scandinavium* species by the lack of lysine decarboxylase and inability to utilise  $\beta$ -methyl-D-glucoside as a carbon source. *S. tedordense* sp. nov. is the only species capable of acidification of both D- and L-fucose. All novel species can be distinguished from the type strain of *S. goeteborgense* CCUG 66741<sup>T</sup> by their inability to utilise  $\gamma$ -amino-butyric acid,  $\alpha$ -hydroxy-butyric acid and  $\beta$ -hydroxy-D,L-butyric acid as carbon sources.

**Table 24:** Key phenotypic characteristics that allow for the differentiation of all known members of the genus *Scandinavium*. (1) *Scandinavium goeteborgense* (n = 2), (2)

*Scandinavium hiltneri* (n = 2), (3) *Scandinavium manionii* (n = 3), (4) *Scandinavium tedordense* (n = 1).

Reaction	1	2	3	4
lysine decarboxylase	+	+	-	+
citrate utilization	-	+	-	-
sorbitol	-	+	-	-
<b>Fermentation of:</b>				
D-adonitol	-	+	+	+
dulcitol	-	V <sup>a</sup>	V <sup>a</sup>	-
D-raffinose	-	+	-	+
D-fucose	-	-	-	+
L-fucose	-	V <sup>b</sup>	-	+
D-arabitol	-	+	+	+
<b>Utilisation of:</b>				
β-methyl-D-glucoside	+	+	-	+
fusidic acid	+	+	+	-
myo-inositol	-	+	-	-
D-aspartic acid	+	+	-	+
D-serine	+	+	-	+
minocycline	-	+	-	-
L-pyroglutamic acid	+	+	-	-
D-saccharic acid	+	-	V <sup>b</sup>	+
p-hydroxy-phenylacetic acid	-	+	-	-
D-lactic acid methyl ester	+	+	-	+
tween 40	+	+	-	+
γ-amino-butyric acid	+	-	-	-
α-hydroxy-butyric acid	+	-	-	-
β-hydroxy-D,L-butyric acid	+	-	-	-
α-keto-butyric acid	-	-	-	-
acetoacetic acid	+	+	-	-

+, 90 – 100 % strains +; -, 91 – 100 % strains -; V, variable between species; <sup>a</sup>, positive for type strain; <sup>b</sup>, negative for type strain.

### *Dryocola*

Both species from the proposed novel genus were clearly differentiated from each other based on the phenotypic data as shown in Table 25 (by their fermentation of D-adonitol, D-lactose, D-melibiose, D-raffinose, D-lyxose, D-arabitol, potassium 2-ketogluconate, potassium 5-ketogluconate and palatinose; and utilisation of D-melibiose, D-arabitol, L-arginine and-  $\alpha$ -hydroxy-butyric acid), and from *Cedecea* species (differing reactions to fermentation of L-arabinose, L-rhamnose, methyl- $\alpha$ D-glucopyranoside phenol red and  $\alpha$ -glucosidase and citrate utilisation) as shown in Table 26. Differentiation from *Buttiauxella* based on the information from other studies was less clear, although no single feature is known to be able to differentiate *Buttiauxella* from other members of *Enterobacteriaceae* (Kämpfer, 2015). The most useful phenotypic characteristics for differentiation at the species and genus level are listed in Table 25 and Table 26. All strains tested were observed as both oxidase negative and strongly catalase positive.

**Table 25:** Key phenotypic characteristics that allow for the differentiation between *Dryocola boscaweniae* sp. nov. and *Dryocola clanedunensis* sp. nov. from each other, and existing species of *Cedecea* and *Buttiauxella*. (1) *Dryocola boscaweniae* (n = 3), (2) *Dryocola clanedunensis* (n = 2), (3) *Cedecea davisae* LMG 7682<sup>T</sup>, (4) *Cedecea lapagei* LMG 7863<sup>T</sup>, (5) *Cedecea neteri* LMG 7864<sup>T</sup>, (6) *Buttiauxella agrestis*, (7) *Buttiauxella brennerae*, (8) *Buttiauxella ferragutiae*, (9) *Buttiauxella gavinae*, (10) *Buttiauxella izardii*, (11) *Buttiauxella noackiae*, (12) *Buttiauxella warmboldiae*. Data for *Buttiauxella* (6 -12)\* taken from [22].

Reaction	1	2	3	4	5	6	7	8	9	10	11	12
arginine dihydrolase	+	+	-	+	+	-	-	-	-	-	v	-
ornithine decarboxylase	-	v	+	-	-	+	v	+	-	+	-	-
<b>Fermentation of:</b>												
glycerol	+	v	-	+	-	v	-	-	-	-	-	-
D-arabinose	-	-	-	-	-	v	-	-	v	+	v	+
L-arabinose	+	+	-	-	-	+	+	+	+	+	+	+
D-adonitol	+	-	-	-	-	-	v	-	v	-	-	-
L-rhamnose	+	+	-	-	-	+	+	+	+	+	+	+
methyl- $\alpha$ -D-glucopyranoside	+	+	-	-	-	-	v	-	-	-	-	-

D-lactose	+	-	+	+	+	+	+	-	-	+	-	-
D-melibiose	+	-	-	+	-	v	+	+	-	v	-	-
D-raffinose	+	-	-	+	-	v	+	-	-	-	-	-
D-lyxose	+	-	-	-	-	ND						
L-fucose	-	v	-	-	-	+	-	-	v	+	-	+
D-arabitol	+	-	+	+	+	-	v	-	v	-	-	-
potassium gluconate	+	+	+	-	-	ND						
potassium 2-ketogluconate	+	-	-	-	-	ND						
potassium 5-ketogluconate	+	-	-	-	-	-	+	-	+	v	+	-
phenol red	+	+	-	-	-	ND						
palatinose	+	-	+	+	-	+	+	+	+	V	+	-
$\alpha$ -glucosidase	+	+	-	-	-	ND						
$\alpha$ -galactosidase	+	-	-	-	-	ND						

<i>N</i> -acetyl-D-galactosamine	-	-	ND	ND	ND	+	v	+	+	+	+	+
<b>Utilisation of:</b>												
citrate	-	-	+	+	+	v	v	-	v	+	+	-
D-melibiose	+	-	ND									
D-arabitol	+	-	ND									
L-arginine	-	+	ND	ND	ND	-	-	-	-	-	-	-
L-pyroglutamic acid	+	-	ND									
$\alpha$ -hydroxy-butyric acid	+	-	ND									

+, 90 – 100 % strains +; -, 91 – 100 % strains -; v, variable; ND, not determined

\* Data for the genus *Buttiauxella* were generated using standard biochemical tests, not API 50 CH

**Table 26:** Key phenotypic characteristics that allow for the differentiation between (1) *Dryocola* gen. nov., (2) *Cedecea* and (3) *Buttiauxella*.

Reaction	1	2	3
Arginine dihydrolase	+	(v) <sup>b</sup>	(v) <sup>b</sup>
Ornithine decarboxylase	(v) <sup>b</sup>	v <sup>a</sup>	v <sup>a</sup>
<b>Fermentation of:</b>			
glycerol	v <sup>a</sup>	v <sup>b</sup>	(v) <sup>a</sup>
D-arabinose	-	-	(v) <sup>a</sup>
L-arabinose	+	-	+
D-adonitol	v <sup>a</sup>	-	(v) <sup>b</sup>
L-rhamnose	+	-	+
methyl- $\alpha$ D-glucopyranoside	+	-	(v) <sup>b</sup>
D-lactose	v <sup>a</sup>	+	v <sup>a</sup>
D-melibiose	v <sup>a</sup>	v <sup>b</sup>	(v) <sup>b</sup>
D-raffinose	(v) <sup>a</sup>	v <sup>b</sup>	v <sup>a</sup>
D-lyxose	v <sup>a</sup>	-	ND
L-fucose	-	-	(v) <sup>a</sup>
D-arabitol	v <sup>a</sup>	+	(v) <sup>b</sup>
potassium gluconate	+	v <sup>a</sup>	ND
potassium 2-ketogluconate	v <sup>a</sup>	-	ND
potassium 5-ketogluconate	v <sup>a</sup>	-	v <sup>b</sup>
phenol red	+	-	ND
palatinose	v <sup>a</sup>	(v) <sup>a</sup>	v <sup>a</sup>

α-glucosidase	+	-	ND
α-galactosidase	v <sup>a</sup>	-	ND
N-acetyl-D-galactosamine	-	ND	v <sup>a</sup>
<b>Utilisation of:</b>			
citrate	-	+	(v) <sup>a</sup>
L-arginine	v <sup>b</sup>	ND	-
α-hydroxy-butyric acid	v <sup>a</sup>	ND	ND

+, 90 – 100 % strains +; -, 91 – 100 % strains -; V, variable between species; (V), variable within species; <sup>a</sup>, positive for type strain; <sup>b</sup>, negative for type strain; ND, not determined

#### *Leclercia and Sylvania*

The new species and genus described all present phenotypically unique traits when tested with commercial kits, which can be used for their differentiation from each another and their closest relatives. *Leclercia* and *Sylvania* gen. nov. can be distinguished based on a number of traits including fermentation of D-arabinose and utilisation of p-hydroxy-phenylacetic acid and fusidic acid. *L. tamurae* sp. nov. can be differentiated from *L. adecarboxylata* based on the positive reaction to sorbitol and the inability to utilise D-adonitol or D-arabitol among other traits, while *Sylvania* gen. nov. species can be discriminated by reactions to indole production, rhamnose and sucrose fermentation and pectin utilisation. Table 27 and Table 28 show the most useful phenotypic characteristics used for the differentiation between species of *Leclercia* and *Sylvania* gen. nov., respectively and Table 29 shows those for the differentiation between the two genera. The full results for reactions to each biochemical test can be found in the protologues in the supplementary material.

**Table 27:** Key phenotypic characteristics for differentiation of (1) *Leclercia adecarboxylata* (n = 4), (2) *Leclercia pneumoniae* 49125<sup>T</sup> and (3) *Leclercia tamurae* sp. nov. (n = 5).

Reaction	1	2	3

sorbitol	-	-	+
sucrose	V <sup>a</sup>	-	-
<b>Acidification of:</b>			
D-adonitol	+	ND	-
dulcitol	V <sup>a</sup>	ND	+
methyl- $\alpha$ -D-glucopyranoside	-	ND	V <sup>b</sup>
D-trehalose	+	+	V <sup>a</sup>
D-raffinose	V <sup>a</sup>	ND	-
D-lyxose	V <sup>b</sup>	ND	V <sup>a</sup>
D-arabitol	+	-	-
potassium 2-ketogluconate	+	ND	-
potassium 5-ketogluconate	-	-	V <sup>b</sup>
palatinose	V <sup>a</sup>	-	-
malonate	V <sup>b</sup>	+	+
<i>N</i> -acetyl- $\beta$ -glucosaminidase	-	-	V <sup>a</sup>
<b>Utilisation of:</b>			
sucrose	-	-	V <sup>a</sup>
stachyose	V <sup>a</sup>		-

D-salicin	+	ND	V <sup>a</sup>
3-methyl glucose	-	ND	V <sup>a</sup>
D-aspartic acid	+	ND	-
pectin	V <sup>a</sup>	ND	-
citric acid	-	ND	V <sup>a</sup>
$\alpha$ -keto-glutaric acid	-	ND	-
D-malic acid	V <sup>a</sup>	ND	V <sup>b</sup>
potassium tellurite	-	ND	V <sup>b</sup>
tween 40	+	ND	V <sup>a</sup>
$\alpha$ -hydroxy-butyrlic acid	V <sup>a</sup>	ND	V <sup>a</sup>
$\beta$ -hydroxy-D,L-butyrlic acid	-	ND	V <sup>b</sup>
formic acid	+	ND	V <sup>b</sup>
<b>Resistant to:</b>			
D-serine	-	ND	V <sup>a</sup>
nalidixic acid	V <sup>b</sup>	ND	+
troleandomycin	-	ND	V <sup>a</sup>

+, positive reaction; -, negative reaction; V, variable within species; <sup>a</sup>, positive for type strain; <sup>b</sup>, negative for type strain.

**Table 28:** Key phenotypic characteristics for differentiation of (1) *Silvania hatchlandensis* sp. nov. (n = 2) and (2) *Silvania confinis* sp. nov. H4N4<sup>T</sup>

Reaction	1	2
indole production	+	-
rhamnose	+	-
sucrose	+	-
<b>Acidification of:</b>		
methyl- $\alpha$ -D-mannopyranoside	V <sup>b</sup>	-
methyl- $\alpha$ -D-glucopyranoside	V <sup>b</sup>	-
D-lyxose	-	+
D-tagatose	-	+
phenol red	+	-
$\beta$ -glucuronidase	+	-
malonate	V <sup>b</sup>	-
<b>Utilisation of:</b>		
stachyose	+	-
N-acetyl-D-galactosamine	+	-
L-pyroglutamic acid	+	-
pectin	+	-

quinic acid	+	-
α-keto-glutaric acid	+	-
D-malic acid	+	-
<b>Resistant to:</b>		
D-serine	-	+
troleandomycin	-	+
nalidixic acid	-	+
potassium tellurite	-	+

+, positive reaction; -, negative reaction; V, variable within species; <sup>a</sup>, positive for type strain; <sup>b</sup>, negative for type strain.

**Table 29:** Key phenotypic characteristics for differentiation between (1) *Leclercia*\* (n = 9) and (2) *Silvania* gen. nov. (n = 3)

Reaction	1*	2
indole production	+	V <sup>a</sup>
sorbitol	V <sup>a</sup>	+
rhamnose	+	V <sup>a</sup>
sucrose	V <sup>b</sup>	V <sup>a</sup>
<b>Acidification of:</b>		
D-arabinose	-	+
D-adonitol	V <sup>b</sup>	-
methyl-α-D-mannopyranoside	-	V <sup>b</sup>

methyl- $\alpha$ -D-glucopyranoside	V <sup>b</sup>	V <sup>b</sup>
D-raffinose	V <sup>b</sup>	+
D-lyxose	V <sup>a</sup>	V <sup>b</sup>
D-tagatose	-	V <sup>b</sup>
D-arabitol	V <sup>b</sup>	-
potassium 2-ketogluconate	V <sup>b</sup>	-
potassium 5-ketogluconate	V <sup>b</sup>	+
phenol red	+	V <sup>a</sup>
palatinose	V <sup>b</sup>	-
$\beta$ -glucuronidase	-	V <sup>a</sup>
malonate	V <sup>a</sup>	V <sup>b</sup>
<i>N</i> -acetyl- $\beta$ -glucosaminidase	V <sup>a</sup>	-
<b>Utilisation of:</b>		
sucrose	V <sup>a</sup>	+
stachyose	V <sup>b</sup>	V <sup>a</sup>
<i>N</i> -acetyl-D-galactosamine	-	V <sup>a</sup>
3-methyl glucose	V <sup>a</sup>	+
D-aspartic acid	V <sup>b</sup>	+
D-serine	V <sup>a</sup>	+
L-pyroglutamic acid	V <sup>a</sup>	V <sup>a</sup>

pectin	V <sup>b</sup>	V <sup>a</sup>
quinic acid	-	V <sup>a</sup>
p-hydroxy-phenylacetic acid	+	-
citric acid	V <sup>a</sup>	+
α-keto-glutaric acid	-	V <sup>a</sup>
D-malic acid	V <sup>a</sup>	V <sup>a</sup>
α-hydroxy-butyric acid	V <sup>a</sup>	-
β-hydroxy-D,L-butyric acid	V <sup>b</sup>	-
formic acid	V <sup>b</sup>	-
<b>Resistant to:</b>		
fusidic acid	+	-
D-serine	+	V <sup>b</sup>
Troleandomycin	V <sup>a</sup>	V <sup>b</sup>
nalidixic acid	V <sup>a</sup>	V <sup>b</sup>
potassium tellurite	V <sup>b</sup>	V <sup>b</sup>

+, positive reaction; -, negative reaction; V, variable within species; <sup>a</sup>, positive for type strain; <sup>b</sup>, negative for type strain. \* *Leclercia adecarboxylata* and *Leclercia tamurae* sp. nov.

#### 5.2.9. Fatty acid and methyl ester analysis:

##### *Scandinavium*

The major fatty acids were identified as C<sub>16:0</sub>, C<sub>18:1</sub> ω7c, C<sub>17:0</sub> cyclo, and summed feature 3 (C<sub>16:1</sub> ω7c and/or C<sub>16:1</sub> ω6c). Table 30 details the FAMES profiles for all species of *Scandinavium*. The fatty acid profiles of all strains of all species analysed were very similar.

**Table 30:** The major fatty acid methyl ester (FAME) average % peaks and standard deviation recorded for species of *Scandinavium*. (1) *Scandinavium goeteborgense* CCUG 66741<sup>T</sup>, (2) *Scandinavium hiltneri* (*n* = 2), (3) *Scandinavium manionii* (*n* = 2), (4) *Scandinavium tedordense* TWS1a<sup>T</sup>.

<b>Saturated fatty acids</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
C <sub>12:0</sub>	2.9	3.2 (± 0.18)	3.2 (± 0.14)	3.5
C <sub>14:0</sub>	5.2	7.5 (± 0.18)	7.0 (± 0.95)	7.4
C <sub>16:0</sub>	30.8	31.0 (± 0.37)	31.1 (± 0.95)	32.8
<b>Unsaturated fatty acids</b>				
C <sub>18:1</sub> ω7c	15.6	16.5 (± 1.48)	16.2 (±0.96)	14.7
<b>Cyclopropane fatty acids</b>				
C <sub>17:0</sub> cyclo	16.4	14.7 (± 0.35)	14.7 (± 4.36)	13.1
<b>Summed features</b>				
2: C <sub>14:0</sub> 3-OH and/or iso-C <sub>16:1</sub>	7.7	7.7 (± 0.24)	7.7 (± 0.07)	8.1
3: C <sub>16:1</sub> ω7c and/or C <sub>16:1</sub> ω6c	17.9	16.5 (± 0.84)	16.3 (± 3.37)	18.0

### *Dryocola*

The major fatty acids for both novel species were C<sub>16:0</sub>, C<sub>18:1</sub> ω7c, C<sub>17:0</sub> cylco, summed features 2 (C<sub>14:0</sub> 3-OH and/or iso-C<sub>16:1</sub>) and 3 (C<sub>16:1</sub> ω7c and/or C<sub>16:1</sub> ω6c). While the FAMES profiles were mostly consistent between genera, differences between the novel genus, *Cedecea* and

*Buttiauxella* were observed in the amounts of C<sub>14:0</sub> and C<sub>16:0</sub>. The FAME profiles for all the strains including *Cedecea* and *Buttiauxella* can be seen in Table 31, *Buttiauxella* results were obtained from (Kämpfer, Meyer and Müller, 1997).

**Table 31:** The average percentage of peak areas making up the fatty acid methyl ester composition of *Dryocola* gen. nov., *Cedecea* and *Buttiauxella* species. (1) *Dryocola boscaweniae* sp. nov. (n = 3), (2) *Dryocola claudunensis* sp. nov. (n = 2), (3) *Cedecea davisae* (LMG 7862<sup>T</sup>), (4) *Cedecea lapagei* (LMG 7863<sup>T</sup>), (5) *Cedecea neteri* (LMG 7864<sup>T</sup>), (6) *Buttiauxella agrestis* (n = 13), (7) *Buttiauxella ferragutiae* (n = 5), (8) *Buttiauxella gavinae* (n = 11), (9) *Buttiauxella brennerae* (n = 7), (10) *Buttiauxella izardii* (n = 12), (11) *Buttiauxella noackiae* (n = 14), (12) *Buttiauxella warmboldiae* (n = 5). Data for 6 – 12 taken from (Kämpfer, Meyer and Müller, 1997).

<b>Saturated fatty acids</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>	<b>12</b>
C <sub>12:0</sub>	3.9 (± 0.6)	4.2 (± 1.2)	4.8	2.3	3.9	2.9 (± 1.4)	1.0 (± 1.1)	3.0 (± 0.6)	2.6 (± 0.7)	2.6 (± 3.0)	2.9 (± 1.0)	3.3 (± 0.8)
C <sub>14:0</sub>	5.8 (± 0.7)	6.9 (± 0.4)	4.6	2.6	3.9	6.1 (± 1.8)	6.5 (± 1.2)	6.9 (± 1.2)	6.7 (± 0.8)	6.4 (± 0.8)	6.6 (± 1.7)	5.0 (± 0.4)
C <sub>16:0</sub>	34.0 (± 0.8)	31.3 (± 1.3)	35.9	33.1	29.6	26.0 (± 8.77)	27.1 (± 2.1)	26.1 (± 3.1)	24.9 (± 3.1)	26.8 (± 5.0)	28.3 (± 6.6)	25.6 (± 2.1)
C <sub>17:0</sub>	0.4 (± 0.1)	0.5 (± 0.1)	-	0.7	1.6	1.1 (± 3.4)	-	1.4 (± 2.4)	1.3 (± 2.3)	1.4 (± 2.9)	1.3 (± 1.6)	1.0 (± 0.1)
<b>Unsaturated fatty acids</b>												
C <sub>18:1 ω7c</sub>	11.3 (± 1.8)	16.0 (± 0.4)	11.6	20.2	21.2	15.8 (± 4.2)	15.4 (± 2.2)	14.0 (± 2.3)	15.3 (± 2.9)	15.5 (± 3.6)	14.9 (± 3.1)	18.0 (± 2.6)

<b>Cyclopropane fatty acids</b>												
C <sub>17:0</sub> cylco	13.6 (± 3.9)	12.6 (± 3.9)	10	11.07	9.34	10.75 (± 9.68)	7.64 (± 2.01)	12.67 (± 4.34)	10.59 (± 7.47)	11.12 (± 4.96)	14.03 (± 8.76)	8.75 (± 1.73)
C <sub>19:0</sub> ω8c	0.93 (± 0.76)	2.1 (± 1.0)	0.9	2.5	0.9	-	-	-	-	-	-	-
<b>Summed features</b>												
2: C <sub>14:0</sub> 3-OH and/or iso-C <sub>16:1</sub>	9.3 (± 1.0)	8.8 (± 1.5)	10.1	9.6	8.5	9.1 (± 3.4)	9.1 (± 1.1)	9.4 (± 1.5)	9.3 (± 1.5)	9.1 (± 2.5)	8.5 (± 2.1)	8.8 (± 0.5)
3: C <sub>16:1</sub> ω7c and/or C <sub>16:1</sub> ω6c	19.9 (± 2.9)	15.8 (± 1.8)	14.3	11.3	15.8	22.5 (± 12.3)	30.4 (± 4.9)	20.8 (± 5.6)	21.7 (± 6.1)	20.3 (± 5.3)	17.8 (± 8.0)	25.6 (± 3.1)
5: C <sub>18:2</sub> ω6,9c and/or C <sub>18:0</sub> ante	0.7 (± 0.4)	0.5 (± 0.4)	1.4	0.6	0.5	-	-	-	-	-	-	-

### *Leclercia and Sylvania*

Based on the results generated by the Sherlock Microbial Identification System Version 6.4 (MIDI Inc.), the key fatty acids were C<sub>16:0</sub>, C<sub>18:1</sub> ω7c, summed feature 3 (C<sub>16:1</sub> ω7c and/or C<sub>16:1</sub> ω6c). The fatty acid profiles for each strain can be seen in Table 13. Minor differences can be observed between C<sub>18:1</sub> ω7c which is higher in *Leclercia* species and summed feature 3 (C<sub>16:1</sub> ω7c and/or C<sub>16:1</sub> ω6c) which is higher in *Sylvania* gen. nov. species.

**Table 32:** The major fatty acid methyl ester (FAME) average % peaks and standard deviation for *Leclercia* and *Sylvania* gen.nov. (1) *Leclercia adecarboxylata* (n = 2), (2) *Leclercia tamurae* (n = 4), (3) *Sylvania hatchlandensis* H19S6<sup>T</sup>, (4) *Sylvania confinis* H4N4<sup>T</sup>

<b>Saturated fatty acids</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
C <sub>12:0</sub>	3.7 (± 0.0)	3.8 (± 0.2)	3.5	2.9
C <sub>14:0</sub>	5.3 (± 0.1)	5.4 (± 0.1)	5.5	5.2
C <sub>16:0</sub>	26.4 (±1.1)	28.4 (± 1.1)	25.8	24.4
<b>Cyclopropane fatty acids</b>				
C <sub>17:0</sub> cyclo	7.1 (± 1.9)	9.3 (± 1.9)	6.7	2.8
<b>Unsaturated fatty acids</b>				
C <sub>18:1</sub> ω7c	21.4 (± 0.2)	20.0 (± 0.2)	18.4	14.9
<b>Summed features</b>				
2: C <sub>14:0</sub> 3-OH and/or iso-C <sub>16:1</sub>	7.5 (± 0.0)	8.7 (± 0.0)	7.4	7.8
3: C <sub>16:1</sub> ω7c and/or C <sub>16:1</sub> ω6c	24.3 (± 1.2)	22.0 (± 1.2)	27.2	35.7

#### 5.2.10. Virulence genes analysis of *Scandinavium* species:

Using the parameters provided, concise inference to the genes presents in these bacteria with results of BLASTP comparisons to the VFDB demonstrated they possess 210–237 virulence genes. Furthermore, several interesting enzymes were identified in KofamKOALA using the relevant BRITE protein family identifications from the KEGG Mapper Reconstruction results. The most noteworthy enzymes identified against the databases included pectinase, adhesin/invasion protein homologues, proteins related to the assembly and utilisation of flagella and pili and the core genes required for a T6SS and associated secreted proteins. The presence of a T6SS was also identified by searching against the SecReT6, database with 156–208 different genes being aligned when querying the annotated genomes. Alignments for the membrane complex (*TssJ*, *TssM*, *TssL*, and *TagL*), baseplate (*TssK*, *TssF-G*, and *TssE*), spike (*PAAR* and *TssI*), sheath and tube (*TssB*, *TssC*, and *TssD*) and the distal end (*TssA*) were all identified with high sequence identity, namely to sequences from members of genera such as *Enterobacter*, *Klebsiella* and *Yersinia* for all strains, excluding H17S15<sup>T</sup> (the type strain of *S. manionii*) which lacked the membrane complex. The majority of genes required for a functioning T2SS were also identified although either *gspO* or *gspS* appear to be absent. Unsurprisingly, given the clinical importance of the type species, many proteins associated with disease development in humans were also identified.

OrthoFinder assigned 34,172 genes (96.5 % of total) to 4,957 orthogroups. 50 % percent of all genes were in orthogroups with 8 or more genes (G50 = 8) and were contained in the largest 2,044 orthogroups (O50 = 2,044). There were 3,222 orthogroups with all species present and 2,867 of these consisted entirely of single-copy genes. Next the virulence genes from the VFDB comparison were compared and OrthoFinder assigned 1,187 genes (99.6 % of total) to 145 orthogroups. Fifty percent of all genes were in orthogroups with 8 or more genes (G50 = 8) and were contained in the largest 55 orthogroups (O50 = 55). There were 100 orthogroups with all species present and 62 of these consisted entirely of single-copy genes. This meant the virulence genes identified could be used to infer the phylogenomic position of all the species of *Scandinavium* investigated in this work.

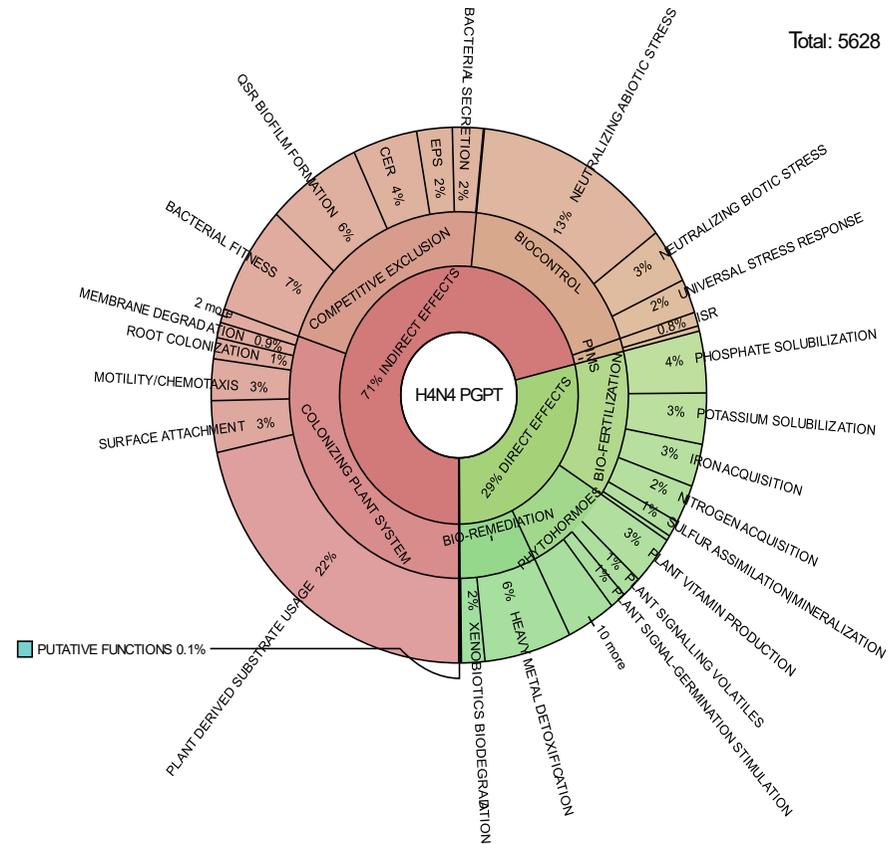
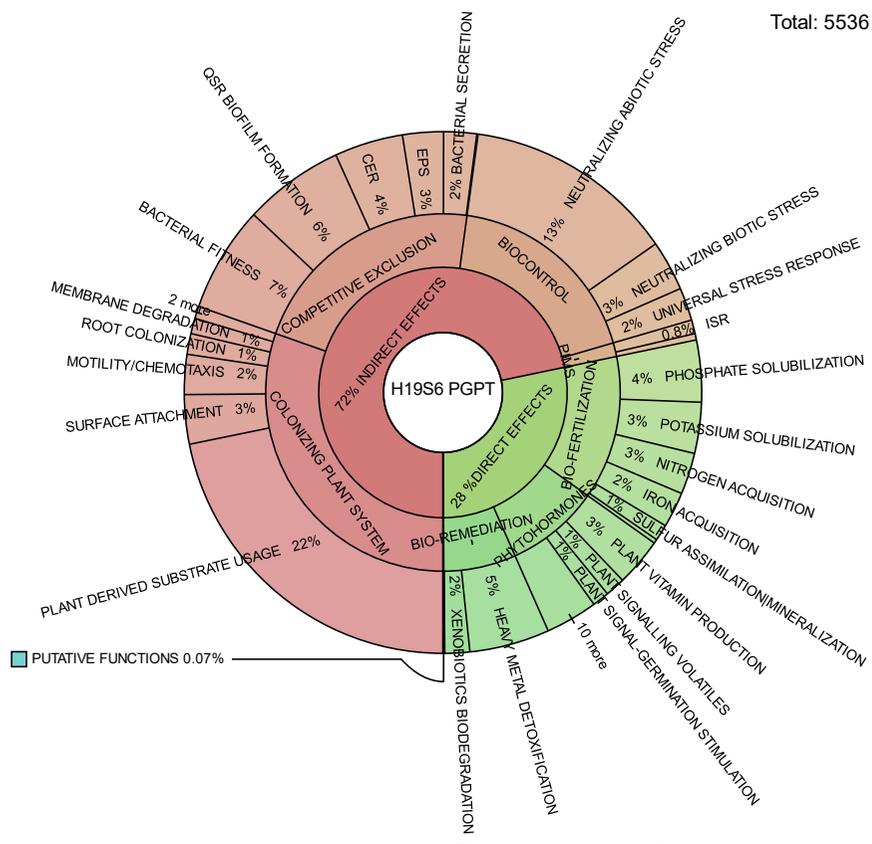
Finally, BLAST results from the RAST server against each novel genome demonstrated that the novel quinolone resistance pentapeptide repeat protein QnrB96 is present with high

homology in all the *Scandinavium* species. The homologue in *S. hiltneri* sp. nov. showed 95 % sequence identity to the complete amino acid sequence, while *S. manionii* sp. nov. showed 96 % and *S. tedordense* sp. nov. showed 94 % similarity. Unsurprisingly the highest homology was observed in H5W5, the strain determined to belong to *S. goeteborgense*, with 99 % sequence identity. The majority of strains displayed 7–11 amino acid substitutions in the *qnrB* protein sequence when compared to the type strain of *S. goeteborgense*, excluding H5W5 which had a single substitution.

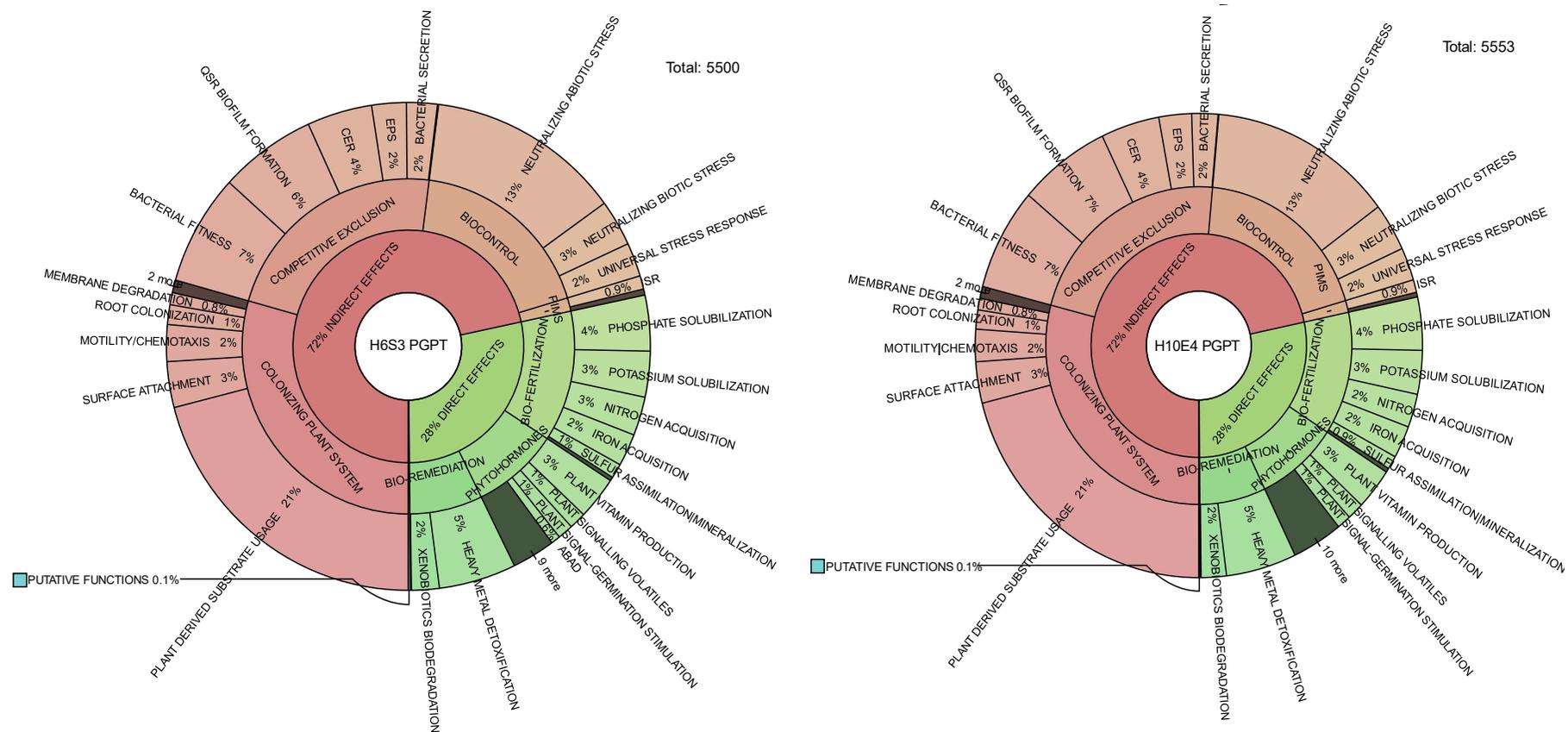
#### 5.2.11. Plant growth promoting genes analysis of *Leclercia* and *Silvania* species:

To investigate the potential of *L. adecarboxylata*, *L. tamurae* sp. nov. and species of *Silvania* gen. nov. as Plant growth promoting bacteria playing a positive role in the soil, their PGPT were investigated computationally. The results from the DIAMOND MEGAN pipeline comparison against the PLant-associated BActeria web resource (PLaBAse) database revealed larger numbers of important plant interaction proteins through the PGPT viewer and KEGG orthology viewer. The resulting PGPT data showed that each submitted annotated genome had between 5,500 – 5,638 PGPTs aligned to known proteins. The majority produced indirect effects such as stress relief and biocontrol, competitive exclusion and genes involved in colonising the plant system. Of the direct effects, the main categories of the genes were involved in bioremediation, phytohormone production and biofertilisation. The krona plots produced show the PGPT possessed by the type strains of the novel species and *L. adecarboxylata* H10E4 (Figures. 46 and 47). Traits of interest included potassium and phosphate solubilisation, nitrogen and iron acquisition, sulphur assimilation and carbon dioxide fixation, features which all directly aid plant growth by increasing nutrient availability. Thirteen percent of the PGPT involved abiotic stress responses to neutralise salinity, osmotic, nitrosative/oxidative, herbicidal, and acidic stress, which are predisposing environmental factors in decline disease (Denman *et al.*, 2022). It has been demonstrated previously that highly acidic soils are known to contribute to AOD symptoms (Brown *et al.*, 2018), especially in parkland systems where many of the strains in the present study were isolated from. A small number of zinc heavy metal resistance genes responsible for *L. adecarboxylata* MO1s plant growth-promoting association (Kang *et al.*, 2021) were identified in all species, although most of the heavy metal resistance genes were related to iron. Few differences could be seen between the *Leclercia* and *Silvania* gen. nov. species although H4N4<sup>T</sup> had more alignments

and the largest number of PGPTs identified. However, given their highly conserved AAI values of 90 – 91 %, this is unsurprising and a further implication of their phylogenetic relatedness.



**Figure 46:** Krona plot representation of the major plant growth-promoting traits found in *Silvania hatchlandensis* sp. nov. (H19S6T) and *Silvania confinis* sp. nov. (H4N4T). Identification of PGPTs was performed by BlastP and HMMER annotation against the PGPT-BASE. Text files of the annotation were downloaded, and Krona plots were made using the 'ktImportText' command in Bioconda. Depth of annotation is shown to level three of six, excluding pathways, gene names and accession numbers. QSR = Quorum sensing response, CER = Cell envelope remodelling, EPS = Exopolysaccharide production, PIMS = Plant immune system stimulation and ISR = Induction of systemic resistance.



**Figure 47:** Krona plot representation of the major plant growth-promoting traits found in *Leclercia adecarboxylata* (H10E4) and *Leclercia tamurae* sp. nov. (H6S3<sup>T</sup>). Identification of PGPTs was performed by BlastP and HMMER annotation against the PGPT-BASE. Text files of the annotation were downloaded, and Krona plots were made using the 'ktImportText' command in Bioconda. Depth of annotation is shown to level three of six, excluding pathways, gene names and accession numbers. QSR = Quorum sensing response, CER = Cell envelope remodelling, EPS = Exopolysaccharide production, PIMS = Plant immune system stimulation, ISR = Induction of systemic resistance and ABAD = Abscisic acid degradation

The assessment made for each strain annotated genome comparison against plant bacterial only interaction factors (proteins) or PIFAR, suggested that the novel species were all capable of interaction with plants, but the identified interaction factors were related to virulence. Between 31 – 32 % of factors were toxins (syringomycin and toxoflavin), 17 – 19 % were exopolysaccharides (namely amylovoran), 8 – 9% of *Silvania* gen. nov. and 11 – 12 % of *Leclercia* factors were for detoxification (of plant compounds such as isothiocyanate), and ~15 % were adhesion and metabolism genes. Between 0.6 – 0.9 % (*Leclercia*) and 2 % (*Silvania* gen. nov.) of the identified bacterial plant interaction markers were PCWDE which are key markers of phytopathogens. The features identified through PIFAR such as EPS, toxins and PCWDE implicate the novel isolates as having pathogenic potential. These genes products are associated with the invasion, colonisation and degradation of plant tissue (Toth, Pritchard and Birch, 2006). However, many of these genes are also used by PGPB for the colonisation of plants, where they continue to have a positive effect. Nonetheless, the identified pathogenicity traits complicate the potential role of these isolates as PGPB concerning oak (Monteiro *et al.*, 2012).

Results from the comparison to the virulence factor database (VFDB), however showed 126 - 140 proteins from the novel strains were aligned to known virulence proteins from other pathogens with the vast majority related to motility, immunomodulation and adhesion. Some T6SS effector delivery system proteins were identified, although no complete set of the assembly proteins and no secreted effector proteins were identified in the alignments. These results imply that the novel isolates have low pathogenic potential, or if they are pathogenic they possess novel ability which cannot be identified through database comparison.

Overall, it can speculate that these isolates may have a positive role in the rhizosphere through several important direct PGPT genes such as heavy metal detoxification, biofertilisation and phytochemical signalling which all aid plant growth and resilience. Alongside gene products with/predicted to have direct effects are other positive PGPTs such as stress relief for osmotic, heat, salinity and competitive exclusion genes which indirectly benefit plants. However, based on the alignments made in both the VFDB and through the PIFAR database the novel isolates here all contain genes related to virulence such as motility, adhesion, and Immune modulation genes. These genes could implicate a potential for pathogenicity, although all genes identified could also be utilised by PGPB for colonisation of

the plant endosphere. While it cannot be concluded on the precise role these isolates play in this niche, it is probable based on comparison to *L. adecarboxylata* MO1, that the novel strains isolated in the present study promote plant growth through their action in the rhizosphere, especially in relation to heavy metal detoxification (Kang *et al.*, 2021).

### 5.3. Conclusions

Based on the genomic, genotypic, chemotaxonomic and phenotypic data this body of work represents the isolation and identification of eight new species and two new genera. The full protologs including the emended genus descriptions for *Leclercia* and *Scandiavium* can be found in the Supplementary data file, along with the three related publications.

### 5.4. Discussion

The work presented in this chapter shows the application of the polyphasic taxonomic approach to eight new species spread over four genera, two of which are novel. Individually much can be said about each of the genera, but these conclusions are better left to their respective publications. Instead on a general level this chapter demonstrates how little is known about the oak rhizosphere microbiome. This study utilised specific enrichment of one family of bacteria leading to the recovery and identification of eight new species. Altered enrichment of other families could lead to the identification of numerous other new species and genera. Within this work a group of potentially novel plant pathogens have been isolated, which also appear to be present in weeping lesions on other broadleaf hosts, namely *Tilia*, from which a novel species *Brenneria tiliae* has recently been isolated (Kile *et al.*, 2022). This has also led to the isolation of a potentially interesting group of novel PGPB that appear to be predominantly associated with the healthy oak rhizosphere. Based on the computational analysis performed in this work both groups of bacteria deserve further attention based on their ability to interact with plants. The novel *Leclercia* species offers the most interesting avenue for bioremediation, especially considering the ability of *Leclercia adecarboxylata* MO1 to reduce cadmium stress when applied to other plants (Kang *et al.*, 2021).

One of the biggest contributions of this chapter is the clarity given to some rarely isolated genera. The identification of the novel genus *Dryocola*, has begun to elucidate a relation between *Cedecea*, a rarely isolated group of human pathogens, and *Buttiauxella*, a widely distributed and commonly isolated group of bacteria which have been recorded playing the

role of pathogens in humans and PGPB in the rhizosphere of several plants (Wu *et al.*, 2018; Farmer, 2015; Ferragut, Izard and Gavini, 1981). Likewise, the addition of new species of *Leclercia* and *Scandinavium* help reveal the taxonomic position of these infrequently isolated genera, that show low similarity to even their closest taxonomic neighbours. Furthermore, the addition of the new genus *Silvania*, combined with the novel species of *Leclercia* have highlighted the incorrect taxonomic identification of the species *Leclercia pneumoniae*, which itself appears to be a novel, single species genus. Hopefully with the addition of these bacteria, further clarification the order Enterobacterales has been granted and new closely related species and genera will be more clearly defined.

# Chapter 6. Long Read 16S rRNA Gene Sequencing of the Oak Rhizosphere Microbiome

## 6.1. Introduction:

The rhizosphere represents the primary point of contact between plants and soil, and so plays an essential role in plant health (Yu *et al.*, 2019). Microorganisms can have both positive effects, through nutrient mobilisation and phytohormone production as well as antagonistic roles as potential phytopathogens (Raaijmakers *et al.*, 2009). The holistic approach to understanding which bacteria are present or absent in specific rhizosphere samples has become the standard approach in recent years. This is especially apparent within agriculture where soil microbiomes can be significantly altered by crop management practices with equally significant effects on crop yield (Li *et al.*, 2017). However, generally little information is known about how different changes alter the complete soil microbial taxa (Fadiji, Kanu and Babalola, 2021).

The use of whole community comparison allows the discrimination of differences in the microorganisms present, highlighting taxa that are overrepresented in one scenario in comparison to another (Fadiji and Babalola, 2020). This has led to breakthroughs concerning AOD, including the discovery of the shift in the endophytic microbiome of oaks suffering from diseases. Both 16S rRNA gene metabarcoding and a multi-omics approach have identified two distinct microbiomes correlating to oak health (Broberg *et al.*, 2018; Sapp *et al.*, 2016). Recently, community analysis of rhizosphere samples revealed that these differences are not limited to the endophytic oak microbiome. For example, healthy oaks presented non-extreme soils with neutral to slightly acidic pH, moisture content, available carbon and nitrogen and a multitude of bacteria that can provide benefits to the host. Meanwhile, low and mid-stage AOD oaks exhibited lower soil pH and their microbial composition differed significantly from healthy oaks (Pinho *et al.*, 2020). However, these findings were limited to family-level identifications, allowing some broad conclusions based on the majority function of microorganisms at this level but not differentiation based on species-level function. However, they have been further supported by the findings that ammonia-oxidising bacteria (AOB) are significantly associated with asymptomatic oaks and like the previous study this association was correlated with soil pH. Interestingly, denitrifying bacteria did not follow the same correlation but instead were influenced by the soil C:N ratio and the abundance of AOB. This highlighted the potential to utilise specific bacteria to reduce stress on declining oaks by altering C:N ratio and deacidifying soils (Scarlett *et al.*, 2021).

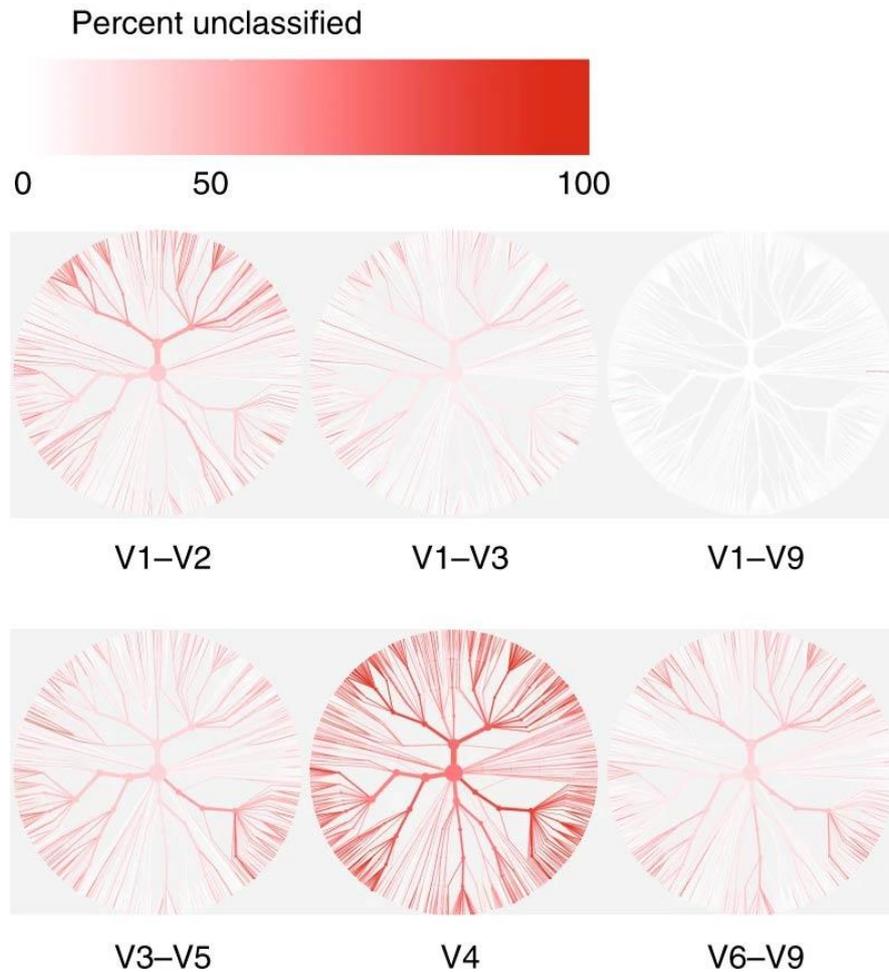
Given these findings does the rhizosphere community influence the decline spiral in which less beneficial rhizosphere microbiomes limit the ability of a tree to respond to negative environmental stimuli? Or does the health status of the tree directly affect the composition of microbial life that its rhizosphere can host? Considering host root exudates significantly shape the community structure of bacteria in the rhizosphere, it is reasonable to expect the latter. Plants sink up to 20 % of their photosynthetic net gain into the soil, which comes at a significant cost to the plant which appears to continuously secrete exudates and low-weight anti-microbial compounds under regulation (Olanrewaju *et al.*, 2019). Stressed trees, however, are known to store carbon. For example, in Norway spruce when net carbon gain is lower than amount utilised, carbon will be preferentially allocated to non-structural carbon reserves (Huang *et al.*, 2021). Interestingly the deposition of high amounts of carbon-based root exudates is associated with young plants, while more mature plants use this carbon in their shoots (Pausch and Kuzyakov, 2018). Moreover, abiotic stress which contributes to the predisposing stage of AOD has well-recorded changes in root exudates and as such the composition of the rhizosphere microbiome (Vives-Peris *et al.*, 2020). From these studies, speculation can be made that the health status of the tree will in some way influence the rhizosphere microbiome and there will be a distinguishable difference between healthy and diseased tree rhizospheres.

But what is the most appropriate method to observe the bacterial rhizosphere and compare them between samples? Previous studies have utilised several different sequencing methods (Fadiji and Babalola, 2020). Of these, PCR amplification and sequencing of genes with different sequence compositions have been routinely used to evaluate microbial species diversity in samples. The 16S rRNA gene has been the gold standard of prokaryotic identification, in part due to its presence in most bacterial, archaeal, and mitochondrial genomes (Gray, 2012; Acinas *et al.*, 2004). Sequencing is generally performed on one of the nine hypervariable regions (Bartram *et al.*, 2011) followed by alignment against a 16S rRNA gene database, such as the SILVA ribosomal RNA gene database project (Quast *et al.*, 2013). However, in more recent years as technological advances in DNA sequencing have reduced the price of sequencing, whole genome shotgun metagenomics has gained favour. Total DNA is extracted from the sample and then fragmented, these fragments are sequenced and

aligned against other databases such as RefSeq (O’Leary *et al.*, 2016) or GenBank (Benson *et al.*, 2013).

While both methods are frequently used to investigate complex environmental microbiomes, there are pros and cons to each approach. A metastudy utilised National Ecological Observatory Network data to compare 16S rRNA gene amplicon sequencing to whole genome shotgun (WGS) metagenomics. Both sequencing methods identified highly similar bacterial phyla, but WGS identified the microbes with higher resolution while also allowing the identification of less dominant members of the microbiome (Brumfield *et al.*, 2020). However, the high biodiversity of WGS can lead to insufficient sequencing depths, which in turn requires more reads for an already expensive process. This biodiversity goes further than the prokaryotic scope of this study, covering all eukaryotic, viral and fragmented relic DNA present as well, which in turn requires further sequencing depth (Semenov, 2021). The level of sequencing required in a sample is immense with studies showing that even 300 Gbp of sequencing data was still not sufficient to represent deep coverage of the soil community (Prosser, 2015). So, while WGS metagenomics is the superior method in the assessment of the microbial community analysis, it was both beyond the scope and financial range of this study. As such 16S rRNA gene sequencing was both the suitable and affordable choice for the assessment of the bacterial community composition of oak rhizosphere soil samples.

However, 16S rRNA gene sequencing is limited in its ability to reach species-level descriptions. Traditional sequencing focuses on one to several of the nine hypervariable regions in the ~1500b bp gene, with different combinations of the hypervariable regions performing better based on sample origin and or study focus (Sirichoat *et al.*, 2021). However, the full 16S rRNA gene sequence can provide high taxonomic resolution, to both species and strain levels (Johnson *et al.*, 2019). A comparisons of species identified from the same dataset for each hypervariable region compared to the full gene can be seen below (Figure 48).



**Figure 48:** Phylogenetic trees on the identified taxonomy for each variable region (V1-9) of the 16S rRNA gene. The colour of each branch correlates to the percentage of sequences within each clade that were not identified at the species level. The figure is adapted from the results presented by Johnson *et al.*, 2019.

These results are not isolated, with species-level findings seen in studies assessing microbial diversity in anaerobic digesters and the spatiotemporal effect of bacterial communities in polluted estuaries (Hongxia *et al.*, 2021; Lam *et al.*, 2020). However, to obtain the full 16S rRNA gene many utilise PacBio long-read sequencing, which is prone to error rates higher than short-read sequencing, with a singular long-read error rate for PacBio 1 being around 11 - 15 % (Rhoads and Au, 2015) though PacBio sequel II sequencing offers better accuracy this comes with a price increase. How can full 16S rRNA gene sequencing for species-level bacterial identification with low error rates be achieved?

An emerging method is synthetic long read (SLR) sequencing, which was first commercialised in 2014, though was not compatible with amplicon sequencing due to the use of DNA identifiers assigned to a well with multiple molecules. More recent SLR methods have since

utilised unique molecular identifiers that allow each molecule to be identified during sequencing (Jeong *et al.*, 2021). Loop Genomics published their commercially available option showing error base rates of 0.005 % over full-length 16S rRNA genes, allowing for species identification from complex samples and exact sequence variant identification in mock microbial communities (Callahan *et al.*, 2020). Likewise, an independent study showed that the use of SLR 16S rRNA gene sequencing was a suitable tool for investigating complex microbiomes such as the gut microbiota in comparison to the V3-V4 variable region, achieving species-level identification (Jeong *et al.*, 2021).

As such, in the scope of this study identification of the bacterial composition of the oak rhizosphere the 16S rRNA gene is more suitable than WGS. To overcome the main limitation of 16S rRNA gene sequencing (obtaining species-level delimitation) SLR sequencing was utilised, which provides a low error rate. It was expected that the AOD symptomatic microbiome would show significantly reduced alpha and beta diversity in comparison to the healthy oak microbiome, due to the inability of the tree to support a healthier more diverse bacterial community. As such this work aimed to use SLR sequencing to compare the microbiome of healthy and diseased oak. This data then would highlight differences in the abundance and diversity of healthy vs diseased oak microbiomes. Differences identified at the species level were then to be used to inform distinct differences in specific bacterial presence and absence between diseased and healthy oaks, along with their associated protein pathways and known environmental roles to infer effects on oak health.

## 6.2. Results

### 6.2.1. Whole Site Composition

When comparing samples based on location and health, the difference between woodland and parkland (site) samples shows the largest effect. For each level of increase in taxonomic rank the shared features become fewer and the unique taxa larger (

Table **33**). Similar effects can be seen in comparisons of woodland healthy to woodland diseased, and parkland healthy to parkland diseased although they are significantly smaller, with each taxonomic rank continuing to share features. The least different of the comparisons

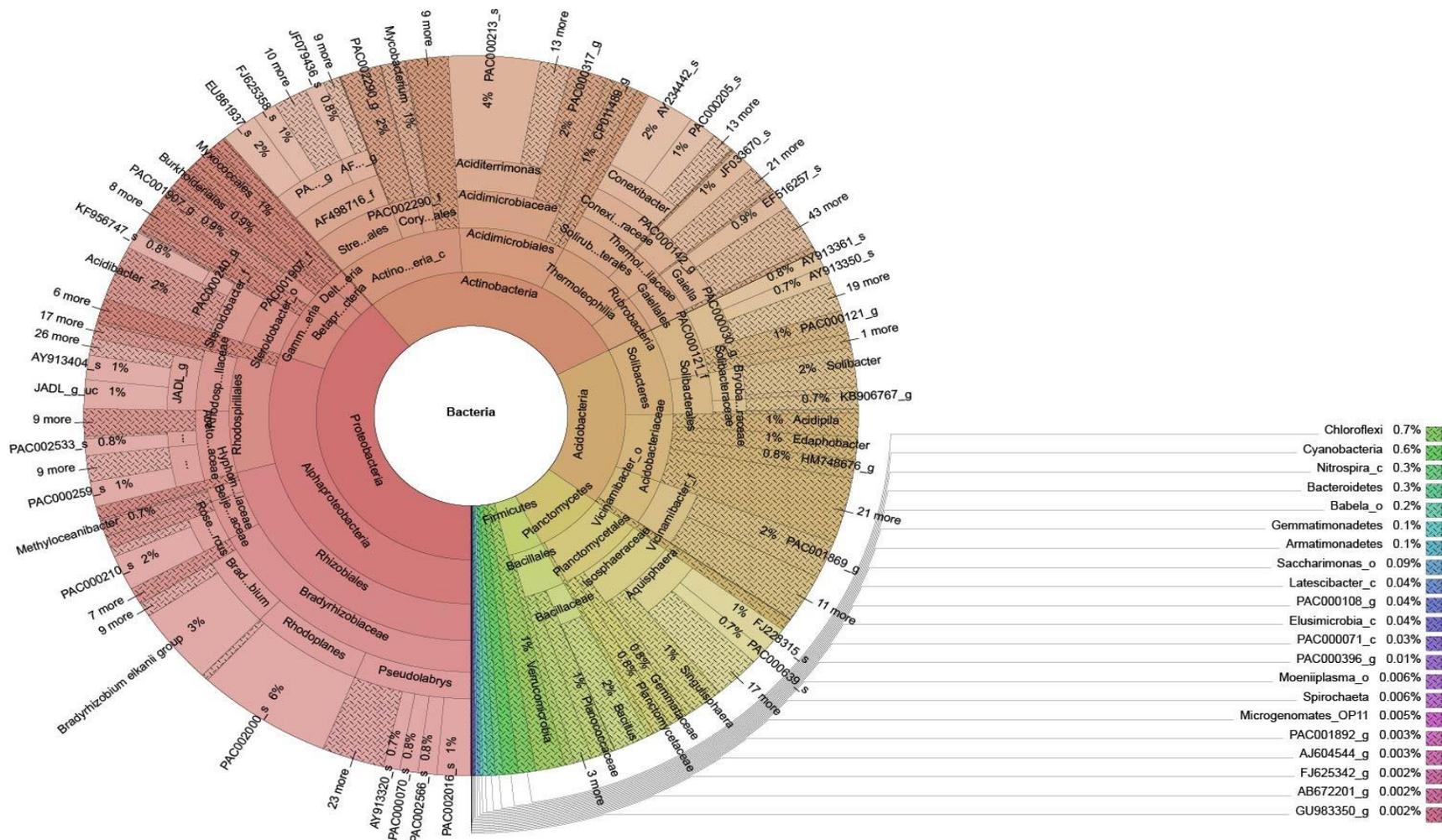
is the complete healthy to diseased, which share the highest number of features and lowest number of unique features, demonstrating the inability to combine the woodland and parkland samples based on health status.

**Table 33:** The number of shared and unique taxa identified at each taxonomic rank between different paired comparisons. The averaged taxonomic compositions of the microbiome taxonomic OTU identifications were used for each set of sample divisions. The estimated taxonomic composition of taxa was set to 1.0 % to remove OTUs that made up less than 1.0 % of the overall microbiome composition. The more closely related a sample the more shared identifications are seen between each comparison, while further distance between samples is indicated by the unique OTUs at each taxonomic ranking.

	Woodland (W) & Parkland (P)			Healthy (H) & Diseased (D)			Woodland Healthy (WH) & Woodland Diseased (WD)			Parkland Healthy (PH) & Parkland Diseased (PD)		
	Shared OTUs	Unique OTUs		Shared OTUs	Unique OTUs		Shared OTUs	Unique OTUs		Shared OTUs	Unique OTUs	
Taxa	W & P	W	P	H & D	H	D	WH & WD	WH	WD	PH & PD	PH	PD
Phylum	5	1	1	6	0	0	5	0	1	7	1	0
Class	11	2	5	16	0	0	13	0	1	15	0	1
Order	11	5	7	19	0	0	17	1	2	18	0	0
Family	11	8	9	21	2	2	18	1	5	16	2	1
Genus	7	18	11	24	2	0	21	4	3	17	3	2
Species	0	14	13	6	2	4	7	6	4	6	2	6

The differences between the sequencing results are visually represent below in the Krona plots (Figure 49 – Figure 54). The site classification exhibits the most differences seen in the

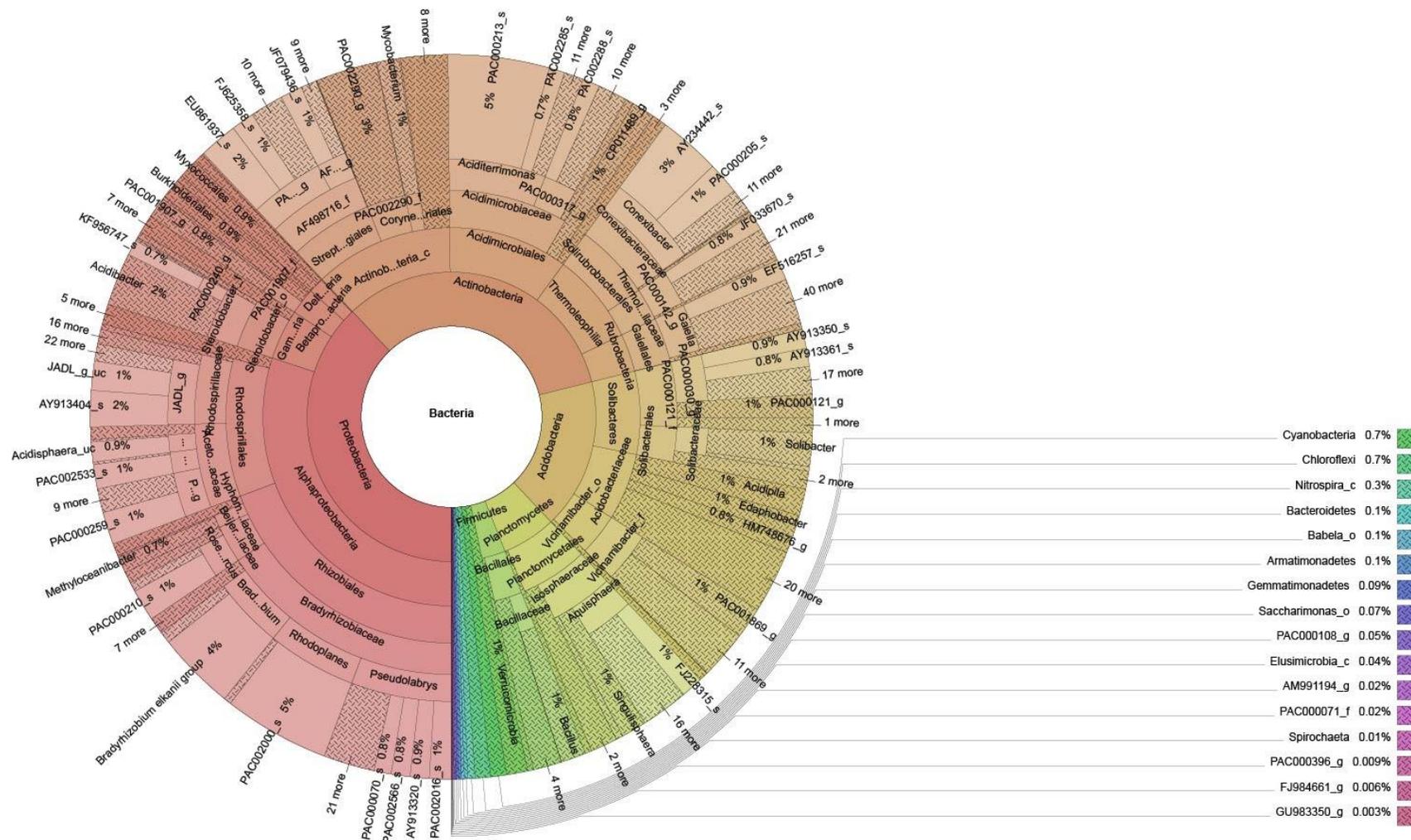
date with the most abundant taxa in parkland being Acidobacter accounting for 26 % of the bacteria identified, while the Proteobacteria are the largest phylum in the woodland dominating 39 % of the identified bacteria (Figure 49 and Figure 50). Other visually apparent differences include the Firmicutes and Verrucomicrobia which are abundant phyla in the parkland making up 15 and 6 % respectively, while in woodland samples they constitute a smaller 4 and 1 % of the site-level identifications. Even within specific phyla, the differences are striking, for example, the Proteobacteria which dominate the woodland are mainly formed of the Alphaproteobacteria, more specifically the *Bradyrhizobiaceae* of the Rhizobiales. Meanwhile, in parkland samples, the distribution between the Alpha, Beta and Deltaproteobacteria is more even. Within the Alphaproteobacteria the *Bradyrhizobiaceae* are still the most abundant (44 %), but the *Hyphomicrobiaceae* are also highly abundant at 27 % while they constitute only 2 % in woodland samples.



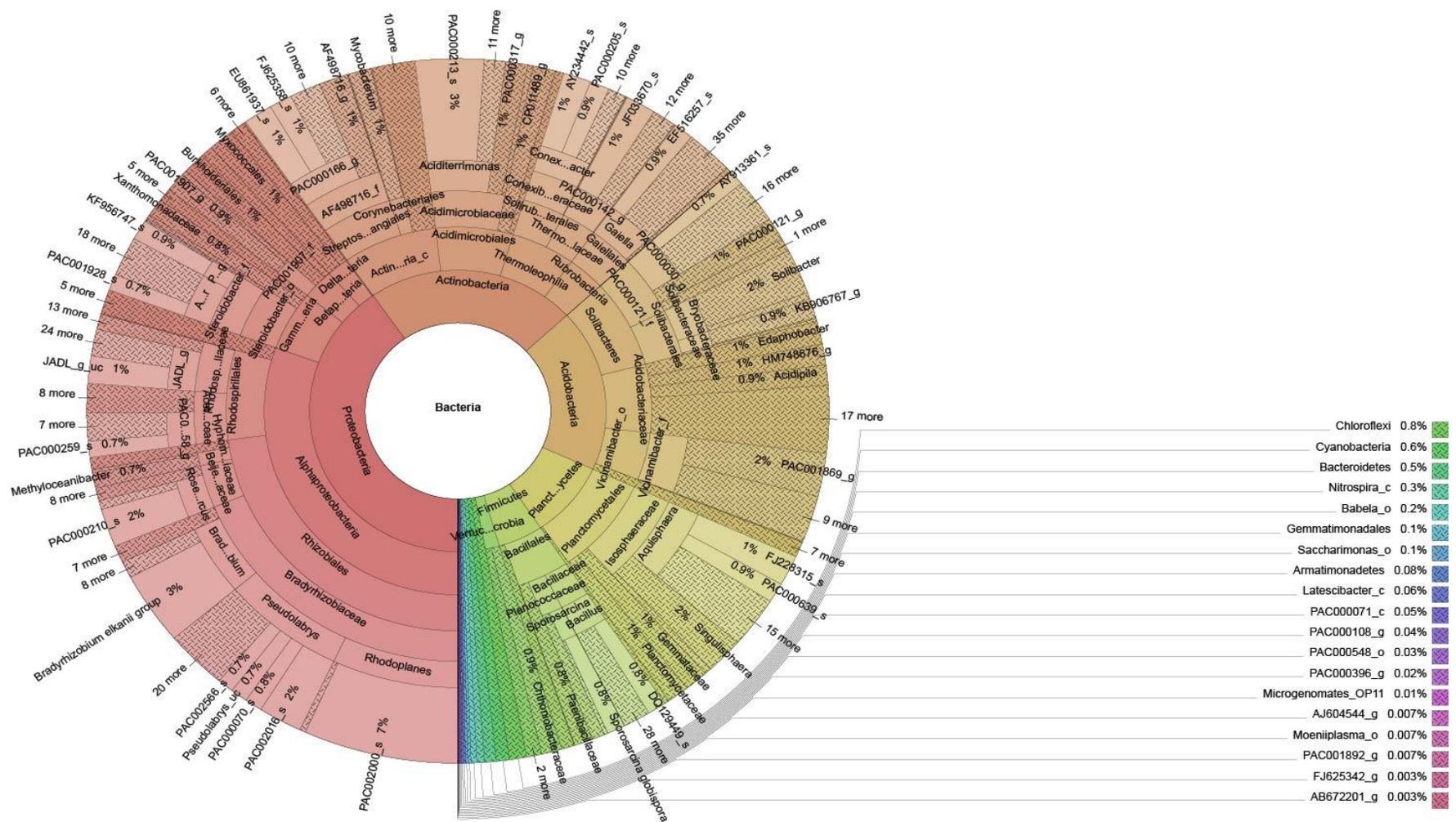
**Figure 49:** Krona plot of the woodland OTU identifications from the EzBiocloud Pipeline. Samples are organised by the abundance with red samples on the bottom left of the plot being the most abundant, going clockwise showing a decrease in abundance, samples take the same order of taxa from phylum to family level. The interactive plot can be accessed through the link in Suppl. Table. S5.



The healthy and diseased woodland samples Krona plots demonstrate how little variation can be observed between the samples based on health status (Figure 51 and Figure 52). Some minor differences can be seen between the percentage abundance of these bacteria, with the *Bradyrhizobiaceae* taking a larger proportion of diseased samples, with the difference being a higher abundance of some other taxa in the healthy samples. The most notable difference is the 9 % representation of the Planctomycetes in diseased samples and the 6 % representation in the healthy samples, demonstrating how similar the composition of the samples was.

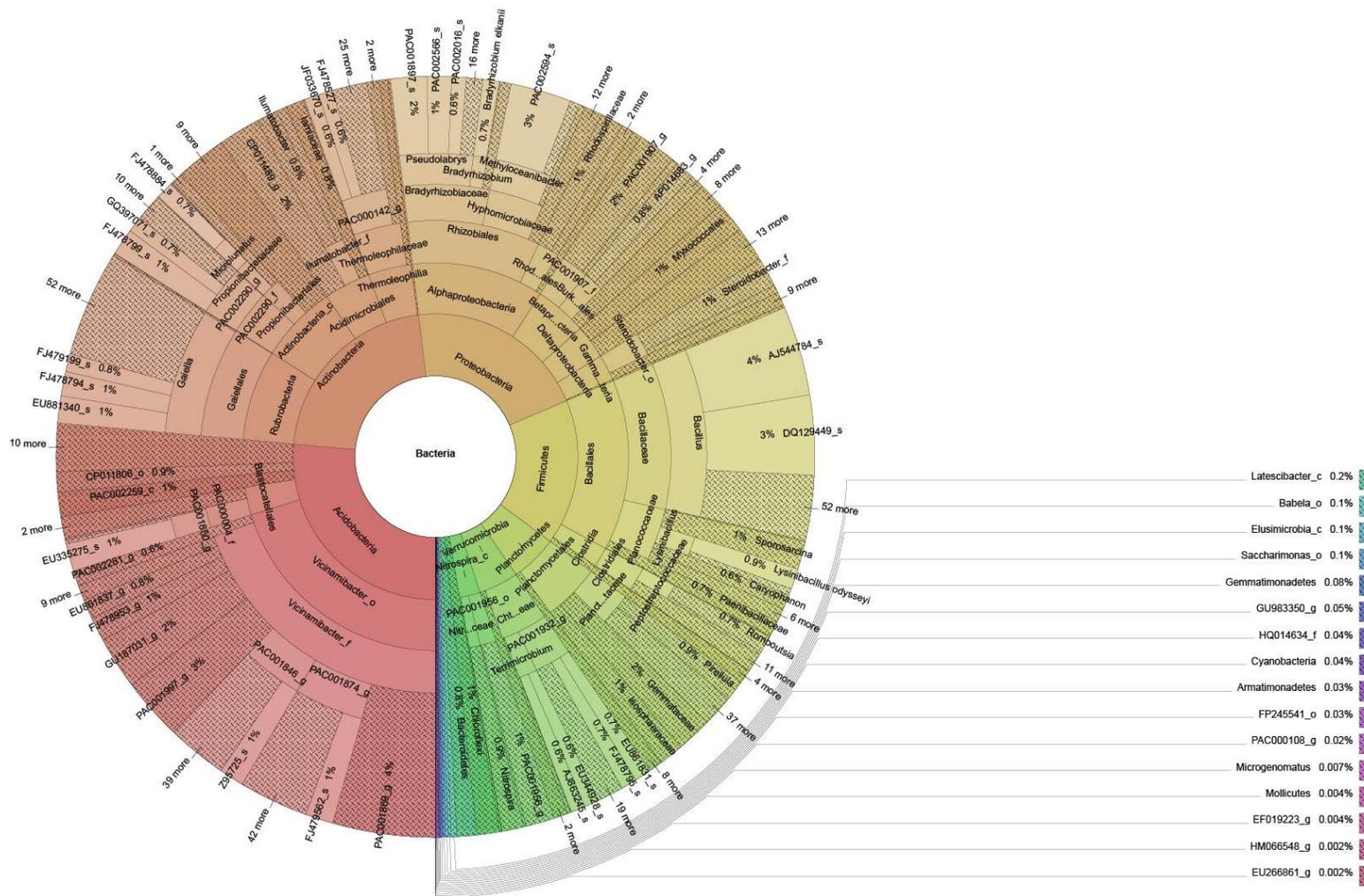


**Figure 51:** Krona plot of the woodland healthy OTU identifications from the EzBiocloud Pipeline. Samples are organised by the abundance with red samples on the bottom left of the plot being the most abundant, going clockwise showing a decrease in abundance, samples take the same order of taxa from phylum to family level. The interactive plot can be accessed through the link in Suppl. Table. S5.

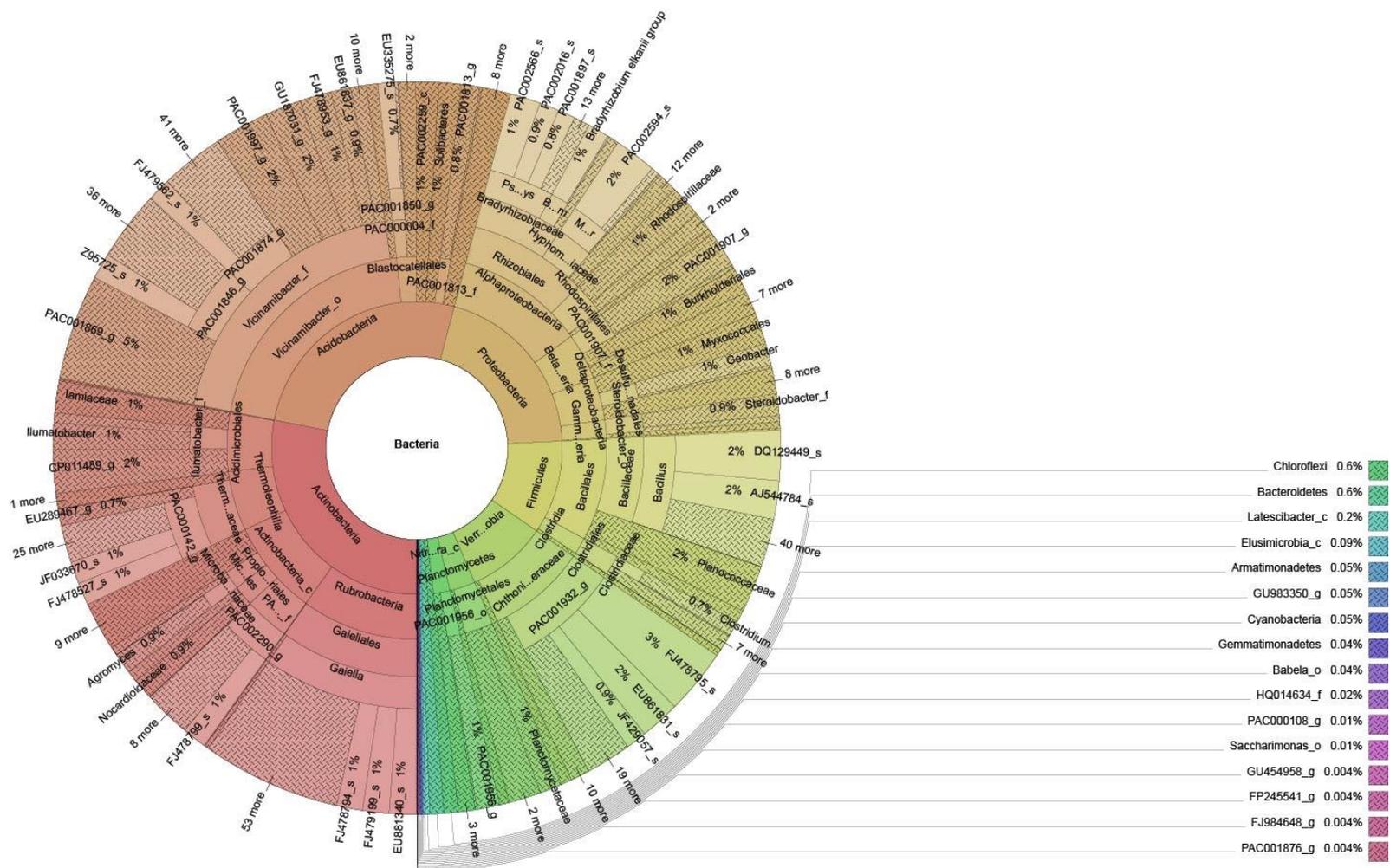


**Figure 52:** Krona plot of the woodland diseased OTU identifications from the EzBiocloud Pipeline. Samples are organised by the abundance with red samples on the bottom left of the plot being the most abundant, going clockwise showing a decrease in abundance, samples take the same order of taxa from phylum to family level. The interactive plot can be accessed through the link in Suppl. Table S5.

The final Krona plots show a larger variation in taxonomic composition can be seen between healthy and diseased trees in the parkland (Figure 53 and Figure 54) than those seen in the woodland trees (Figure 51 and Figure 52). First, the diseased samples show their largest phylum to be the Actinobacteria, while healthy samples show Acidobacteria to be their most abundant phylum. The composition of Proteobacteria is nearly identical between the two samples at 20 %, but the Firmicutes show a larger representation in healthy samples at 16 % while in diseased they only constitute 10 %. Finally, the difference between Verrucomicrobia and Planctomycetes demonstrates the variable nature of the samples with them being 9 and 3 % of phyla in parkland diseased and 4 and 5 % of phyla in parkland healthy.



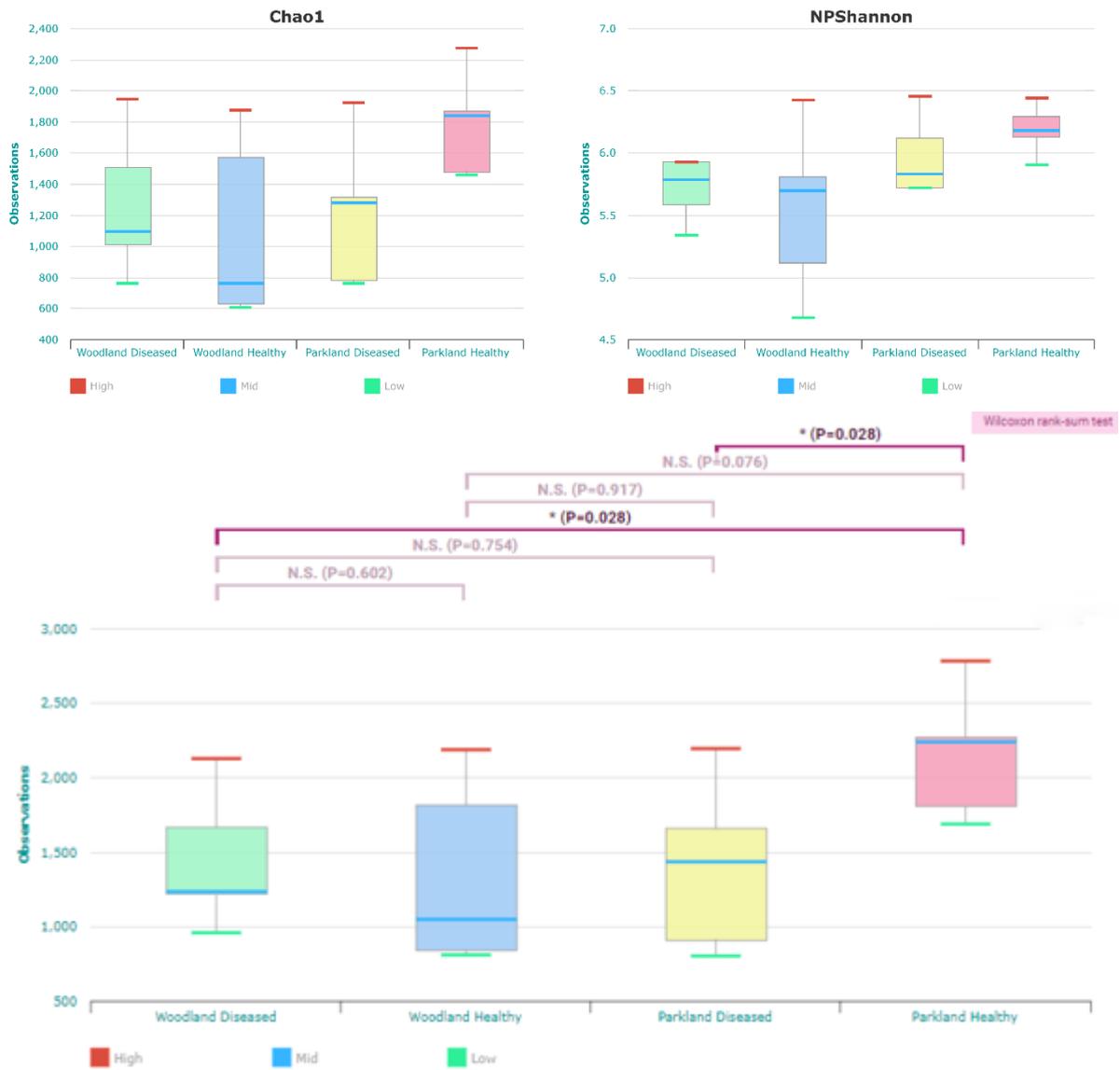
**Figure 53:** Krona plot of the parkland healthy OTU identifications from the EzBiocloud Pipeline. Samples are organised by the abundance with red samples on the bottom left of the plot being the most abundant, going clockwise showing a decrease in abundance, samples take the same order of taxa from phylum to family level. The interactive plot can be accessed through the link in Suppl. Table. S5.



**Figure 54:** Krona plot of the parkland diseased OTU identifications from the EzBiocloud Pipeline. Samples are organised by the abundance with red samples on the bottom left of the plot being the most abundant, going clockwise showing a decrease in abundance. The interactive plot can be accessed through the link in Suppl. Table. S5.

#### 6.2.2. Diversity Statistics:

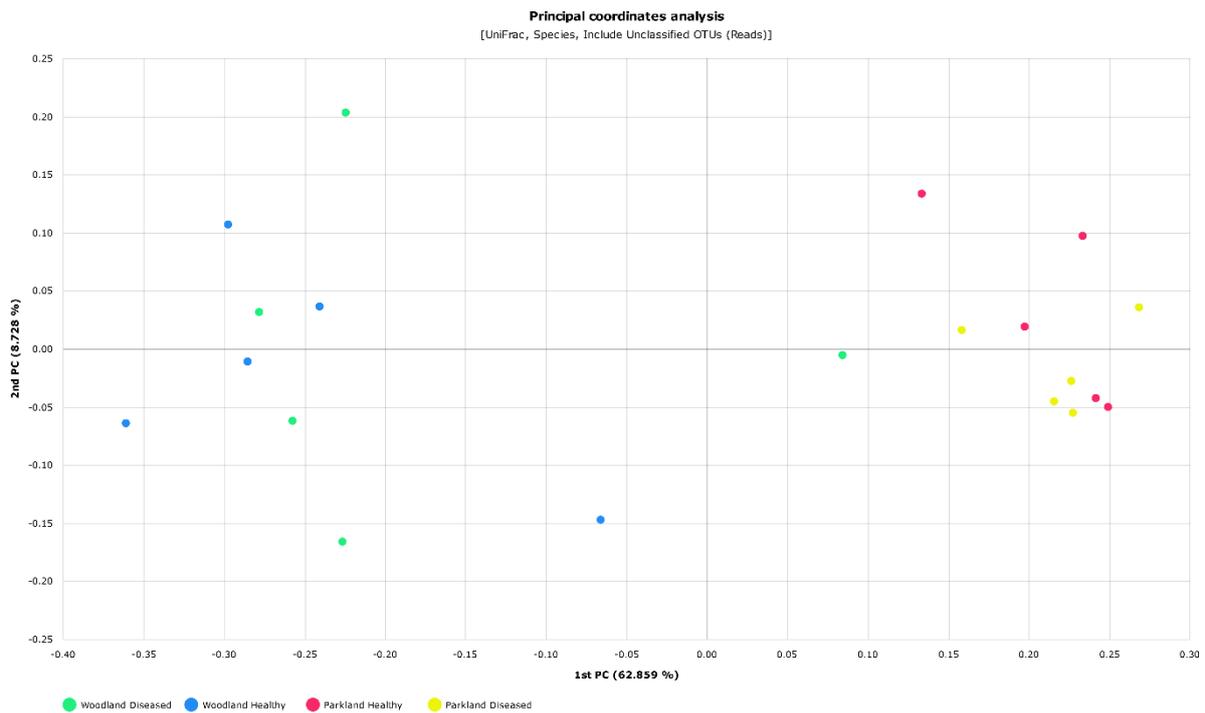
The limited difference between samples is highlighted by the alpha diversity (Figure 55). Species richness performed via Chao1 analysis showed no significant difference between any of the samples based on the different groups. The non-parametric Shannon diversity index, which similarly showed no significant difference between any sample. Using significance to identify the relationship between samples highlights a trend in which the lowest amount of difference is seen between the two diseased samples, then between the two woodland samples, next between the two healthy samples and finally the largest difference between the two parkland samples. Phylogenetic diversity however was shown to be significant between both woodland and parkland diseased samples when compared with parkland healthy samples, indicating an important change between the two based on health status.



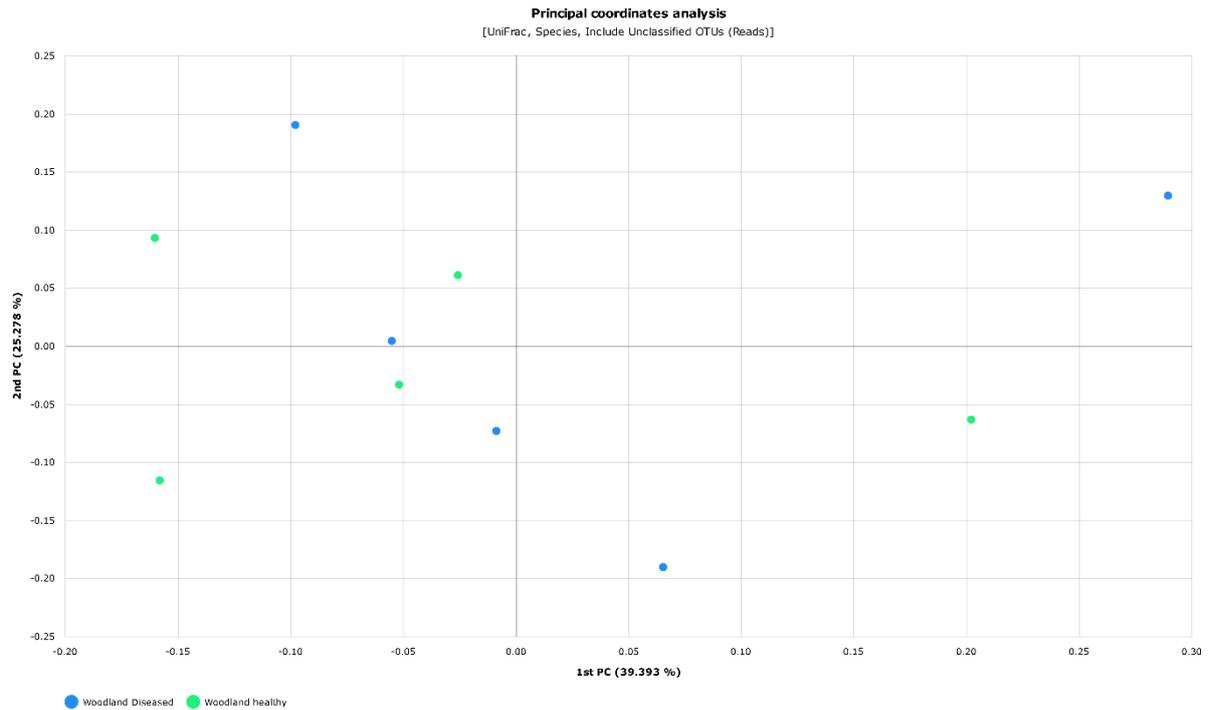
**Figure 55:** The alpha diversity statistics taken for the whole site comparisons, with health as the separator. Chao1 indicated the species richness observed in the samples with no significant difference identified ( $P > 0.05$ ). The non-parametric Shannon diversity index (NPS Shannon) was used, with significance being identified by  $P < 0.05$ , again no significance between any of the comparisons was identified. Phylogenetic diversity was also assessed, and significance bars were included here as significance where  $P < 0.05$  was identified between the woodland diseased and parkland healthy samples and parkland diseased and parkland healthy samples.

The beta diversity statistics, presented through an unweighted UniFrac principal coordinate analysis (PCoA) also revealed interesting trends in the data. The PCoA analysis for the full site data in which a clear separation based on the location of samples is apparent, being clearly presented by the 1<sup>st</sup> PC, though sample H19 does appear to cluster closer to the parkland samples rather than the woodland samples it is associated with (Figure 59). Furthermore, the

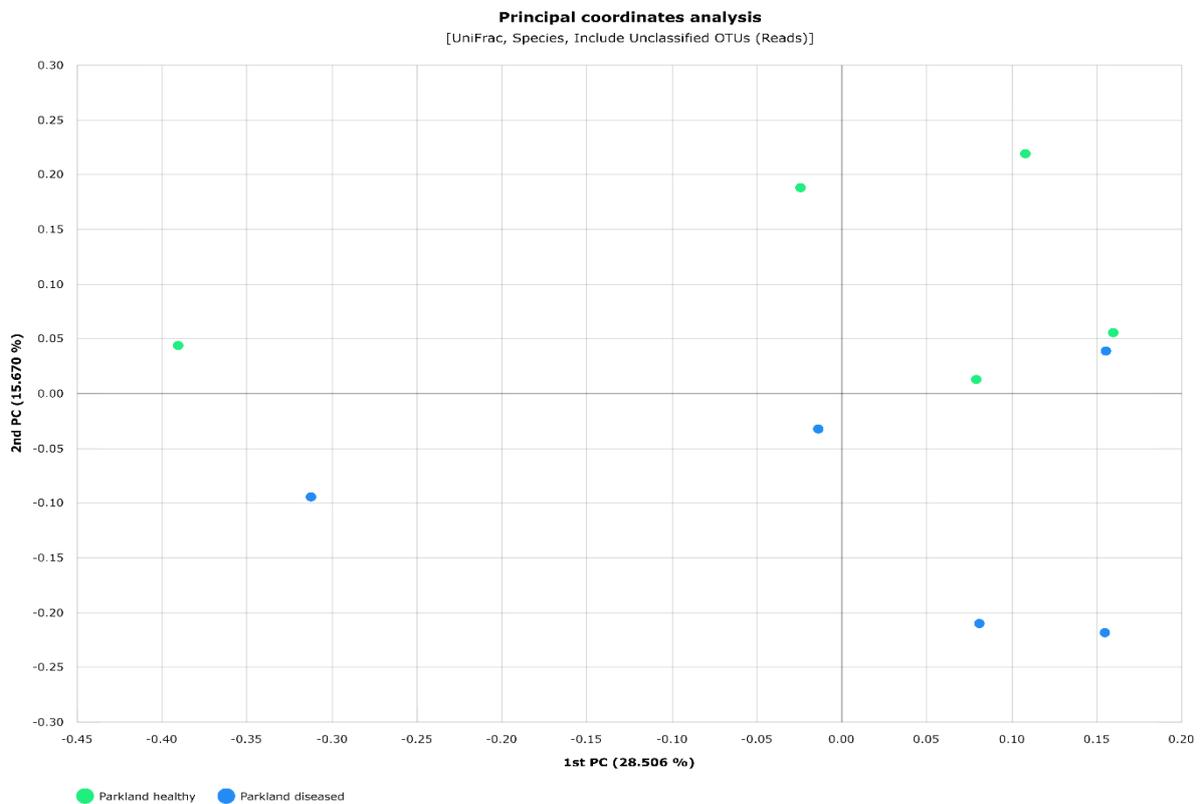
distribution of woodland samples is much larger than parkland samples, indicating a higher level of spread identified in the analysis. the lack of relation based on health status of the woodland samples in comparison with each other is clear in the PCoA, they are highly dispersed with no clear clustering, highlighting the lack of relation between samples (Figure 57). However, differences based on health may be more apparent in parkland samples which show a clearer clustering and distance based on disease status (Figure 58). Healthy samples cluster in a similar fashion but with a higher 2<sup>nd</sup> Principal Coordinate than diseased samples.



**Figure 56:** The beta diversity unweighted UniFrac principal coordinate analysis including unclassified reads for the full site analysis. The 1<sup>st</sup> PC with the largest effect separates samples based on location while the second PC appears to separate samples based on identity (tree of origin). Health is poorly represented in the principle co-ordinates with low separation being shown.

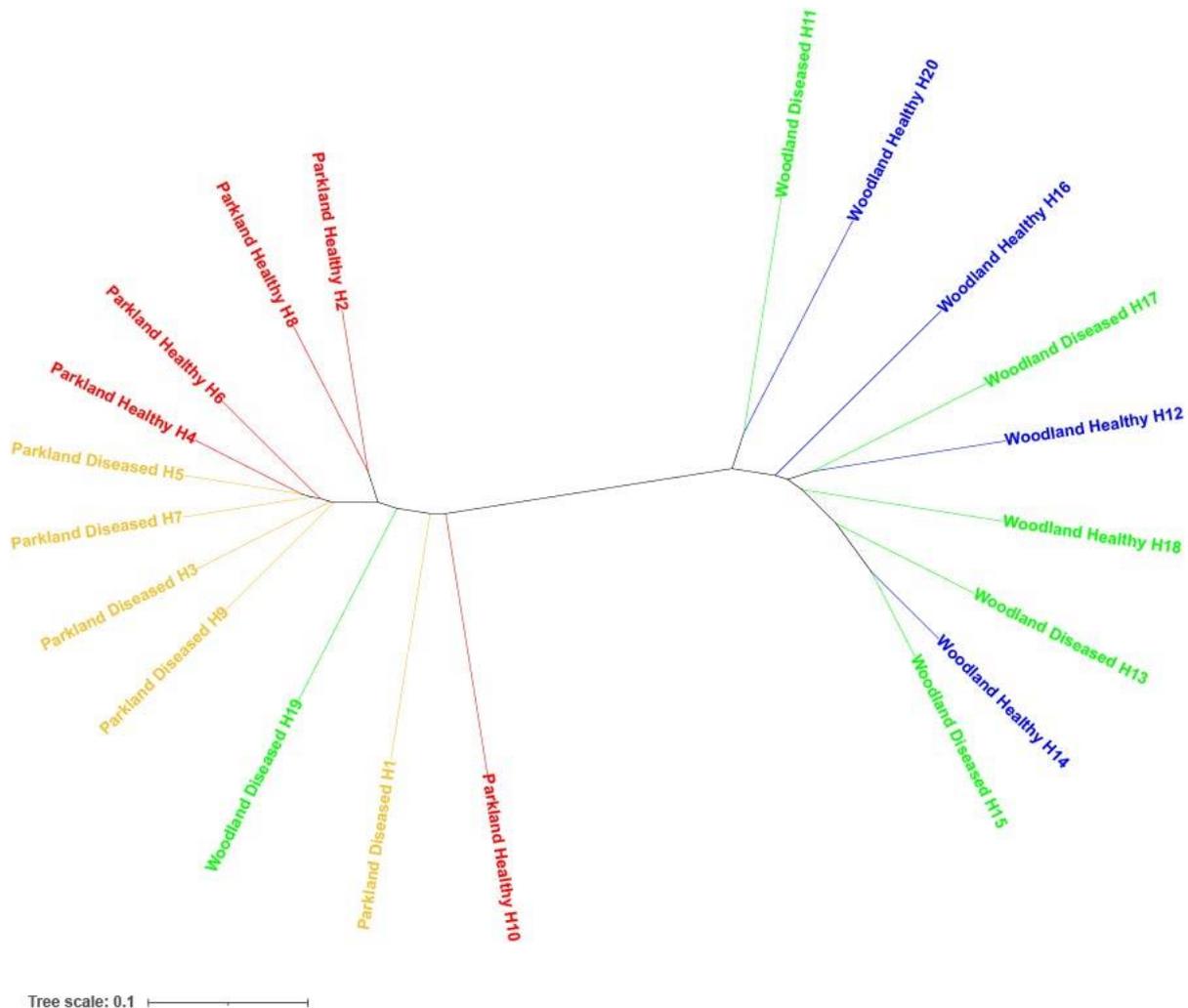


**Figure 57:** The beta diversity unweighted UniFrac principal coordinate analysis including unclassified reads for the woodland diseased and woodland healthy. The 1<sup>st</sup> PC with the largest effect separates samples based on identity (tree of origin) while the second PC appears to separate samples based on health. Sadly despite the well represented distribution of the principal co-ordinates (65 %) low clustering based on the metadata given is seen.



**Figure 58:** The beta diversity unweighted UniFrac principal coordinate analysis including unclassified reads for the parkland diseased and parkland healthy. The 1<sup>st</sup> PC with the largest effect separates samples based on identity (tree of origin) while the second PC appears to separate samples based on health. Despite a low coverage of the principal co-ordinates good clear clustering of samples based on the metadata is seen, with healthy and diseased samples separating clearly.

The unweighted pair group method with arithmetic mean (UPGMA) clustering tree shown also supports the PCoA results, demonstrating that woodland and parkland samples cluster separately, excluding H19 which clusters within the parkland samples (Figure 59). The tree also shows that the clustering of parkland samples appears to be based on health status rather than on their paired sample. For example, H2 and H8, and H6 and H4 are closely related and H5, H7, H3 and H9 similarly show a close relationship. Meanwhile, the woodland samples show random dispersal neither based on health status nor tree pairings.



**Figure 59:** UPGMA unrooted UniFrac tree representation of the phylogenetic clustering of all samples. Parkland diseased samples are shown in yellow, parkland healthy in red, woodland

diseased in green and woodland healthy in blue. Paired samples are H1 and H2, H3 and H4 and so on.

### 6.2.3. LEfSe analysis:

Linear discriminant (LDA) effect size (LEfSe) analysis was used to identify taxonomic biomarkers and functional biomarkers (proteins) with a P value < 0.05 to determine where differences arose for groupings.

#### *Woodland and Parkland*

The differences between site-level samples were highlighted further by identifying 718 species-level bacteria significantly associated with either woodland or parkland rhizosphere samples. Likewise, LEfSe analysis was used to identify functional biomarkers with a statistical difference ( $P < 0.05$ ) via phylogenetic investigation of communities by reconstruction of unobserved states (PICRUSt) between sites. The identification of 419 functional biomarker genes, indicates that the differences seen at the species level are also likely mirrored in the hypothetical function of those species. Due to the number of identified genes, a precise understanding of these differences remains difficult. However, interesting virulence features such as secretion systems could be manually searched for. Type III Secretion System (T3SS) proteins were shown to be significantly associated with parkland samples. Several other secretion system proteins were also found to be significantly associated with parkland rhizosphere soil microbiomes. No significant difference was identified in any module through the LEfSe PICRUSt analysis. The valine, leucine and isoleucine degradation pathways were found to be the only significantly different pathways between woodland and parkland samples.

#### *Woodland Healthy and Woodland Diseased*

Twenty-one differentially abundant bacteria could be identified, five of which appeared to show an association with the healthy rhizosphere samples, while the remaining 16 bacteria associated with diseased rhizosphere samples. Unlike the site-level comparisons, only one of the taxonomic biomarkers was specifically associated with just one healthy sample which was an unclassified member of the *Paenibacillus*. Likewise, LEfSe PICRUSt analysis was also performed, which identified 12 differentially associated proteins, though all showed P values around 0.047, which was near the cut-off. Of these 12 identified proteins, *Salmonella* plasmid virulence protein B appears the most relevant to pathogenicity, with a weak association to

woodland diseased trees. The cytochrome b6f complex was shown to have a low association with diseased rhizosphere samples.

#### *Parkland Healthy and Parkland Diseased*

Differences between the parkland healthy and diseased samples were observed with 107 differentially abundant bacteria being identified. Of 107, only nine appeared to be significantly associated with diseased samples. Those significantly associated appeared to be members of the *Gaiellaceae*, *Acidimicrobiaceae*, *Nocardioideae* and *Vicinamibacter* though the significance was small with  $P = 0.0472$ . Furthermore, though they showed a larger association with diseased samples, they were still seen to be present in healthy samples. On the contrary, healthy rhizosphere samples contained the remaining 98 differentially abundant bacteria, 39 of which were exclusively associated with healthy samples including species belonging to *Luteolibacter*, *Porphyrobacter*, *Gemmata* and *Paenibacillus*. LEfSe PICRUST analysis of functional biomarkers identified 20 orthologous proteins, which were all associated with parkland diseased samples. The proteins identified seemed of little impact on plant health aside from the accessory colonization factor AcfA. The Type II general secretion module was the only feature to be identified as differentially abundant and was associated with healthy rhizosphere samples. Likewise, the only pathway that was identified was the *Vibrio cholerae* biofilm formation pathway, which was also significantly associated with healthy rhizosphere samples.

#### 6.2.4. PERMANOVA results

To look for statistical significance between each of the groups, permutational multivariate ANOVA (PERMANOVA) was performed for pairwise comparisons (Table 34). The largest significance was seen between woodland and parkland samples with a P value of 0.001 being recorded, for the full site data sets as well with further qualifiers of diseased and healthy. As previously indicated, woodland healthy pairwise comparison to woodland diseased samples was non-significant ( $P = 0.547$ ). However, the parkland healthy pairwise comparisons to parkland diseased samples showed a small significance value ( $P = 0.045$ ). The Pseudo-F values also provided information with site samples separated on health status having the smallest effect. The largest effect was between woodland healthy and parkland diseased, with both woodland samples showing a smaller effect size than parkland healthy, indicating a similar dispersal.

**Table 34:** The PERMANOVA results for each pairwise comparison from the full site datasets, including the Pseudo-F and q values. The number of permutations was set at 999 and significant P-values (< 0.05) they are indicated by \*.

Pair 1	Pair 2	pseudo-F	P-value	q-value
Healthy	Diseased	0.350	0.884	0.884
Parkland	Woodland	20.796	0.001*	0.001
Woodland Diseased	Woodland Healthy	0.896	0.547	0.547
	Parkland Diseased	17.87	0.001*	0.0015
	Parkland Healthy	16.525	0.001*	0.0015
Woodland Healthy	Parkland Diseased	27.387	0.001*	0.0015
	Parkland Healthy	25.698	0.001*	0.0015
Parkland Diseased	Parkland Healthy	2.346	0.045*	0.054

#### 6.2.5. Parkland paired samples

While the site-level results are interesting, it is difficult to draw meaningful conclusions from the soil over such a large area due to the known heterologous nature of soil over large spatial differences. This study was designed to consider both site-level differences and see if these differences are present in paired trees, where spatial variation is minimalised. As such re-analysis of the paired trees using the same SLR 16S rRNA gene data was performed to understand individual differences.

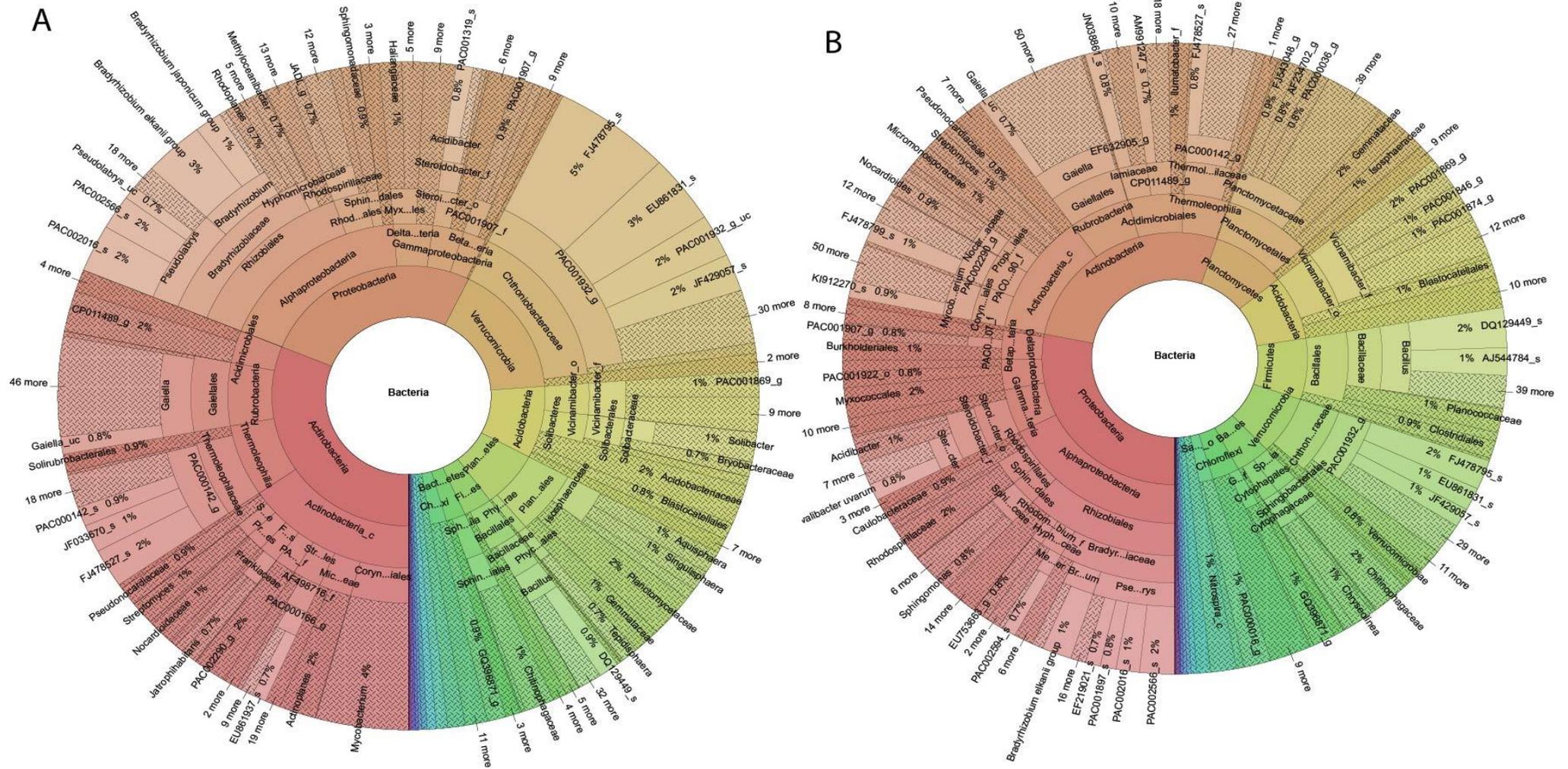
#### *Samples parkland diseased H1 and parkland healthy H2*

The numbers from the composition of the microbiome show divisions between the paired trees in which parkland healthy tree H2 supports a more diverse range of bacteria, while parkland diseased tree H1 shows less diversity and a higher abundance of specific groups at a range of taxonomic ranks, this is demonstrated through the unique taxa numbers in Table

35. The Krona plots in show the vast difference in composition between samples, with the distribution of the top three phyla being different between the two samples (Figure 60).

**Table 35:** The number of shared and unique taxa identified at each taxonomic rank for paired parkland samples parkland disease H1 and parkland healthy H2. Averaged taxonomic compositions of the microbiome taxonomic profiles were used with the estimated taxonomic composition of taxa set to 1.0 %.

Taxonomic Rank	Shared taxa	Unique to parkland diseased (H1)	Unique to parkland healthy (H2)
Phylum	9	0	1
Class	14	2	3
Order	19	6	4
Family	16	8	6
Genus	9	6	5
Species	3	7	3



**Figure 60:** Krona plots of the OTU identifications from the EzBiocloud Pipeline. A = parkland diseased sample H1 and B = parkland healthy sample H2. Samples are organised by abundance with red samples at the bottom left of the plot being the most abundant, going clockwise showing a decrease in abundance, coloured from orange-yellow-green-blue. Links to the interactive plots can be found in Suppl. Table S5.



**Figure 61:** The alpha and beta diversity stats generated from the parkland disease H1 and parkland healthy H2 microbiome taxonomic profiles. Both The Chao1 and NPS Shannon demonstrate that the root regardless of location is the least diverse rich microbiome, with diseased samples then showing a reduced richness and diversity compared to health samples. This is also shown in the PCoA with samples type being the 1<sup>st</sup> PC and most important feature, with health status being the 2<sup>nd</sup> PC with a reduced impact.

The diversity statistics reveal new trends in the paired datasets, with exorhizosphere samples showing higher species richness by Chao1 than endorhizosphere samples, though not at significant levels (Figure 61). Likewise, healthy samples show the slightly higher but non-significant richness and non-parametric Shannon diversity. The beta diversity shows that a difference between paired exorhizosphere samples is observable, as is a difference between endorhizosphere samples, but the largest difference is between the principal coordinate of soil and root. The PERMANOVA results (Table 36) support this observation as the pseudo-F value indicated the smallest effect observable is through the soil to soil, then soil to root from

the same tree, then soil to root from the paired tree. However, none of the P values were significant, indicating a conserved microbiome dispersal in all samples.

**Table 36:** PERMANOVA results for paired samples parkland diseased H1 and parkland healthy H2. The test was performed with 999 permutations and significance was identified if  $P < 0.05$ .

Pair 1	Pair 2	pseudo-F	P-value	q-value
H1	H2	2.050	0.096	0.096
H1 Soil	H1 Root	6.046	0.101	0.1212
	H2 Soil	3.880	0.101	0.1212
	H2 Root	9.456	0.101	0.1212
H1 Root	H2 Soil	6.405	0.091	0.1212
	H2 Root	3.442	0.333	0.333
H2 Soil	H2 Root	6.434	0.101	0.1212

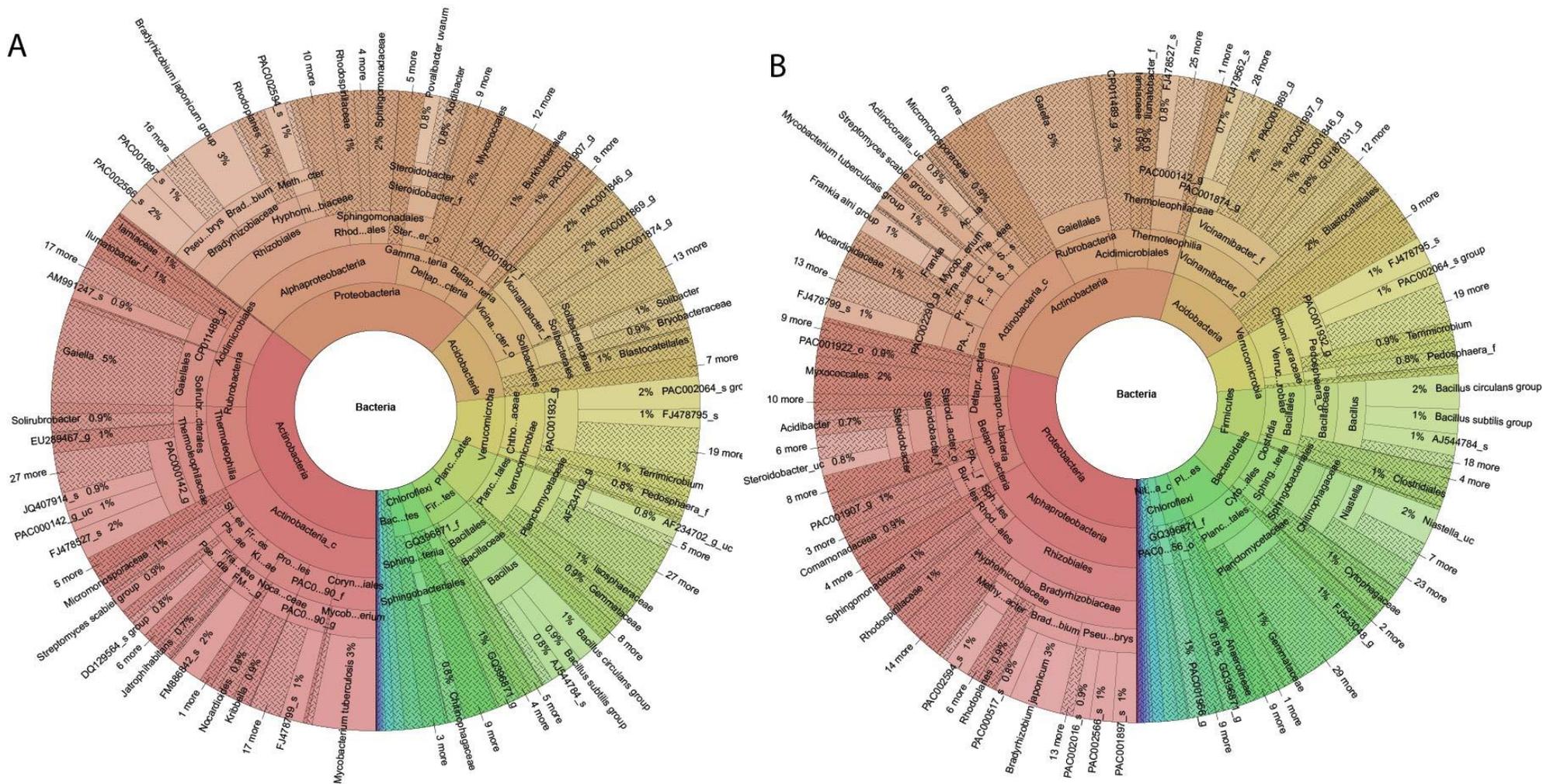
However, differences between the samples could still be identified and as such 453 differentially abundant taxa were identified by LEfSE analysis. Of these identifications, 242 were at the species level and 142 of those differentially abundant species were significantly associated with rhizosphere samples of healthy trees. More importantly, unlike in the grouped LEfSE analysis, PGPB can be identified in each pairing, but their roles appear to differ somewhat. For example, *Chrysolinea*, *Gaiella* and large amounts of *Planctomycetaceae* appeared associated with healthy rhizosphere sample H2. Meanwhile, the diseased rhizosphere showed more associations with *Solibacter*, *Mycobacterium*, and *Streptomyces*. LEfSE analysis also identified 104 functional biomarkers, of these 23 were associated with parkland diseased H1 including MFS transporter, DHA1 family, multidrug resistance protein, carboxylesterase 1 and beta-lactamase class C, all of which play important roles in antibiotic resistance. Parkland healthy H2 showed a larger number of associated proteins, with those of interest including the beta-lactamase class C, Enterobactin, and a type IV secretion subsystem protein.

*Samples parkland diseased H3 and parkland healthy H4*

Samples H3 and H4 represent interesting, paired trees in which a large proportion of the microbiome composition above 1 % is consistent between trees, while a small but equal distribution of tree-specific taxa can be observed at each phylogenetic level as shown in Table 37. The corona plot visually represents the high level of similarity seen between parkland diseased tree H3 and parkland healthy tree H4 though the most abundant taxa are different they are proportionally the same between samples (Figure 62).

**Table 37:** The number of shared and unique taxa identified at each taxonomic rank for paired parkland samples parkland diseased tree H3 and parkland healthy tree H4. Averaged taxonomic compositions of the microbiome taxonomic profiles were used with the estimated taxonomic composition of taxa set to 1.0 %.

Taxonomic Rank	Shared taxa	Unique to parkland diseased (H3)	Unique to parkland healthy (H4)
Phylum	9	0	1
Class	15	2	3
Order	26	4	4
Family	21	5	5
Genus	11	3	4
Species	2	5	4



**Figure 62:** Krona plots of the OTU identifications from the EzBiocloud Pipeline. A = parkland diseased sample H3 and B = parkland healthy sample H4. Samples are organised by abundance with red samples on the bottom left of the plot being the most abundant, going clockwise showing a decrease in abundance. Links to the interactive plots can be found in Suppl. Table S5.



**Figure 63:** The alpha and beta diversity stats generated from parkland disease tree H3 and parkland healthy tree H4 microbiome taxonomic profiles. Both The Chao1 and NPS Shannon demonstrate that the root regardless of location is the least diverse rich microbiome, with diseased samples then showing a reduced richness and diversity compared to health samples. This is also shown in the PCoA with samples type being the 1<sup>st</sup> PC and most important feature, with health status being the 2<sup>nd</sup> PC with a reduced impact.

The diversity plots show similar information to those of parkland diseased H1 and parkland healthy tH2 in Figure 61 with both soil samples show the highest species richness and diversity observations, though the roots of diseased parkland tree H3 show a higher average richness and diversity than those of parkland healthy H4 (Figure 63). Again, no significant difference can be identified through the alpha diversity statistics, indicating similarity in all samples. The beta diversity plot again shows that the biggest observable difference is between the root and soil components of the samples. However, this time there is no clear separation between the diseased and healthy samples as there was in the first pairing. The soil samples show

almost identical coordinates, and the root samples show a clustering due to a lower first principal coordinate but otherwise no relation to their respective root sample or the exorhizosphere soil sample. The PERMANOVA results (Table 38) shows that soil-to-soil samples and then the root-to-root samples are the most similarly related, followed by a closer relation of the root sample to the related soil sample than that of the paired tree. Again, no significance was identified through pairwise comparisons.

**Table 38:** PERMANOVA results for paired parkland disease samples H3 and the parkland healthy sample H4. The test was performed with 999 permutations and significance was identified if  $P < 0.05$ .

Pair 1	Pair 2	pseudo-F	P-value	q-value
H3	H4	0.636	0.636	0.606
H3 Soil	H3 Root	4.828	0.101	0.152
	H4 Soil	0.404	0.614	0.614
	H4 Root	6.700	0.101	0.152
H3 Root	H4 Soil	6.009	0.091	0.1512
	H4 Root	1.938	0.333	0.3996
H4 Soil	H4 Root	7.782	0.101	0.1512

The close relation of the samples shown through the diversity analysis was also revealed through the taxonomic biomarker comparison which again mirrored the small variation between trees. One hundred and seven differentially abundant taxa were identified, of these 62 were identified as differentially abundant species, only 14 were associated with the healthy sample H4, while the remaining 48 were more abundant in the rhizosphere of parkland diseased H3. Species associated with H4 belonged to *Flavitalea*, *Gemmatimonas* and uncultured members of the *Planctomycetaceae*. H3 however, showed stronger associations with *Streptomyces*, *Pseudocardia* and *Gaiella*. The Functional biomarker discovery showed 37 specific proteins assigned to each tree via LEfSe analysis. Twenty-five proteins were associated with the H3 rhizosphere sample with the remaining 12 showing higher relative abundance in parkland healthy tree H4. None of the identified proteins in parkland diseased

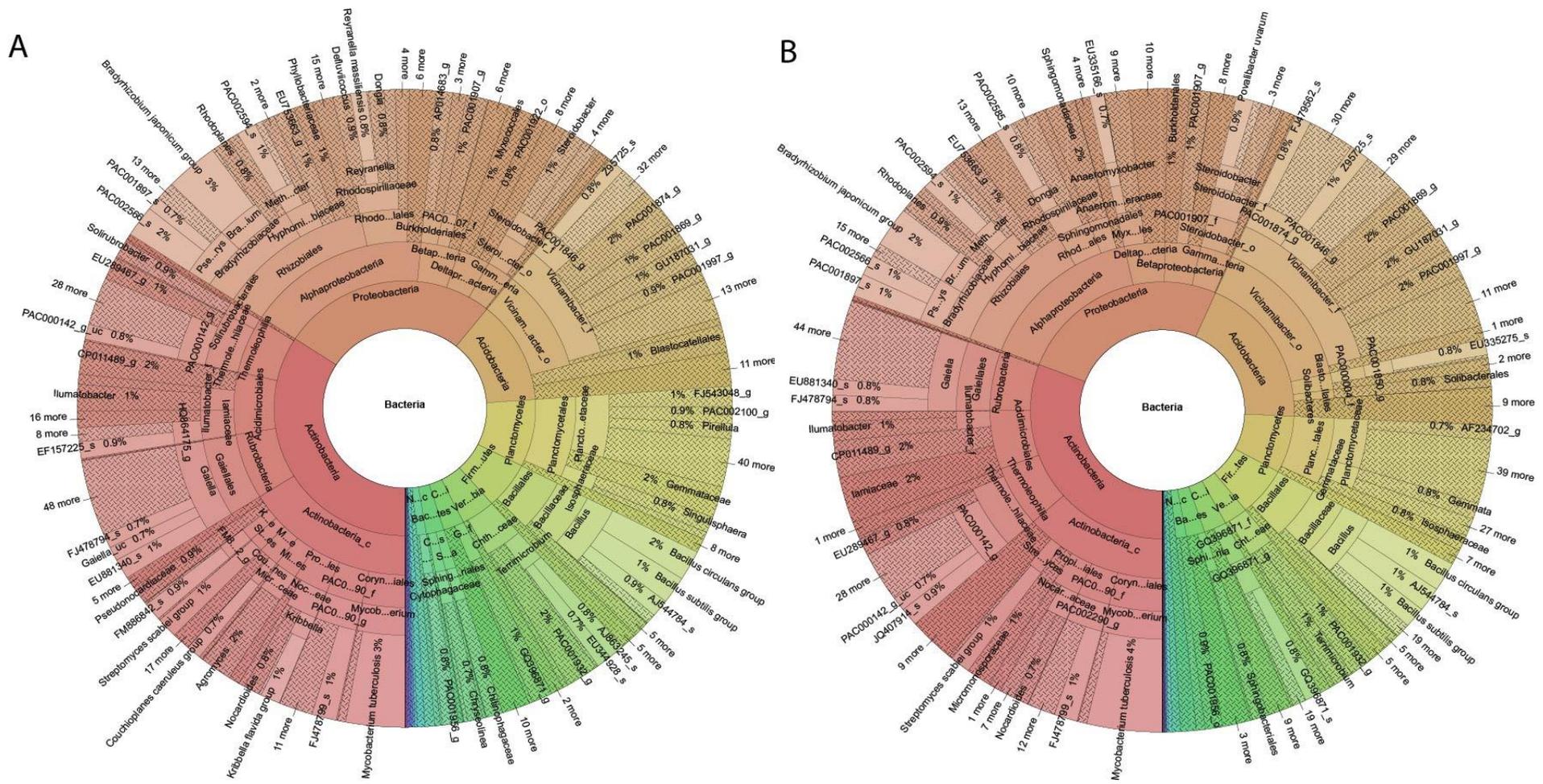
H3 showed a clear role in plant health, but parkland healthy H4 had specific associations with a HlyD family secretion protein which belongs to the Type I Secretion System, and a membrane fusion protein, multidrug efflux system.

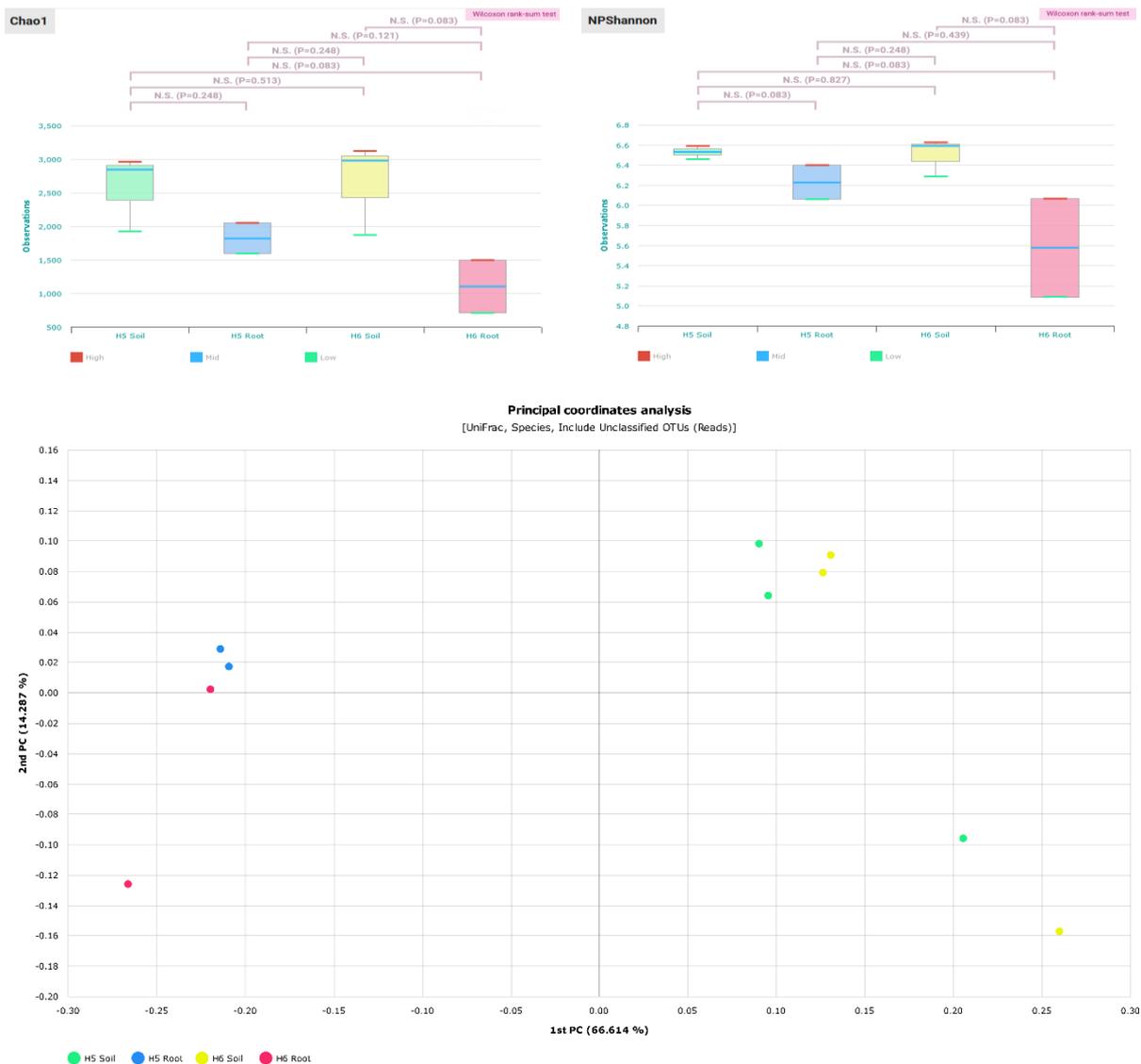
*Samples Parkland diseased H5 and parkland healthy H6*

Samples H5 and H6 display the most spatial variation of all the paired samples. Despite this, they demonstrate a remarkably similar microbiome composition as shown in Table 39. However, H5 does appear to support a larger number of unique taxa at each rank excluding species, where most of the identifications fall under the 1.0 % estimated taxonomic composition cut-off value. Again parkland diseased tree H5 and parkland healthy tree H6 demonstrates that there is very few differentially abundant taxa between spatially related healthy and AOD symptomatic samples (Figure 64).

**Table 39:** The number of shared and unique taxa identified at each taxonomic rank for paired parkland samples parkland diseased H5 and parkland healthy H6. Averaged taxonomic compositions of the microbiome taxonomic profiles were used with the estimated taxonomic composition of taxa set to 1.0 %.

Taxonomic Rank	Shared taxa	Unique to parkland diseased (H5)	Unique to parkland healthy (H6)
Phylum	9	0	0
Class	16	1	0
Order	19	6	1
Family	21	6	2
Genus	14	6	3
Species	2	1	3





**Figure 65:** The alpha and beta diversity stats generated from the parkland diseased tree H5 and parkland healthy H6 microbiome taxonomic profiles. Both The Chao1 and NPS Shannon demonstrate that the root regardless of location is the least diverse rich microbiome, with diseased samples then showing a reduced richness and diversity compared to health samples. This is also shown in the PCoA with samples type being the 1<sup>st</sup> PC and most important feature, with health status being the 2<sup>nd</sup> PC with a reduced impact.

The previous trend is also seen in the paired comparison of parkland diseased tree H5 and parkland healthy tree H6 samples, with exorhizosphere soil showing a closer composition to one another, than to their respective endorhizosphere samples, with the same being true for the endorhizosphere samples (Figure 65). Soil is richer and more diverse, after which diseased roots appear to be more species-rich and diverse. Again, no significance was identified, and the beta diversity showed similar clustering effects, though H5 diseased root samples appear highly similar, while H6 healthy roots show more spread. The PERMANOVA results in Table

40 support all the previous statements, with the whole H5 to H6 and then soil to soil showing the smallest pseudo-F values, though interestingly the H6 root composition showed a slightly closer relation to the H5 soil, than H6 soil.

**Table 40:** PERMANOVA results for paired parkland diseased tree H5 and parkland healthy tree H6. The test was performed with 999 permutations and significance was identified if  $P < 0.05$ .

Pair 1	Pair 2	pseudo-F	P-value	q-value
H5	H6	0.362	0.726	0.726
H5 Soil	H5 Root	7.286	0.101	0.152
	H6 Soil	0.632	0.53	0.53
	H6 Root	6.937	0.101	0.152
H5 Root	H6 Soil	7.377	0.091	0.152
	H6 Root	1.272	0.333	0.3996
H6 Soil	H6 Root	6.991	0.101	0.152

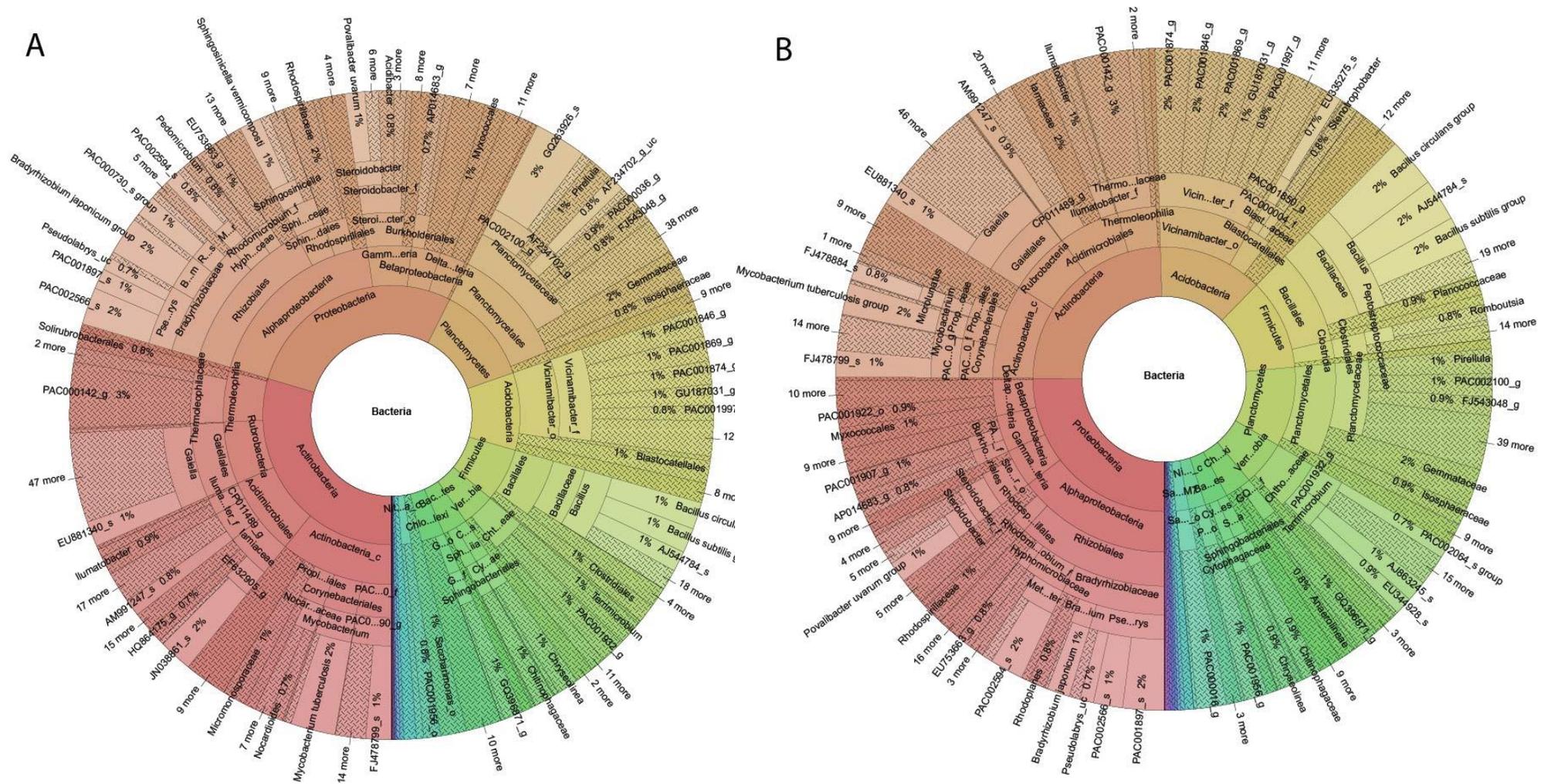
The taxonomic biomarker comparison showed 113 differentially abundant taxa being identified between diseased tree H5 and healthy tree H6. Fifty seven of these differentially abundant taxa were identified at the species level, of which only seven were associated with healthy tree H6 while the remaining 50 were associated with diseased tree H5. Healthy tree H6 species were reported as *Desulforhopalus* sp., *Streptomyces canus* and *Actinoallomurus coprocola*. While diseased tree H5 showed larger associations with *Defluviicoccus*, *Streptomyces* and *Mycobacterium*. The functional biomarkers revealed 40 differentially abundant proteins, of which 17 were associated with H6 and the remaining 23 with H5. While proteins associated with H6 showed normal cell regulatory roles, an MFS transporter, ACDE family and multidrug resistance protein were also seen as significantly associated. Meanwhile, H5 showed upregulated marker aligning to endoglucanase, a cellulase family member, and chitinase both of which play degradation roles.

*Samples parkland diseased H7 and parkland healthy H8*

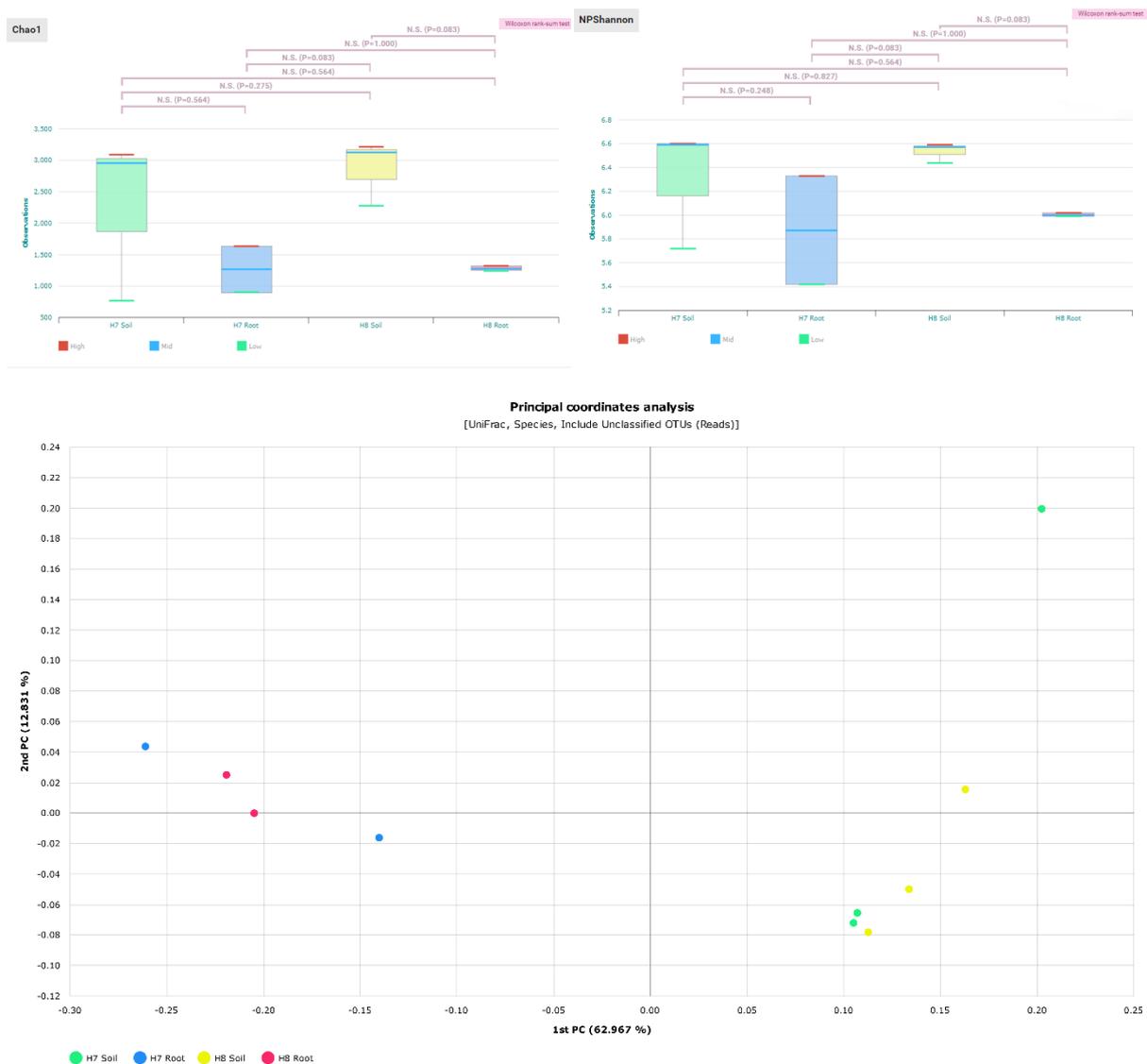
Parkland diseased H7 and parkland healthy H8 samples were closely related which appear to be similar in their taxonomic microbiome composition. From phylum to family level, the two samples hardly deviate at all in taxa which make up over 1 % of the microbiome (Table 41). Even at genus and species level, both trees appear to support very similar microbiomes, with only a small amount of equally spread deviation between the two samples. The krona plots show that although the samples are similar the dominant phyla are different in each sample, indicating a larger difference in the composition of samples (Figure 66).

**Table 41:** The number of shared and unique taxa identified at each taxonomic rank for paired parkland samples diseased tree H7 and healthy tree H8. Averaged taxonomic compositions of the microbiome taxonomic profiles were used with the estimated taxonomic composition of taxa set to 1.0 %.

Taxonomic Rank	Shared taxa	Unique to parkland diseased (H7)	Unique to parkland healthy (H8)
Phylum	9	0	0
Class	18	0	2
Order	20	1	1
Family	21	1	1
Genus	13	3	6
Species	4	4	4



**Figure 66:** Krona plots of the OTU identifications from the EzBiocloud Pipeline. A = parkland diseased sample H7 and B = parkland healthy sample H8. Samples are organised by abundance with red samples on the bottom left of the plot being the most abundant, going clockwise showing a decrease in abundance. Links to the interactive plots can be found in Suppl.table. S5.



**Figure 67:** The alpha and beta diversity stats generated from the parkland diseased tree H7 and the parkland healthy tree H8 microbiome taxonomic profiles. Both The Chao1 and NPS Shannon demonstrate that the root regardless of location is the least diverse rich microbiome, with diseased samples then showing a reduced richness and diversity compared to health samples. This is also shown in the PCoA with samples type being the 1<sup>st</sup> PC and most important feature, with health status being the 2<sup>nd</sup> PC with a reduced impact.

The alpha diversity results show that healthy soil shows the highest species richness via Chao1, with diseased soil showing a similar but average number of observations, though the lowest number of observations in the diseased samples was 750 observations lower (Figure 67). Both the root samples showed similar observations though the diseased showed a slightly higher average and larger spread. This is supported by NPS Shannon diversity, which showed soil samples as more diverse, with parkland diseased tree H7s soil showing a higher average.

The root samples showed parkland healthy tree H8 had a slightly higher average and lower deviation between samples, while parkland diseased tree H7 showed large variability between the number of observations, no significance was identified between any of the samples. This is echoed in the beta UniFrac PCoA plot which shows a large distance between H7 diseased root samples. The same effect is seen in the other beta diversity plots in which the most distinct difference is between the endorhizosphere root samples against the exorhizosphere soil samples, independent of the tree of origin. H8 healthy soil samples show lower deviation than parkland diseased H7, though both samples have a high level of variability. The PERMANOVA results in Table 42 support these conclusions, indicating soil samples, independent of the tree, have the lowest pseudo-F value, followed by root-to-root comparisons. Like the PERMANOVA results for parkland diseased tree H5 and parkland healthy tree H6 in Table 8, H8 endorhizosphere roots show a more similar composition to the parkland diseased tree H7 exorhizosphere soil samples, with the largest effect in the dataset being seen between the parkland healthy H8 soil and roots.

**Table 42:** PERMANOVA results for paired samples from parkland diseased tree H7 and parkland healthy tree H8. The test was performed with 999 permutations and significance was identified if  $P < 0.05$ .

Pair 1	Pair 2	pseudo-F	P-value	q-value
H7	H8	0.561	0.572	0.572
H7 Soil	H7 Root	4.635	0.101	0.152
	H8 Soil	0.834	0.53	0.53
	H8 Root	6.534	0.101	0.152
H7 Root	H8 Soil	7.758	0.091	0.152
	H8 Root	1.346	0.333	0.3996
H8 Soil	H8 Root	12.499	0.101	0.152

Only 78 taxonomic biomarkers were discovered using LeFSE analysis, of these only 42 were assigned at the species level, with 21 being significantly associated with parkland diseased tree H7 and 21 being associated with parkland healthy tree H8. H8 healthy samples showed the largest associations with species belonging to the *Gemmataceae*, *Ilumatobacter* and *Anaerolinaceae*, H7 diseased samples meanwhile showed associations with *Kribbella*, *Variibacter* and *Streptomyces*. The functional biomarker results were limited with 16 proteins identified, again these were evenly distributed. Parkland diseased tree H7 showed interesting proteins such as streptomycin 3"-adenylyltransferase which mediates antibiotic resistance (Prabhu, Vidhyavathi and Jeyakanthan, 2017) and 2'-hydroxybiphenyl-2-sulfinate desulfinate which is responsible for the hydrolysis release of sulphities which can be pesticides and plant-growth regulators (Sandler and Karo, 1992). Parkland healthy tree H8 however, presented no proteins with clear environmental or plant health roles, though a minor pathway for porphyrin and chlorophyll metabolism was more significantly associated with it.

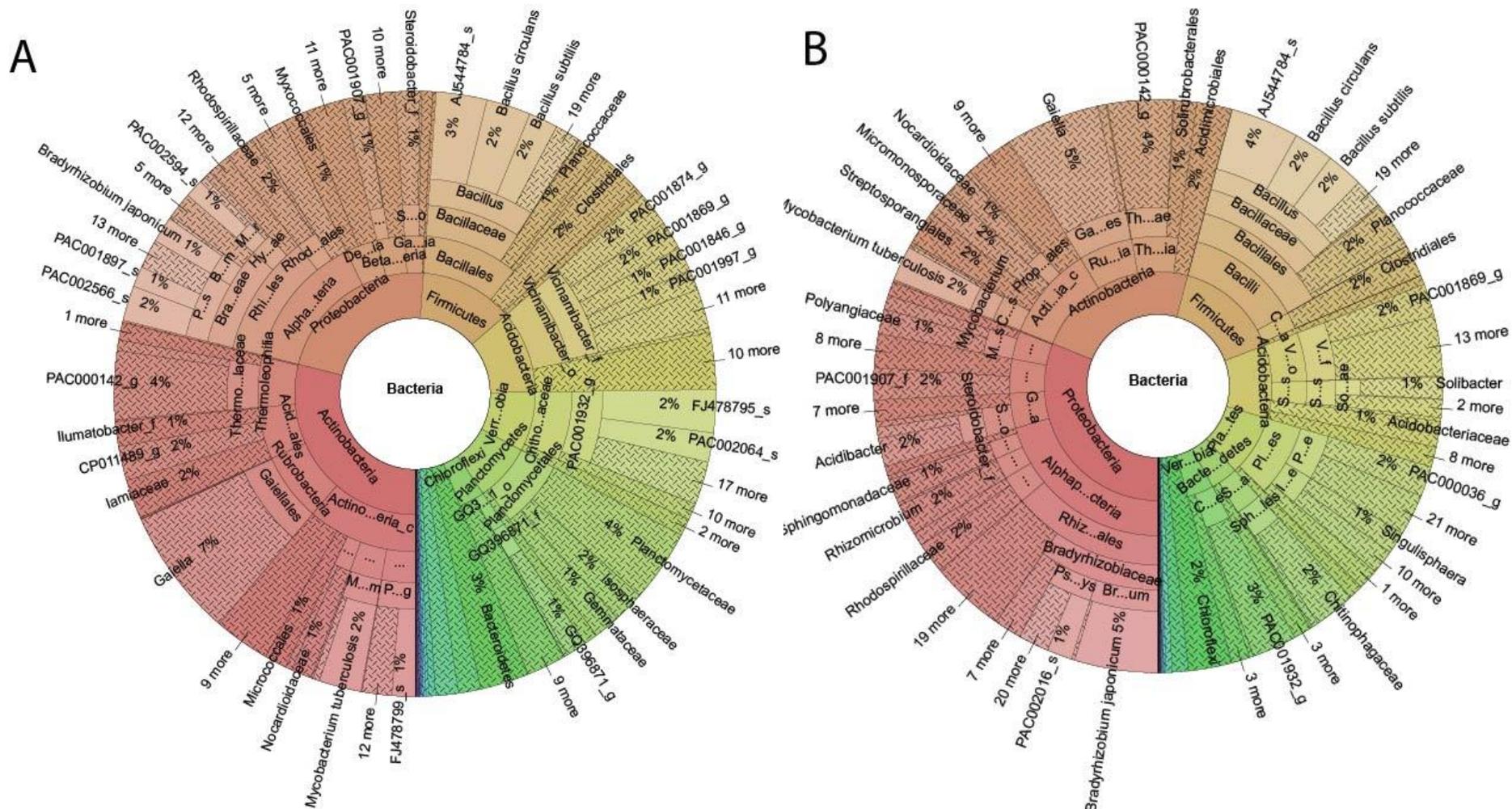
*Samples parkland diseased H9 and parkland healthy H10*

The final pairing offers a unique perspective based on their location on the other side of the parkland area of Hatchlands separated by the woodland which can be seen in the original sampling map in chapter one (Figure 23). As Table 43 shows, they do differ considerably more than some of the other pairings, with parkland healthy H10 supporting more unique taxa above the 1.0 % ETC cut-off than parkland diseased H9 apart from at the species level. This is also well supported by the krona plots in which diseased tree H9 supports Actinobacteria as the largest phyla, while the healthy tree H10 supports Proteobacteria, though otherwise, the order of abundance is very similar between the two samples (Figure 68).

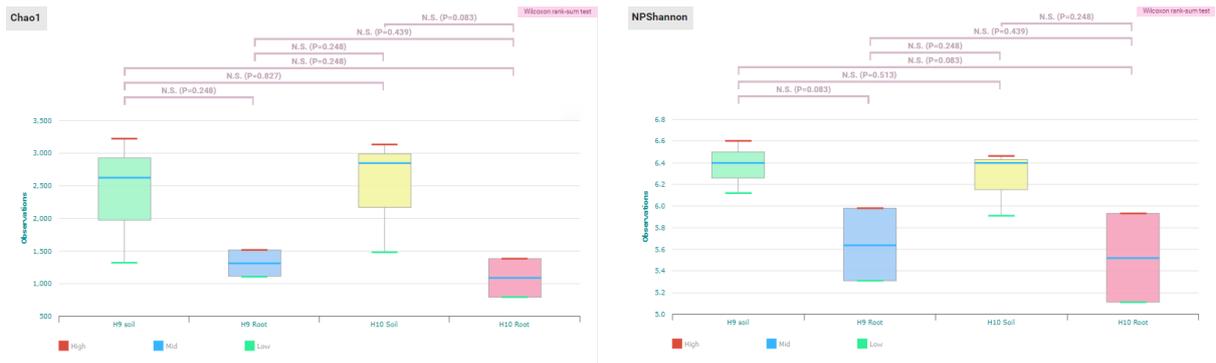
**Table 43:** The number of shared and unique taxa identified at each taxonomic rank for paired samples parkland diseased H9 and parkland healthy H10. Averaged taxonomic compositions of the microbiome taxonomic profiles were used with the estimated taxonomic composition of taxa set to 1.0 %.

Taxonomic Rank	Shared taxa	Unique to parkland diseased (H9)	Unique to parkland healthy (H10)
Phylum	9	0	0
Class	15	2	4

Order	19	2	7
Family	17	6	9
Genus	7	5	8
Species	3	5	1

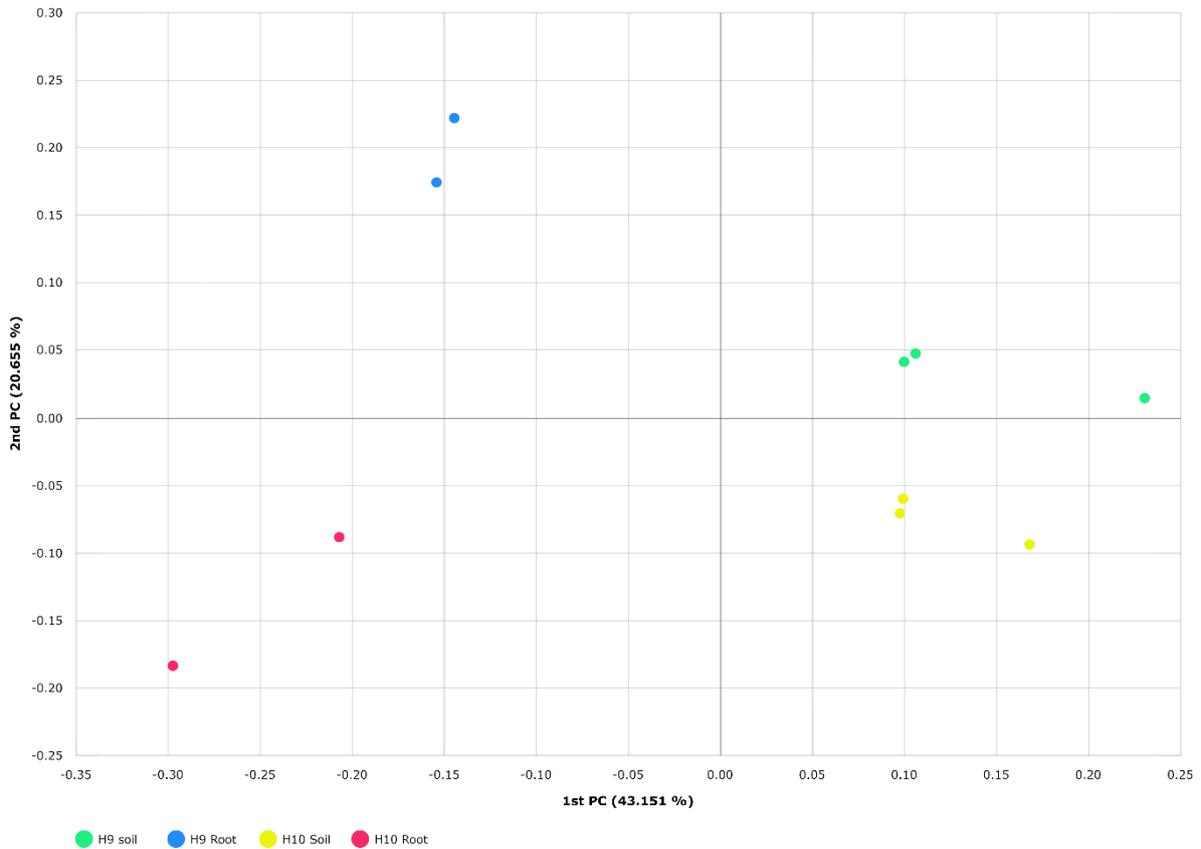


**Figure 68:** Krona plots of the OTU identifications from the EzBiocloud Pipeline. A = parkland diseased sample H9 and B = parkland healthy sample H10. Samples are organised by abundance with red samples on the bottom left of the plot being the most abundant, going clockwise showing a decrease in abundance. Links to the interactive plots can be found in Suppl. Table S5.



### Principal coordinates analysis

[UniFrac, Species, Include Unclassified OTUs (Reads)]



**Figure 69:** The alpha and beta diversity stats generated from the parkland diseased tree H9, and parkland healthy tree H10 microbiome taxonomic profiles.

The diversity statistics show the high level of separation between samples. For both alpha diversity plots, soil shows the highest species richness and diversity with healthy soil showing higher richness, but slightly lower diversity. The root samples were less rich and diverse with H9 diseased roots showing more observations in alpha diversity than H10 healthy roots (Figure 69). However, none of the identified differences was shown to be statistically significant, implying a close dispersal between the paired samples. The beta diversity plot demonstrates the same trend as parkland diseased sample H1 and parkland healthy sample

H2, with the clearest separation being between roots and soil, but with soil samples forming two clear clusters based on their tree of origin. The difference between the root samples was large, highlighting a closer relation to their respective endorhizosphere samples than to the samples from the paired tree. The PERMANOVA results in Table 44 support the beta plot in Figure 69 with a nearly statistically significant difference between the samples. When breaking them down, the clearest relations are the same with soil to soil and then root to root showing the smallest pseudo-F values, followed by soil to their respective root sample and finally root to the soil of the paired tree.

**Table 44:** PERMANOVA results for paired samples parkland diseased H9 and parkland healthy H10. The test was performed with 999 permutations and significance was identified if  $P < 0.05$ .

Pair 1	Pair 2	pseudo-F	P-value	q-value
H9	H10	1.800	0.055	0.055
H9 soil	H9 Root	5.078	0.101	0.1212
	H10 Soil	2.821	0.101	0.1212
	H10 Root	5.768	0.101	0.1212
H9 Root	H10 Soil	6.984	0.091	0.1212
	H10 Root	3.052	0.333	0.333
H10 Soil	H10 Root	5.839	0.101	0.1212

The differences identified in Figure 69 and Table 44 were reflected in the taxonomic biomarker analysis which showed 386 differentially abundant taxa. Of these, 219 were identified at the species level, and of these 122 were predominantly associated with parkland diseased tree H9 while the remaining 97 were associated with parkland healthy tree H10. H9 showed high associations with *Gaiella*, *Anaerolinea*, *Gemmata* and *Mycobacterium*, while H10 showed associations with *Solibacter*, *Koribacter*, *Caulobacter*, *Saccharimonas* and *Rhizomicrobium*. The LEfSe functional biomarker analysis also supported this by the

identification of 77 significantly differentially abundant pathways, of which 59 were associated with diseased tree H9 with the remaining 18 being associated with healthy tree H10. Some of the interesting proteins identified included Type III secretion inner rod protein HrpB2, Type IV pilus assembly protein PilC and a cysteine desulfuration protein SufE, all with association to parkland diseased H9. Meanwhile, parkland healthy tree H10 had a HlyD family secretion protein (of the T1SS), a methyl-accepting chemotaxis protein and an MFS transporter, DHA1 family, multidrug resistance protein.

### 6.3. Discussion:

The work presented in this chapter demonstrates a two-pronged approach to assessing rhizosphere soil microbiome composition. The first approach is more standardised where samples are viewed in groups to identify differences between variables at a site level. These results from this section reveal an interesting trend, in which woodland and parkland samples show a high level of variation between their microbiome composition. This is to be expected as different environments are thought to support different bacteria. For example, the woodland environment was full of different tree species and other floor-based plants with a layer of debris on the topsoil. This was being actively degraded, and as such the soil was generally dark brown in colour, loosely packed and full of small decaying plant matter. These features are indicators of good soil health, which is generally related to diverse and rich microbiomes (Sokolov *et al.*, 2020). However, this was not reflected in the diversity statistics for the site comparisons where parkland rhizosphere samples were richer and more diverse. This could correlate with the humic substances in woodland soil, which can interfere with sequencing (Sharma *et al.*, 2012; Sutton and Sposito, 2005).

When separating the samples into healthy and diseased woodland and parkland, the variability and lack of trends in the woodland samples became apparent. This is most likely due to the higher level of variability in the surrounding environment, because of plant matter and other plant interactions. As such, the more sterile environment of the parkland appeared to offer clearer samples to work with. Parkland samples showed a clear separation based on health status in which the healthy rhizosphere supported richer and more diverse samples, with a significant difference in the phylogenetic diversity observed between them. Moreover, the difference was identified by PERMANOVA analysis, though with a small P value of 0.045. Why a difference would exist here and not at the woodland site is a very interesting point. As

noted, the woodland samples originated from apparently healthy soil, while the parkland soil samples were compact, light brown and sandy, containing little decaying matter. Previous studies at the same site have also shown that the parkland soil is more acidic and lower in nutrient content (Pinho *et al.*, 2020). Therefore, it is logical for the rhizosphere effect previously discussed to have larger effects on the microbiome composition of the soil. This is significant when considering that AOD has been seen to cause severe disruption to tree exudate profiles (Wiley, 2020). The poor-quality soil may be unable to support rich and diverse microbiomes alone. Instead, the rhizosphere effect plays an essential role in supporting the microbiome richness and diversity, and when disrupted notable effects on the microbiome are then seen. This leads to a primary conclusion that the disease status of oak plays a role in the rhizosphere composition at the site level, depending on where samples originate from.

However, site-level comparisons are of questionable relevance, considering the known heterogeneous nature of the soil, in which small spatial effects have huge effects on microbiome compositions, which only increase over larger distances. This study was designed to also observe the effect in paired samples where spatial variation was minimalised. Due to the cost and work involved, this analysis was limited to the parkland samples where a difference was identified. To add a further layer of analysis, both endorhizosphere (root) and exorhizosphere (soil) samples were sequenced to determine how the effect varied over the whole of the rhizosphere. However, when removing spatial separation, the effects of diseased and healthy are also diminished. In each pairing, no significance between any sample type was found for the alpha diversity or PERMANOVA tests performed. However, a trend was consistently identified in each pairing in which the soil samples were the most diverse and rich in terms of alpha diversity, but they were also the most related and this was visible in the beta diversity plots. This shows that by reducing the spatial variation, samples showed a clear relation. The most interesting outcome of this was that despite the similarity between soil sample populations, the root samples showed a consistent difference, from each other and the soil. It is already known that the majority of root colonisers originate from the soil, most specifically the rhizosphere, and that root exudates affect endosphere colonisers (Compant *et al.*, 2021). Therefore, two explanations can be inferred for the root sample differences. Either the original composition of the endophytic composition of the root was different, or

more likely as it is proven that endophytic colonisers alter with age, that the bacteria recruited from the exorhizosphere into the endorhizosphere alter based on the tree. This could be affected by AOD, leading to the altered endosphere rhizosphere microbiome, though further work would be required to reveal this connection.

One consistently interesting result shown through each of the paired sample analyses is the differentially abundant species analysis. Though conclusive statements are hard to make from these results, trees are supporting different bacteria, with potentially significant roles relating to how they respond to the environment. For example, all diseased rhizosphere samples excluding parkland diseased tree H9 showed higher correlation to *Streptomyces* species, a genus renowned for producing numerous antibiotic substances (Amaresan *et al.*, 2020). Hypothetically, the *Streptomyces* strains could be specifically recruited by trees for their antagonistic effect on deleterious microorganisms, which would be especially beneficial for weakened trees which are susceptible to infection. A generalisation is harder to make for the healthy samples as the more abundant species were more unique per tree than those in diseased, however, genera like *Gemmatimonas*, *Gaiella* and *Planctomycetaceae* were seen as significantly associated with many healthy root samples. While little is known about *Gemmatimonas*, *Gaiella* and *Planctomycetaceae*, they are potentially interesting genera. *Gaiella* for example has been shown to metabolise complex chemicals in soil, such as herbicides and has been suggested for bioremediation because of these properties (Pertile *et al.*, 2021). Likewise, the *Planctomycetaceae* are unique anaerobic nitrifiers, called ammox bacteria (Buckley *et al.*, 2006). This could be of significance as nitrogen-cycling bacteria are known to be associated with healthy oak, and the parkland soil at Hatchlands was previously shown to be compact, which increases the anaerobic potential of soil (Pinho *et al.*, 2020; Batey and McKenzie, 2006).

Overall, differences can be seen between the woodland and parkland samples, as well as the parkland diseased and parkland healthy samples. However, using the paired model analysis that the study was designed for, the differences between healthy and diseased trees cannot be observed in paired samples. However, despite the similarity of the exorhizosphere samples, endorhizosphere samples appear different from the soil and each other, highlighting a potentially interesting avenue for further investigation. Furthermore, the differentially

abundant bacteria present in each sample highlight interesting differences, which despite the lack of statistical significance, might still play important roles in plant health.

# Chapter 7. General Discussion

The overall aims of this thesis were to identify if rhizosphere soil could function as a reservoir for the bacteria associated with AOD lesions; isolate and taxonomically classify potentially interesting novel species of bacteria associated with the oak rhizosphere and assess the community composition of the oak rhizosphere to see if structural differences exist in correlation with the health status of the tree. Each chapter in this body of work represents a contribution to answering these questions.

While the direct impact of Chapter three on the understanding of AOD is very limited, it does give a rational on how to approach sampling the rhizosphere from start to finish. Currently, there is no standardised approach to collecting, storing, and processing soil, generating limitations on the ability to compare separate studies (Barillot *et al.*, 2013). By demonstrating where variability arises in the sampling process, future studies can utilise the most appropriate method, minimalizing post-sampling variation and increasing the ability for comparison to other studies. Sample collection was designed to minimise the effect of spatial variation on the rest of the study. Two different environments over one site were picked, with five pairs of healthy and diseased trees being identified, and four samples per tree collected for each tree. While most studies appreciate the effect of spatial separation in rhizosphere studies, methods to reduce the effect do not appear commonly considered, and the use of paired environmental samples to reduce it was not identified in the literature (Vetterlein *et al.*, 2020). Short-term storage of samples was shown to have little effect on the microorganisms that were recovered, using non-specific methods or the DNA extracted which has previously been suggested (Lauber *et al.*, 2010). However, the way that soil samples are processed in the lab was shown to have a significant effect on the CFU and morphotypes of bacteria isolated as well as the quality and quantity of DNA extracted. While this was optimised for this work based on these samples and with the equipment available, it only highlights a portion of the sample collection process that requires standardisation. Further work would be required, including testing a larger number of soil types, and using further equipment such as blenders and stomachers to find the best tool for substrate disruption. Equally important would be to test these methods on different target groups, such as other families of bacteria, to see if they are broadly applicable or if specific processing methods are required for different study types. Perhaps the most important contribution of the optimisation work is the first application of LAMP for the confirmation of rhizosphere soil as

originating from oak. LAMP has been used for the identification of several different samples including the identification of plant pathogens as a cost-effective and rapid colour change system (Pritchard *et al.*, 2016; Czajkowski *et al.*, 2015). As noted in Chapter three, it is not standard practice to identify where rhizosphere samples originate from, which is of little surprise as visual root identification requires specialist knowledge and other sequencing options are both time-consuming and financially unviable for many studies. The rapid identification method ratifies these concerns while also proving a reliable method to confirm the identity of root samples. The results for the LAMP section also indicate that this should be standard practice for rhizosphere studies of a similar nature to this work. Over a third (37.5 %) of rhizosphere samples collected in this study did not originate from oak, without LAMP identification these samples would have been included in the rest of this work which could have altered and confounded the study outcomes. Instead, it can say with certainty that the HRM identifications of AOD bacteria in rhizosphere soil, the novel isolates identified and the full-length 16S rRNA gene sequencing analysis are all directly related to oak, and any impact found is specific. As such it would be good practice to make this the standardised approach to in vivo environmental studies relating to specific plants.

A large meta-study identified that within plant disease research, bacteria appear as the second largest cause of stress within the biotic stress field (Gimenez, Salinas and Manzano-Agugliaro, 2018). One issue with plant bacterial diseases is that unlike viruses, which almost exclusively require an insect vector, transport and dispersal can be through, rain, wind, soil, seeds/pollen and other biotic factors such as animals, people and insects (Purcell, 2009). Dispersal of pathogens leads to the infection of a new host, completion of their life cycles, potential epidemics as well as increased gene flow, causing the evolution of the pathogen population (West, 2014). As such understanding dispersal is an essential focus for pathologists as it enables early detection and the ability to influence and control the spread of pathogens. Within the field of bacterial plant disease, the gram-negative bacteria *Erwinia*, *Pseudomonas* and *Xanthomonas* are seen as the most destructive, economically important pathogens (Sharma, Gautam and Wadhawan, 2014). These three genera exemplify the wide range of dispersal strategies, with *Erwinia amylovora* being seen to use insect vectors, water, soil and orchard worker clothes (Santander *et al.*, 2020; Choi *et al.*, 2019; West, 2014). *Xanthomonas* is mostly associated with the movement of contaminated plant matter, mainly seeds (An *et*

*al.*, 2019), though a portion of its life cycle is spent outside of hosts in the soil where it is a key member of the rhizosphere microbiome, making up 0.7 – 7 % of the community composition (Bhattacharyya *et al.*, 2018; Xu *et al.*, 2018). Due to the high numbers of *Xanthomonas* in rhizosphere soil, it has been speculated that it may be a reservoir of infection (Zhao, Damicone and Bender, 2002). Meanwhile, *Pseudomonas syringae*, a model plant pathogen, is shown to be a ubiquitous epiphyte, which was originally considered to be its main environmental niche. However, recent work has shown that *P. syringae* is also found in aquatic habitats, such as alpine lakes where it shows its highest number of genetic groups, some of which are unobserved elsewhere. The identification of these reservoirs has shown that *P. syringae* is associated with the water cycle and has ice nucleation activity that can influence snow and rainfall, which leads to its ubiquitous dispersal through the water cycle (Morris, Monteil and Berge, 2013).

Chapter four discussed at the length closely related genera of bacteria to those identified in the AOD lesion microbiome and the probability of where the associated bacteria might be isolated from. The first major contribution of chapter four is the knowledge that all the AOD bacteria are capable of surviving in the soil matrix for extended periods, which differs from the previous conclusions drawn from a similar experiment, in which *Brenneria goodwinii* became viable but non-cultivable upon spiking into the soil environment (Pettifor *et al.*, 2020). From all tested sample types *Rahnella victoriana* appeared the most ubiquitously dispersed environmental bacterium, being easily recovered from soil, leaves and acorns as well as some other niches which were performed in this work but not included (bark swabs and branch cuttings). This was as expected, as *Rahnella* is well-recorded as being a hardy, ecologically diverse bacteria found in numerous environments (Janda and Abbott, 2021). Given that it is seen in a large number of AOD lesion microbiomes, but not all, and that it is seen to have a secondary role in the disease (Doonan *et al.*, 2019), it seems reasonable to say that findings support the hypothesis that *Rahnella* is most likely just playing an opportunistic role due to its wide dispersal. *Gibbsiella quercinecans* shows a similar dispersal with identifications made from soil, leaves and acorns, but it is less frequently recovered. More importantly, it was significantly more associated with acorns than any other niche, potentially indicating that maybe *G. quercinecans* is an oak endophyte, which would explain the high proportion of lesions samples it is recovered from and the primary role it plays in lesion

development (Doonan *et al.*, 2019). While *Lonsdalea britannica* and *B. goodwinii* were only identified in a few samples their sole recovery from acorns strongly implies they are inherited members of the seed endophytic community. Though if true, this does not explain the infrequent isolation of *Lonsdalea* in lesion samples, except through its poor ability to grow from mixed cultures compared to the other members of the AOD lesion microbiome (Brady *et al.*, 2012). The implication of three of the members of the lesion microbiome being endophytic in acorns, suggesting they are inherited members of the lesion microbiome is concerning. Firstly, if endophytes are becoming pathogenic it would suggest trees showing symptoms are already severely damaged, as to not be able to keep their endophytic bacteria from becoming pathogenic. These concerns were also noted with *Brenneria salicis*, which is endophytic in willow, where it also causes disease (Maes *et al.*, 2017). The problems noted are the inability to stop the spread of endophytes in the same way as free-living pathogens and monitoring for presence no longer enables early prediction of disease (Maes *et al.*, 2017).

Pathogens that attack plants through soil have the potential to be treated using biocontrol agents. For example, *Ralstonia solanacearum* the cause of bacterial wilt diseases of tobacco, has been suppressed through the addition of *Bacillus amyloliquefaciens* Z9M (Hu *et al.*, 2021). Other alternatives include the addition of chitin material to soil (Hjort *et al.*, 2014) or the use of microwaves (Nelson, 2015). Likewise, control of pathogens that enter through the aerial section of the plant can be mitigated using different control methods. For example, *Xanthomonas campestris* pv. *pruni* leaf spot infection can be reduced by 58 % complete foliar infection to 22 % by pre-treatment with phage (Jones *et al.*, 2008). Alternative methods also include a range of fungal and bacterial biopesticides, which are now favoured due to the reduced environmental pollution associated with agricultural chemicals (Ritika and Utpal, 2014). Treatment of diseases in the stem of plants can also be performed with injections of antibiotics, although this is both labour-intensive and expensive, and must be frequently repeated to relieve or delay symptoms and as such is not practical for agricultural or forest trees (McManus and Stockwell, 2001). Some tree diseases have been treated with phage such as the canker of *Xanthomonas citri* subsp. *citri* using bacteriophages combined with acibenzolar-S-methyl which is a fungicide, applied with sprays and soil drenching (Ibrahim, Saleh and Al-Saleh, 2017). A few commercially available phage treatments against fire blight exist including AgriPhages and Erwiphage, though their efficiency is hindered by many

environmental features and require developments in their delivery strategies and formulas (Vu and Oh, 2020; Buttmer *et al.*, 2017). This has been suggested as a method for control of emerging tree diseases such as AOD but phage control of these bacteria is yet to be published (Grace *et al.*, 2021). Another option could be the use of rhizosphere bacteria or endophytic members of healthy oak to control pathogen levels, which has been shown to work with a large number of model systems previously (Ciancio, Pieterse and Mercado-Blanco, 2019; De Silva *et al.*, 2019).

Chapter five leads on from the identification of AOD bacteria using specific enrichment, by also isolating and identifying other members of the Enterobacterales from the same samples. The first publication was for the classification of the three novel species of *Scandinavium*. While many of the reported bacteria in the rhizosphere show plant growth-promoting properties, the new isolates are reminders that soil remains a playground for microbial activity with bacteria demonstrating pathogenic traits also identified in this work (Turner, James and Poole, 2013). The reason that these bacteria were initially hypothesised as phytopathogenic was due to their significant association and isolation from roots originating from AOD-positive oak. However, suspicions were further roused when several other strains which were included in this work originated from *Tilia* (lime) with bleeding lesions that are symptomatically similar to AOD, where a novel species of *Brenneria*, *B. tiliae* was also isolated and identified (Kile *et al.*, 2022). Support was then given to the hypothesis based on the identification of numerous virulence genes through whole genome database comparison. The isolates contained the genes required for a functioning type VI secretion system, near complete type II secretion system and many virulence genes. However, pathogenicity trials were not performed with these new strains and as such the isolation point and identification of virulence genes were only used to speculate on their pathogenic potential. Going forward it would be beneficial to identify if these isolates could cause disease through different pathogenicity trials.

A novel genus, *Dryocola*, with two novel species proved to be another taxonomic group of interest, though with a somewhat more complicated background. *Dryocola* forms a clear cluster between *Cedecea* and *Buttiauxella*, and while both have been recorded as being pathogenic in humans they show rather striking differences in their environmental roles (Thompson and Sharkady, 2020; Patra *et al.*, 2018). *Buttiauxella* are frequently associated

with being positive, plant growth-promoting bacteria, which are seen to increase the growth of roots and shoots as well as the overall dry mass of banana seedlings when inoculated in the roots system (Araújo *et al.*, 2021), and as plant growth and cadmium accumulation in *Sedum alfredii* (Wu *et al.*, 2018). Meanwhile, there are numerous reports of *Cedecea* causing disease in different fruiting bodies of mushrooms (Liu *et al.*, 2021; Yan *et al.*, 2019). Though not directly related to plant disease, their ability to infect and cause disease in different eukaryotic organisms does paint them in a different light. As such there was little indication as to what role *Dryocola* might play in the environment and comparison to databases and metagenome comparisons were utilised. The results which were not included in this work demonstrated a large repertoire of virulence genes including a T3SS, numerous effectors, PCWDE and other key genes associated with phytopathogens. Again, further work with pathogenicity trials would be beneficial to understand if these traits are expressed in the phenotype of species of *Dryocola*.

The final set of classifications were for three species spread between two genera, one novel and the other belonging to *Leclercia*. *Leclercia* was composed of a single species of bacteria, which is pathogenic in humans with infections often arising from the environment (Keren *et al.*, 2014; De Baere *et al.*, 2001). However, much like *Buttiauxella*, the species is seen to have a beneficial role to plants in the environment, with strain MO1 associated with heavy metal detoxification (Kang *et al.*, 2021) and strain LSE-1 noted for nitrogen fixing and indole acetic acid producing properties that promote plant growth (Kumawat *et al.*, 2019). As such it was highly suspected that the novel species of *Leclercia*, *L. tamurae* would show similar properties. This was confirmed with database comparisons for plant interaction and plant growth-promoting genes, with *L. tamurae* showing highly similar identified profiles to those of *L. adecarboxylata*. The novel genus *Silvania* was also expected to have similar potential to that of the *Leclercia* species and, again using database comparisons, promising plant growth-promoting bacterial species were identified, with larger numbers of beneficial genes for plant growth being identified in both *Silvania* species than *Leclercia* species.

Overall, the classification work in this chapter has made several contributions to the field of bacterial taxonomy as well as AOD research. The most obvious contribution is illuminating the taxonomic evolutionary relationship between several interesting *Enterobacteriaceae*. However, the three papers all go further than the required polyphasic taxonomic approach

that is currently still used to classify bacteria. Each publication has considered the potential role of the novel species in the environment that they were isolated from and use genome comparisons to justify these predictions. However, a limitation of the work is highlighted here in that none of these predictions utilised *in vitro* trials to see if their genetic potential was seen in the phenotype. This would have been most interesting for the *Dryocola* species, based on the large amount of potentially novel T3SS effectors identified. Nonetheless, the use of these predictions makes the publications more useful when/if the species are identified in future studies. A further contribution was the demonstration of how enrichment can lead to the isolation of large numbers of novel isolates, potentially helping to remedy the fact that so few bacteria can be isolated from soil under laboratory conditions (Steen *et al.*, 2019). This work focused on one enrichment method made for the former family *Enterobacteriaceae*, though now it may be seen as an enrichment method for the order Enterobacterales, due to the ability to recover bacteria from other families such as *Rahnella* from the family Yersiniaceae. This methodology allowed the recovery of nine novel species of bacteria in this work from one sample type; though other potential novel isolates were also obtained it was not possible to classify them all in this work. Further use of different enrichment media could lead to larger portions of the currently uncultivable section of the soil microbiome being identified, resulting in the isolation of more new and interesting isolates. Moreover, it demonstrates the need to continue to work towards preserving our native species and other afflicted species of oak from AOD as they may support unknown numbers of bacteria of potential interest and importance. Possibly sequencing projects could identify groups of interest and then obtain pure isolates using specific enrichment based on the groups physiological growth properties.

The final chapter in this work presented two sequencing projects performed on forty of the oak rhizosphere samples that were identified in chapter one. The first section of site-level analysis could be seen as a proof of concept, a traditional sequencing project. Twenty samples, five healthy and diseased samples from the parkland and woodland were sequenced, with extractions performed on the two positively identified rhizosphere samples for each tree, which were then pooled for sequencing. The results were promising with a difference identified between the woodland and parkland samples, which was expected as soil type and environment have consistently been cited as the most important factors

contributing to microbiome composition in multiple different papers (Deakin *et al.*, 2018; Vos *et al.*, 2013; Martiny *et al.*, 2011). However, the site analysis also revealed a small difference between the healthy and diseased samples within the parkland. This is not the first time that differences in the rhizosphere microbiome of oak have been seen based on the presence of disease. Oak at Hatchlands Park have previously been shown to show lower soil pH with significantly different microbiome compositions than those seen for the healthy trees at the same site. One of the main conclusions was that less extreme soil conditions and plant growth-promoting microbiota were associated with healthy trees (Pinho *et al.*, 2020). Likewise, another study revealed links between ammonia-oxidising bacteria that showed a significant association with asymptomatic oak rhizospheres (Scarlett *et al.*, 2021). These studies both concluded that bacteria with positive roles in plant health were significantly associated with healthy oaks. The consensus of the sequencing data obtained from Hatchlands samples came to a similar conclusion when looking at data collected from the whole site. Parkland healthy oaks supported richer, more diverse rhizosphere microbiomes than their diseased counterparts. Moreover, the bacteria that were identified as being significantly associated with these healthy samples appeared to show PGP properties. *Porphyrobacter* is a genus with limited information available, however, they have been shown to increase plant growth by 60 % in model systems with cucumbers, with one strain COR-2 being patented for its PGP abilities under patent number KR102299675B1 (Zytynska *et al.*, 2020). Another example is *Paenibacillus*, which has previously been identified in microbiome analysis of rhizosphere samples, with links to improved growth conditions for the associated plants. *P. polymyxa* was identified with significant associations, a species previously suggested as a model for host-microbe interactions based on its bio-fertilisation, biocontrol and abiotic stress-reducing abilities (Langendries and Goormachtig, 2021). These are small examples of the 39 bacteria which were exclusively identified in healthy samples, with a further 59 being significantly associated with healthy trees. Thus, for the site-level data, it can be concluded that the health status of the tree appears to have significant effects on the rhizosphere microbiome, based on the knowledge that plant root exudates have significant effects on their microbiome composition (Sasse, Martinoia and Northen, 2018). However, it would be interesting to prove this was true in the case of AOD, to identify if the tree health shaped the microbiome or if microbiome composition was an essential feature in predisposition of oak health.

However, as noted spatial separation is the largest contributor to observable differences in the microbiome. By looking at paired trees differences in the microbiome can be seen in case-by-case samples. This led to different trends being highlighted in the data in which the rhizosphere soil of the paired healthy and diseased trees showed little to no difference. Root endosphere communities showed a larger difference but are more similarly related to the diseased roots than the soil they originated from. This was the expected outcome in some ways, as it is well-recorded that soil supports the largest diversity of bacteria, a subset of which move and colonise the surface and then the endosphere of the root (Compant *et al.*, 2021; Compant, Clément and Sessitsch, 2010). It would be interesting to see if these trends continued further into the tree, with stem and leaf communities, continuing to appear more similarly related to each other than their previous counterpart, but with high variability based on what is recruited into the tree. This has previously been shown in poplars which show low deviation in the variability of the rhizosphere microbiome, with higher variation seen in the endosphere microbiome. The implication here is that at each level of colonisation there is a fine-tuning mechanism for members of the microbiome of each compartment (Beckers *et al.*, 2017). Overall, it is interesting to observe the data in two different ways and see the difference in the outcome of the results. Increased consideration should be given to the sampling methods utilised in large microbiome studies and this set of results shows as much. The oak rhizosphere microbiome does support interesting differences and the role that this has in the development of AOD is worth pursuing. Can the identified groups that are solely associated with the healthy oak rhizosphere be recovered using specific enrichment or other methods? And does their addition to soil help alleviate disease symptoms? The work presented here could function as a stepping-stone into these further areas of interest.

### 7.1. Conclusion

The rhizosphere associated with AOD is an area of high interest. It can clearly support members of the lesion microbiome, and though is unlikely to be the primary source of *B. goodwinii* or *L. britannica*, it may very well function as a source of infection for *R. victoriana* and possibly *G. quercinecans*. Beyond the bacteria associated with AOD, the rhizosphere of oak offers us a fascinating reservoir of undiscovered and potentially important bacterial isolates, though this work only focused on members of the Enterobacterales the number of novel isolates suggests an untapped area for investigation. Finally, the oak rhizosphere

microbiome can be distinguished on several features including sampling location, and by diseased and healthy status of oak in the parkland location. However, in individual paired tree analysis this is not easily observed. Instead, differences based on their soil and endospheres of the roots were identified, which show greater variability than those of the soil. Together these conclusions demonstrate the potential that the rhizosphere of oak has, and the future work section discusses several approaches to understanding this potential.

## 7.2. Future work

- The LAMP assay performed in this work should be further developed to make it possible to use in the field. While it was essential to be able to identify sample origin, the need to perform the assay in labs meant that samples that did not contain oak roots had to be removed from the study. The ability to perform the assay in the field would allow for non-oak samples to be retaken, without the effect of return sampling.
- The enrichment method requires further development, with more suitable broths that do not require pre-enrichment. Alternatively, a wider range of media with more specific selective qualities for members of the AOD lesion microbiome could solve the issue of recovery of species such as *L. britannica* and *B. goodwinii*.
- The enrichment method, followed by HRM analysis for the identification of the AOD bacteria should be trialled on a larger scale. Rhizosphere soil, roots, stems, core samples, acorns, leaves and if possible *Agrilus biguttatus* from a range of different sites should be screened using the methods utilised in this work to allow for the identification of the members of the AOD lesion microbiome. With further sampling trials, the origin and spread of the AOD bacteria might be better understood than was possible in this work.
- Different enrichment schemes should be used to isolate other groups of bacteria from the rhizosphere of oak, which may allow for the taxonomic classification of further novel isolates of interest.
- Virulence assays should be performed with the novel *Scandinaviium* species to identify their pathogenic potential and see if their isolation from both diseased oak rhizosphere and weeping lesions on *Tilia* was coincidental, or if they potentially act as pathogens in these scenarios.

- The effect of the novel species of both *Leclercia* and *Silvania* should be tested on a model system to determine if their identified plant growth-promoting potential can be proven *in vitro*. If they express a beneficial phenotype, could they be used to mitigate disease symptoms in oak?
- Further investigation of the microbiome of oak is warranted. This should be expanded beyond a single site, which was originally the aim of this project.
- Further investigation of the variability in the microbiome through sequencing of the stem and leaves may yield interesting results relating to how differences between the diseased and healthy endosphere microbiome arise.

Long term monitoring of a healthy tree and diseased trees microbiome and root exudate profiles should be performed to identify where the differences observed between healthy and diseased trees originate from.

## References

- Acinas, S.G., Marcelino, L.A., Klepac-Ceraj, V. and Polz, M.F. (2004) Divergence and Redundancy of 16S rRNA Sequences in Genomes with Multiple *rrn* Operons. *Journal of Bacteriology*. [online]. 186 (9), pp.2629–2635.
- Adeolu, M., Alnajar, S., Naushad, S. and Gupta, R.S. (2016) Genome-based phylogeny and taxonomy of the ‘*Enterobacteriales*’: Proposal for enterobacterales ord. nov. divided into the families *Enterobacteriaceae*, *Erwiniaceae* fam. nov., *Pectobacteriaceae* fam. nov., *Yersiniaceae* fam. nov., *Hafniaceae* fam. nov., *Morgane*. *International Journal of Systematic and Evolutionary Microbiology*. [online]. 66 (12), pp.5575–5599.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. (2002) *Molecular Biology of the Cell* 4th edition. New York, Garland Science.
- Alexander, K.N.. (2003a) Changing distributions of *Cantharidae* and *Buprestidae* within Great Britain (Coleoptera). *Proceedings 13th international colloquium European Invertebrate Survey, September 2001*. 91 (September 2001), pp.87–91.
- Alexander, K.N.A. (2003b) *Provisional atlas of the Cantharoidea and Buprestoidea (Coleoptera) of Britain and Ireland* Huntingdon, Ang lia Print Solutio ns.
- Allen, CD., Macalady, AK., Chenchouni, H., Bachelet, D., McDowell, N and Vennetier, M., (2010) A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management*. [online]. 259 (4), pp.660–684.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *Journal of molecular biology*. 215 (3), pp.403–410.
- Amaresan, N., Kumar, M.S., Annapurna, K. and Kumar, K. (2020) *Beneficial Microbes in Agro-Ecology* Amaresan, N., Kumar, M.S., Annapurna, K. and Kumar, K. (eds.) [online]. Academic Press.
- An, S.Q., Potnis, N., Dow, M., Vorhölter, FJ., He, Y.Q and Becker., A. (2019) Mechanistic insights into host adaptation, virulence and epidemiology of the phytopathogen *Xanthomonas*. *Federation of European Microbiological Societies Microbiology Reviews*. [online]. 44 (1), pp.1–32.

Antiabong, J.F., Boardman, W. and Ball, A.S. (2014) What can we learn from the microbial ecological interactions associated with polymicrobial diseases? *Veterinary Immunology and Immunopathology*. [online]. 158 (1–2), Elsevier B.V., pp.30–36. Available from: <http://dx.doi.org/10.1016/j.vetimm.2013.03.009>.

Araújo, R.C. de, Rodrigues, F.A., Nadal, M.C., Ribeiro, M. de S., Antônio, C.A.C., Rodrigues, V.A., Souza, A.C. de, Pasqual, M. and Dória, J. (2021) Acclimatization of *Musa* spp. seedlings using endophytic *Bacillus* spp. and *Buttiauxella agrestis* strains. *Microbiological research*. [online]. 248 (March), p.126750.

Aslam, Z., Yasir, M., Khaliq, A., Matsui, K. and Chung, Y.R. (2010) Mini Review Too much bacteria still unculturable. *Crop & Environment*. 1 (1), pp.59–60.

Asselin, J.E., Eikemo, H., Perminow, J., Nordskog, B., Brurberg, M.B. and Beer, S. V. (2019) *Rahnella* spp. are commonly isolated from onion (*Allium cepa*) bulbs and are weakly pathogenic. *Journal of Applied Microbiology*. [online]. 127 (3), pp.812–824.

De Baere, T., Wauters, G., Huylensbroeck, A., Claeys, G., Peleman, R., Verschraegen, G., Allemeersch, D. and Vaneechoutte, M. (2001) Isolations of *leclercia adecarboxylata* from a patient with a chronically inflamed gallbladder and from a patient with sepsis without focus. *Journal of Clinical Microbiology*. [online]. 39 (4), pp.1674–1675.

Bağcı, C., Patz, S. and Huson, D.H. (2021) DIAMOND+MEGAN: Fast and Easy Taxonomic and Functional Analysis of Short and Long Microbiome Sequences. *Current Protocols*. [online]. 1 (3), pp.1–29.

Bakken, L.R. (1985) Separation and purification of bacteria from soil. *Applied and Environmental Microbiology*. 49 (6), pp.1482–1487.

Bakken, L.R. and Olsen, R.A. (1987) The relationship between cell size and viability of soil bacteria. *Microbial Ecology*. [online]. 13 (2), pp.103–114.

Balint-Kurti, P. (2019) The plant hypersensitive response: concepts, control and consequences. *Molecular Plant Pathology*. [online]. 20 (8), pp.1163–1178.

Barillot, C.D.C., Sarde, C.O., Bert, V., Tarnaud, E. and Cochet, N. (2013) A standardized method for the sampling of rhizosphere and rhizoplan soil bacteria associated to a herbaceous root

system. *Annals of Microbiology*. [online]. 63 (2), pp.471–476.

Bartram, A.K., Lynch, M.D.J., Stearns, J.C., Moreno-Hagelsieb, G. and Neufeld, J.D. (2011) Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end Illumina reads. *Applied and Environmental Microbiology*. [online]. 77 (11), pp.3846–3852.

Bashir, I., War, A.F., Rafiq, I., Reshi, Z.A., Rashid, I. and Shouche, Y.S. (2022) Phyllosphere microbiome: Diversity and functions. *Microbiological Research*. [online]. 254 (May 2021), Elsevier GmbH, p.126888. Available from: <https://doi.org/10.1016/j.micres.2021.126888>.

Batey, T. and McKenzie, D.C. (2006) Soil compaction: Identification directly in the field. *Soil Use and Management*. [online]. 22 (2), pp.123–131.

Beckers, B., De Beeck, M.O., Weyens, N., Boerjan, W. and Vangronsveld, J. (2017) Structural variability and niche differentiation in the rhizosphere and endosphere bacterial microbiome of field-grown poplar trees. *Microbiome*. [online]. 5 (1), Microbiome, pp.1–17.

Benizri, E., Baudoin, E. and Guckert, A. (2001) Root colonization by inoculated plant growth-promoting rhizobacteria. *Biocontrol Science and Technology*. [online]. 11 (5), pp.557–574.

Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J. and Sayers, E.W. (2013) GenBank. *Nucleic Acids Research*. [online]. 41 (D1), pp.36–42.

Berg, G., Grube, M., Schloter, M. and Smalla, K. (2014) Unraveling the plant microbiome: Looking back and future perspectives. *Frontiers in Microbiology*. [online]. 5 (JUN), pp.1–7.

Berg, G., Zachow, C., Müller, H., Philipps, J. and Tilcher, R. (2013) Next-Generation Bio-Products Sowing the Seeds of Success for Sustainable Agriculture. *Agronomy*. [online]. 3 (4), pp.648–656.

Bharti, R. and Grimm, D.G. (2021) *Current challenges and best-practice protocols for microbiome analysis Briefings in Bioinformatics*. [online]. 22 (1), 178–193. Available from: <https://github.com/grimmlab/MicrobiomeBestPracticeReview>.

Bhattacharyya, D., Duta, S., Yu, S.M., Jeong, S.C. and Lee, Y.H. (2018) Taxonomic and functional changes of bacterial communities in the rhizosphere of Kimchi cabbage after seed bacterization with *Proteus vulgaris* JBS202. *Plant Pathology Journal*. [online]. 34 (4), pp.286–

296.

Van Der Biezen, E.A. and Jones, J.D.G. (1998) Plant disease-resistance proteins and the gene-for-gene concept. *Trends in Biochemical Sciences*. [online]. 23 (12), pp.454–456.

Biosca, E.G., González, R., López-López, M.J., Soria, S., Montón, C., Pérez-Laorga, E. and López, M.M. (2003) Isolation and characterization of *Brenneria quercina*, causal agent for bark canker and drippy nut of *Quercus* spp. in Spain. *Phytopathology*. [online]. 93 (4), pp.485–492.

Blaud, A., Menon, M., van der Zaan, B., Lair, G.J. and Banwart, S.A. (2017) *Effects of Dry and Wet Sieving of Soil on Identification and Interpretation of Microbial Community Composition Advances in Agronomy*. 1st edition. [online]. 142, Elsevier Inc. Available from: <http://dx.doi.org/10.1016/bs.agron.2016.10.006>.

Blom, D., Fabbri, C., Connor, E.C., Schiestl, F.P., Klauser, D.R., Boller, T., Eberl, L. and Weiskopf, L. (2011) Production of plant growth modulating volatiles is widespread among rhizosphere bacteria and strongly depends on culture conditions. *Environmental Microbiology*. [online]. 13 (11), pp.3047–3058.

Boath, J.M., Dakhal, S., Van, T.T.H., Moore, R.J., Dekiwadia, C. and Macreadie, I.G. (2020) Polyphasic characterisation of *Cedecea colo* sp. Nov., a new enteric bacterium isolated from the koala hindgut. *Microorganisms*. [online]. 8 (2).

Bolger, A.M., Lohse, M. and Usadel, B. (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*. [online]. 30 (15), pp.2114–2120.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C and Al-Ghalith, G.A., (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*. [online]. 37 (8), pp.852–857.

Brady, C., Allainguillaume, J., Denman, S. and Arnold, D. (2016) Rapid identification of bacteria associated with Acute Oak Decline by high-resolution melt analysis. *Letters in Applied Microbiology*. [online]. 63 (2), pp.89–95.

Brady, C., Arnold, D., McDonald, J. and Denman, S. (2017) Taxonomy and identification of bacteria associated with acute oak decline. *World Journal of Microbiology and Biotechnology*. [online]. 33 (7), Springer Netherlands, pp.1–11.

Brady, C., Asselin, J.A., Beer, S., Brurberg, M.B., Crampton, B., Venter, S., Arnold, D. and Denman, S. (2022) *Rahnella perminowiae* sp. nov., *Rahnella bonaserana* sp. nov., *Rahnella rivi* sp. nov. and *Rahnella ecdela* sp. nov., isolated from diverse environmental sources, and emended description of the genus *Rahnella*. *International Journal of Systematic and Evolutionary Microbiology*.

Brady, C., Denman, S., Kirk, S., Venter, S., Rodríguez-Palenzuela, P. and Coutinho, T. (2010) Description of *Gibbsiella quercinecans* gen. nov., sp. nov., associated with Acute Oak Decline. *Systematic and Applied Microbiology*. [online]. 33 (8), Elsevier GmbH., pp.444–450. Available from: <http://dx.doi.org/10.1016/j.syapm.2010.08.006>.

Brady, C., Hunter, G., Kirk, S., Arnold, D. and Denman, S. (2014a) Description of *Brenneria roseae* sp. nov. and two subspecies, *Brenneria roseae* subspecies *roseae* ssp. nov and *Brenneria roseae* subspecies *americana* ssp. nov. isolated from symptomatic oak. *Systematic and Applied Microbiology*. [online]. 37 (6), Elsevier GmbH., pp.396–401. Available from: <http://dx.doi.org/10.1016/j.syapm.2014.04.005>.

Brady, C., Hunter, G., Kirk, S., Arnold, D. and Denman, S. (2014b) *Rahnella victoriana* sp. nov., *Rahnella bruchi* sp. nov., *Rahnella woolbedingensis* sp. nov., classification of *Rahnella* genomospecies 2 and 3 as *Rahnella variigena* sp. nov. and *Rahnella inusitata* sp. nov., respectively and emended description of the genus *Rahnella*. *Systematic and Applied Microbiology*. [online]. 37 (8), Elsevier GmbH., pp.545–552. Available from: <http://dx.doi.org/10.1016/j.syapm.2014.09.001>.

Brady, C.L., Cleenwerck, I., Denman, S., Venter, S.N., Rodríguez-Palenzuela, P., Coutinho, T.A. and De Vos, P. (2012) Proposal to reclassify *Brenneria quercina* (Hildebrand and Schroth 1967) Hauben et al. 1999 into a new genus, *Lonsdalea* gen. nov., as *Lonsdalea quercina* comb. nov., descriptions of *Lonsdalea quercina* subsp. *quercina* comb. nov., *Lonsdalea quercina* subsp. *iberica* subsp. nov. and *Lonsdalea quercina* subsp. *britannica* subsp. nov., emendation of the description of the genus *Brenneria*, reclassification of *Dickeya dieffenbachiae* as *Dickeya dadantii* subsp. *dieffenbachiae* comb. nov., and emendation of the description of *Dickeya dadantii*. *International journal of systematic and evolutionary microbiology*. [Online] 62 (7), 1592–1602.

Brink, S.C. (2016) Unlocking the Secrets of the Rhizosphere. *Trends in Plant Science*. [online].

21 (3), Elsevier Ltd, pp.169–170. Available from: <http://dx.doi.org/10.1016/j.tplants.2016.01.020>.

Broberg, M., Doonan, J., Mundt, F., Denman, S. and McDonald, J.E. (2018) Integrated multi-omic analysis of host-microbiota interactions in acute oak decline. *Microbiome*. [online]. 6 (1), Microbiome, pp.1–15.

Brown, A. and Benson, H. (2009) *Benson's Microbiological Applications*. Boston, McGraw-Hill Higher Education.

Brown, N., van den Bosch, F., Parnell, S. and Denman, S. (2017a) Integrating regulatory surveys and citizen science to map outbreaks of forest diseases: Acute oak decline in England and Wales. *Proceedings of the Royal Society B: Biological Sciences*. [online]. 284 (1859).

Brown, N., Inward, D.J.G., Jeger, M. and Denman, S. (2015) A review of *Agrilus biguttatus* in UK forests and its relationship with acute oak decline. *Forestry*. [online]. 88 (1), pp.53–63.

Brown, N., Jeger, M., Kirk, S., Williams, D., Xu, X., Pautasso, M. and Denman, S. (2017b) Acute oak decline and *Agrilus biguttatus*: The co-occurrence of stem bleeding and D-shaped emergence holes in Great Britain. *Forests*. [online]. 8 (3).

Brown, N., Jeger, M., Kirk, S., Xu, X. and Denman, S. (2016) Spatial and temporal patterns in symptom expression within eight woodlands affected by Acute Oak Decline. *Forest Ecology and Management*. [online]. 360, Elsevier B.V., pp.97–109. Available from: <http://dx.doi.org/10.1016/j.foreco.2015.10.026>.

Brown, N., Vangelova, E., Parnell, S., Broadmeadow, S. and Denman, S. (2018) Predisposition of forests to biotic disturbance: Predicting the distribution of Acute Oak Decline using environmental factors. *Forest Ecology and Management*. [online]. 407 (November 2017), Elsevier, pp.145–154. Available from: <https://doi.org/10.1016/j.foreco.2017.10.054>.

De Bruijn, F.J. (1992) Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergeneric consensus) sequences and the polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* isolates and other soil bacteria. *Applied and Environmental Microbiology*. [online]. 58 (7), pp.2180–2187.

Brumfield, K.D., Huq, A., Colwell, R.R., Olds, J.L. and Leddy, M.B. (2020) Microbial resolution

of whole genome shotgun and 16S amplicon metagenomic sequencing using publicly available NEON data. *Public Library of Science ONE*. [online]. 15 (2), pp.1–21.

Buchfink, B., Reuter, K. and Drost, H.G. (2021) Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nature Methods*. [online]. 18 (4), Springer US, pp.366–368. Available from: <http://dx.doi.org/10.1038/s41592-021-01101-x>.

Buckley, D.H., Huangyutitham, V., Nelson, T.A., Rumberger, A. and Thies, J.E. (2006) Diversity of *Planctomycetes* in soil in relation to soil history and environmental heterogeneity. *Applied and Environmental Microbiology*. [online]. 72 (7), pp.4522–4531.

Bueno-Gonzalez, V., Brady, C., Denman, S., Plummer, S., Allainguillaume, J. and Arnold, D. (2019) *Pseudomonas daroniae* sp. Nov. and *pseudomonas dryadis* sp. nov., isolated from pedunculate oak affected by acute oak decline in the uk. *International Journal of Systematic and Evolutionary Microbiology*. [online]. 69 (11), pp.3368–3376.

Buesing, N. and Gessner, M.O. (2002) Comparison of detachment procedures for direct counts of bacteria associated with sediment particles plant litter and epiphytic biofilms. *Aquatic Microbial Ecology*. [online]. 27 (1), pp.29–36.

Bulgarelli, D., Rott, M., Schlaeppli, K., Ver Loren van Themaat E., Ahmadinejad, N and Assenza, F.(2012) Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. *Nature*. [online]. 488 (7409), Nature Publishing Group, pp.91–95. Available from: <http://dx.doi.org/10.1038/nature11336>.

Buonaurio, R., Moretti, C., Da Silva, D.P., Cortese, C., Ramos, C. and Venturi, V. (2015) The olive knot disease as a model to study the role of interspecies bacterial communities in plant disease. *Frontiers in Plant Science*. [online]. 6 (June), pp.1–12.

Buttimer, C., McAuliffe, O., Ross, R.P., Hill, C., O'Mahony, J. and Coffey, A. (2017) Bacteriophages and bacterial plant diseases. *Frontiers in Microbiology*. [online]. 8 (JAN), pp.1–15.

Buttner, D. (2012) Protein Export According to Schedule: Architecture, Assembly, and Regulation of Type III Secretion Systems from Plant- and Animal-Pathogenic Bacteria. *Microbiology and Molecular Biology Reviews*. [online]. 76 (2), pp.262–310.

Callahan, B.J., Grinevich, D., Thakur, S., Balamotis, M.A. and Yehezkel, T. Ben (2020) Ultra-accurate microbial amplicon sequencing directly from complex samples with synthetic long reads. *bioRxiv*. [online].

Carrigg, C., Rice, O., Kavanagh, S., Collins, G. and O'Flaherty, V. (2007) DNA extraction method affects microbial community profiles from soils and sediment. *Applied Microbiology and Biotechnology*. [online]. 77 (4), pp.955–964.

Chaparro, J.M., Badri, D. V. and Vivanco, J.M. (2014) Rhizosphere microbiome assemblage is affected by plant development. *International Society for Microbial Ecology Journal*. [online]. 8 (4), Nature Publishing Group, pp.790–803. Available from: <http://dx.doi.org/10.1038/ismej.2013.196>.

Chaparro, J.M., Sheflin, A.M., Manter, D.K. and Vivanco, J.M. (2012) Manipulating the soil microbiome to increase soil health and plant fertility. *Biology and Fertility of Soils*. [online]. 48 (5), pp.489–499.

Charkowski, A., Blanco, C., Condemine, G., Expert, D., Franza, T and Hayes, C. (2012) The Role of Secretion Systems and Small Molecules in Soft-Rot Enterobacteriaceae Pathogenicity. *Annual Review of Phytopathology*. [online]. 50 (1), pp.425–449.

Chisholm, S.T., Coaker, G., Day, B. and Staskawicz, B.J. (2006) Host-microbe interactions: Shaping the evolution of the plant immune response. *Cell*. [online]. 124 (4), pp.803–814.

Cho, J.C. and Tiedje, J.M. (2000) Biogeography and degree of endemism of fluorescent *Pseudomonas* strains in soil. *Applied and Environmental Microbiology*. [online]. 66 (12), pp.5448–5456.

Choi, H.J., Kim, Y.J., Lim, Y. and Park, D.H. (2019) Survival of *Erwinia amylovora* on Surfaces of Materials Used in Orchards. *Research in Plant Disease*. [online]. 25 (2), pp.89–93.

Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahal, D.R and da Costa, M.S. (2018) Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*. [online]. 68 (1), pp.461–466.

Church, D.L., Cerutti, L., Gürtler, A., Griener, T., Zelazny, A. and Emler, S. (2020) Performance and application of 16S rRNA gene cycle sequencing for routine identification of bacteria in the

- clinical microbiology laboratory. *Clinical Microbiology Reviews*. [online]. 33 (4), pp.1–74.
- Ciancio, A., Pieterse, C.M.J. and Mercado-Blanco, J. (2019) Editorial: Harnessing useful rhizosphere microorganisms for pathogen and pest biocontrol-second edition. *Frontiers in Microbiology*. [online]. 10 (AUG), pp.1–5.
- Cianciotto, N.P. (2005) Type II secretion: A protein secretion system for all seasons. *Trends in Microbiology*. [online]. 13 (12), pp.581–588.
- Clarke, S.C. (2004) Collins and Lyne's Microbiological Methods. *British Journal of Biomedical Science*. [online]. 61 (2), pp.113–113.
- Clarridge, J.E. (2004) Impact of 16S rRNA Gene Sequence Analysis for Identification of Bacteria on Clinical Microbiology and Infectious Diseases Impact of 16S rRNA Gene Sequence Analysis for Identification of Bacteria on Clinical Microbiology and Infectious Diseases. *Clinical Microbiology Reviews*. [online]. 17 (4), pp.840–862.
- Coenye, T., Falsen, E., Vancanneyt, M., Hoste, B., Govan, J.R.W., Kersters, K. and Vandamme, P. (1999) Classification of *Alcaligenes faecalis*-like isolates from the environment and human clinical samples as *Ralstonia gilardii* sp. nov. *International Journal of Systematic Bacteriology*. [online]. 49 (2), pp.405–413.
- Compant, S., Cambon, M.C., Vacher, C., Mitter, B., Samad, A. and Sessitsch, A. (2021) The plant endosphere world – bacterial life within plants. *Environmental Microbiology*. [online]. 23 (4), pp.1812–1829.
- Compant, S., Clément, C. and Sessitsch, A. (2010) Plant growth-promoting bacteria in the rhizo- and endosphere of plants: Their role, colonization, mechanisms involved and prospects for utilization. *Soil Biology and Biochemistry*. [online]. 42 (5), pp.669–678.
- Compant, S., Reiter, B., Sessitsch, A., Nowak, J., Clement, C. and Barka, E.A. (2005) Endophytic Colonization of *Vitis vinifera* L. by Plant Growth Promoting Bacterium *Burkholderia* sp. Strain PsJN. *Federation of European Microbiological Societies Microbial Ecology*. [online]. 71 (4), pp.1685–1693.
- Crampton, B., Brady, C. and Denman, S. (2022) *Bacterial Tree Disease Fact Sheets* [online]. Available from: <https://bacterialplantdiseases.uk/new-bacterial-tree-disease-factsheets/>.

- Crampton, B.G., Plummer, S.J., Kaczmarek, M., McDonald, J.E. and Denman, S. (2020) A multiplex real-time PCR assay enables simultaneous rapid detection and quantification of bacteria associated with acute oak decline. *Plant Pathology*. [online]. 69 (7), pp.1301–1310.
- Crowther, T.W. and Grossart, H.P. (2015) The role of bottom-up and top-down interactions in determining microbial and fungal diversity and function. *Trophic Ecology: Bottom-Up and Top-Down Interactions Across Aquatic and Terrestrial Systems*. [online]. pp.260–287.
- Czajkowski, R., Pérombelon, M.C.M., Jafra, S., Lojkowska, E., Potrykus, M., Van Der Wolf, J.M. and Sledz, W. (2015) *Detection, identification and differentiation of Pectobacterium and Dickeya species causing potato blackleg and tuber soft rot: A review Annals of Applied Biology*. [online]. 166 (1), Blackwell Publishing Ltd, 18–38.
- Davis, K.E.R., Joseph, S.J. and Janssen, P.H. (2005) Effects of growth medium, inoculum size, and incubation time on culturability and isolation of soil bacteria. *Applied and Environmental Microbiology*. [online]. 71 (2), pp.826–834.
- Dawkins, T. and Ellwood, F. (2015) Ground rules. *Biologist*. [online]. 62 (5), pp.26–30.
- Deakin, G., Tilston, E.L., Bennett, J., Passey, T., Harrison, N., Fernández-Fernández, F. and Xu, X. (2018) Spatial structuring of soil microbial communities in commercial apple orchards. *Applied Soil Ecology*. [online]. 130, Elsevier B.V., pp.1–12.
- DeAngelis, K.M., Brodie, E.L., DeSantis, T.Z., Andersen, G.L., Lindow, S.E. and Firestone, M.K. (2009) Selective progressive response of soil microbial community to wild oat roots. *International Society for Microbial Ecology Journal*. [online]. 3 (2), pp.168–178.
- Delétoile, A., Decré, D., Courant, S., Passet, V., Audo, J., Grimont, P., Arlet, G. and Brisse, S. (2009) Phylogeny and identification of *Pantoea* species and typing of *Pantoea agglomerans* strains by multilocus gene sequencing. *Journal of Clinical Microbiology*. [online]. 47 (2), pp.300–310.
- Delgado-Baquerizo, M., Guerra, C.A., Cano-Díaz, C., Egidi, E., Wang, J.T., Eisenhauer, N., Singh, B.K. and Maestre, F.T. (2020) The proportion of soil-borne pathogens increases with warming at the global scale. *Nature Climate Change*. [online]. 10 (6), Springer US, pp.550–554. Available from: <http://dx.doi.org/10.1038/s41558-020-0759-3>.

Delmont, T.O., Robe, P., Cecillon, S., Clark, I.M., Constancias, F., Simonet, P., Hirsch, P.R. and Vogel, T.M. (2011) Accessing the Soil Metagenome for Studies of Microbial Diversity †. *Applied and Environmental Microbiology*. [online]. 77 (4), pp.1315–1324. Available from: <http://www.rothamsted.ac.uk/for>.

Denman, S., Doonan, J., Ransom-Jones, E., Broberg, M., Plummer, S and Kirk, S. (2018) Microbiome and infectivity studies reveal complex polyspecies tree disease in Acute Oak Decline. *International Society for Microbial Ecology Journal*. [online]. 12 (2), Nature Publishing Group, pp.386–399. Available from: <http://dx.doi.org/10.1038/ismej.2017.170>.

Denman, S., Brady, C., Kirk, S., Cleenwerck, I., Venter, S., Coutinho, T. and De Vos, P. (2012) *Brenneria goodwinii* sp. nov., associated with acute oak decline in the UK. *International Journal of Systematic and Evolutionary Microbiology*. [online]. 62 (10), pp.2451–2456.

Denman, S., Brown, N., Kirk, S., Jeger, M. and Webber, J. (2014) A description of the symptoms of Acute Oak Decline in Britain and a comparative review on causes of similar disorders on oak in Europe. *Forestry*. [online]. 87 (4), pp.535–551.

Denman, S., Brown, N., Vanguelova, E. and Crampton, B. (2022) Temperate Oak Declines: Biotic and abiotic predisposition drivers In: *Forest Microbiology*. [online]. Academic Press, 239–263.

Denman, S., Kirk, S. and Webber, J. (2010) Managing acute oak decline. *Practice Note - Forestry Commission*. [online]. (15), p.6 pp. Available from: [http://apps.webofknowledge.com/full\\_record.do?product=UA&search\\_mode=GeneralSearch&qid=1&SID=U2zmSzSoyz654hMpjwu&page=1&doc=2](http://apps.webofknowledge.com/full_record.do?product=UA&search_mode=GeneralSearch&qid=1&SID=U2zmSzSoyz654hMpjwu&page=1&doc=2).

Denman, S., Plummer, S., Kirk, S., Peace, A. and McDonald, J.E. (2016) Isolation studies reveal a shift in the cultivable microbiome of oak affected with Acute Oak Decline. *Systematic and Applied Microbiology*. [online]. 39 (7), Elsevier GmbH., pp.484–490. Available from: <http://dx.doi.org/10.1016/j.syapm.2016.07.002>.

Denman, S. and Webber, J. (2009) Oak declines: new definitions and new episodes in Britain. *Quarterly Journal of Forestry*. 103 (4), pp.285–290.

Donaubauer, E. (1987) Occurrence of diseases and pathogens of the oak and its relation to oak decline. *Oesterr. Forstztg*. 98 (3), pp.46–48.

Doonan, J., Denman, S., Gertler, C., Pachebat, J.A., Golyshin, P.N. and McDonald, J.E. (2015) The intergenic transcribed spacer region 1 as a molecular marker for identification and discrimination of Enterobacteriaceae associated with acute oak decline. *Journal of Applied Microbiology*. [online]. 118 (1), pp.193–201.

Doonan, J., Denman, S., Pachebat, J.A. and McDonald, J.E. (2019) Genomic analysis of bacteria in the Acute Oak Decline pathobiome. *Microbial Genomics*. [online]. 5 (1).

Drummond, D.A. and Wilke, C.O. (2008) Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. *Accounts of Chemical Research*. [online]. 134 (2), pp.341–352.

Dunn, J.P., Potter, D.A. and Kimmerer, T.W. (1990) Carbohydrate reserves, radial growth, and mechanisms of resistance of oak trees to phloem-boring insects. *Oecologia*. 83, pp.458–468.

Dykhuizen, D.E. (1998) Santa Rosalia revisited: Why are there so many species of bacteria? *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*. [online]. 73 (1), pp.25–33.

Dykhuizen, D.E. and Kalia, A. (2007) The Population Structure of Pathogenic Bacteria: Stearns, S. and Koella, J. (eds.) *Evolution in Health and Disease*. 2nd edition. [online]. 15 (1), Oxford, Oxford University Press (OUP), 185–198.

Er, T.K. and Chang, J.G. (2012) High-resolution melting: Applications in genetic disorders. *The International Journal of Clinical Chemistry* [online]. 414, Elsevier B.V., pp.197–201. Available from: <http://dx.doi.org/10.1016/j.cca.2012.09.012>.

Evans, H.F., Moraal, L.G. and Pajares, J.A. (2004) Biology, ecology and economic importance of *Buprestidae* and *Cerambycidae*: Lieutier, F., Day, K.R., Battisti, A., Grégoire, J.C. and Evans, H.F. (eds.) *Bark and wood boring insects in living trees in Europe, a synthesis*. Dordrecht, Kluwer Academic Publishers, 447–474.

Fadiji, A.E. and Babalola, O.O. (2020) Metagenomics methods for the study of plant-associated microbial communities: A review. *Journal of Microbiological Methods*. [online]. 170 (December 2019), Elsevier, p.105860. Available from: <https://doi.org/10.1016/j.mimet.2020.105860>.

Fadiji, A.E., Kanu, J.O. and Babalola, O.O. (2021) Metagenomic profiling of rhizosphere microbial community structure and diversity associated with maize plant as affected by cropping systems. *International Microbiology*. [online]. 24 (3), International Microbiology, pp.325–335.

Fan, T.W.M., Lane, A.N., Pedler, J., Crowle, D. and Higashi, R.M. (1997) Comprehensive Analysis of Organic Ligands in Whole Root Exudates Using Nuclear Magnetic Resonance and Gas Chromatography–Mass Spectrometry. *Analytical Biochemistry*. [online]. 68 (251), pp.57–68.

Farmer, J.J. (2015) *Cedecea*. *Bergey's Manual of Systematics of Archaea and Bacteria*. [online]. pp.1–14.

Farmer, J.J., Davis, B.R., Hickman-Brenner, F.W., McWhorter, A., Huntley-Carter, G.P and Asbury, M.A. (1985) Biochemical identification of new species and biogroups of *Enterobacteriaceae* isolated from clinical specimens. *Journal of Clinical Microbiology*. [online]. 21 (1), pp.46–76.

Farris, J.S. (1972) *Estimating Phlogenetic Trees from Distance Matrices*. 106 (951), pp.645–670.

Feil, E.J. and Spratt, B.G. (2001) Recombination and the Population structures of bacterial pathogens. *Annual Review of Microbiology*. 55 (1), pp.561–590.

Fernandes, C., Duarte, L., Naves, P., Sousa, E. and Cruz, L. (2022) First report of *Brenneria goodwinii* causing acute oak decline on *Quercus suber* in Portugal. *Journal of Plant Pathology*. [online]. 143.

Fernández-González, A.J., Martínez-Hidalgo, P., Cobo-Díaz, J.F., Villadas, P.J., Martínez-Molina, E., Toro, N., Tringe, S.G. and Fernández-López, M. (2017) The rhizosphere microbiome of burned holm-oak: Potential role of the genus *Arthrobacter* in the recovery of burned soils. *Scientific Reports*. [online]. 7 (1). Available from: [www.nature.com/scientificreports](http://www.nature.com/scientificreports) [Accessed 13 October 2022].

Ferragut, C., Izard, D. and Gavini, F. (1981) *Buttiauxella*, a new genus of the family *Enterobacteraceae*. *Zentralblatt fur Bakteriologie. Allgemeine Angewandte und Okologische Microbiologie Abt.1 Orig.C Hyg*. [online]. 2 (1), Gustav Fischer Verlag, Stuttgart/New York,

pp.33–44. Available from: [http://dx.doi.org/10.1016/S0721-9571\(81\)80016-6](http://dx.doi.org/10.1016/S0721-9571(81)80016-6).

Finch, J.P., Brown, N., Beckmann, M., Denman, S. and Draper, J. (2021) Index measures for oak decline severity using phenotypic descriptors. *Forest Ecology and Management*. [online]. 485, Elsevier B.V., p.118948. Available from: <https://doi.org/10.1016/j.foreco.2021.118948><https://linkinghub.elsevier.com/retrieve/pii/S0378112721000372>.

Finer, K.R., Fox, L. and Finer, J.J. (2016) Isolation and Characterization of *Agrobacterium* Strains from Soil: A Laboratory Capstone Experience . *Journal of Microbiology & Biology Education*. [online]. 17 (3), pp.444–450.

Fisher, M.M. and Triplett, E.W. (1999) Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. *Applied and Environmental Microbiology*. [online]. 65 (10), pp.4630–4636.

Forest Research (2014) Rapid PRA for Acute Oak Decline. pp.1–27.

Forest Research (2020) *Typical Symptoms Of Acute Oak Decline Forest Research*. [online].2020

Forest Research (2021) *Woodland Area By Species: Broadleaves Forest Research*. [online].2021 [online]. Available from: <https://www.forestresearch.gov.uk/tools-and-resources/statistics/forestry-statistics/forestry-statistics-2019/woodland-area-and-planting/national-forest-inventory/woodland-area-by-species-broadleaves/>.

Fox, G.E., Magrum, L.J., Balcht, W.E., Wolfef, R.S. and Woese, C.R. (1977) Classification of methanogenic bacteria by 16S ribosomal RNA characterization (comparative oligonucleotide cataloging/phylogeny/molecular evolution). *Evolution*. 74 (10), pp.4537–4541.

Franklin, R.B. and Mills, A.L. (2003) Multi-scale variation in spatial heterogeneity for microbial community structure in an eastern Virginia agricultural field. *Federation of European Microbiological Societies Microbiology Ecology*. [online]. 44 (3), pp.335–346.

Gamalero, E., Ingua, G.L., Erta, G.B. and Emanceau, P.L. (2003) Review article Methods for studying root colonization by introduced beneficial bacteria. *Agronomie*. [online]. 23, pp.407–418.

Garritano, S., Gemignani, F., Voegelé, C., Nguyen-Dumont, T., Le Calvez-Kelm, F., De Silva, D.,

- Lesueur, F., Landi, S. and Tavtigian, S. V. (2009) Determining the effectiveness of High Resolution Melting analysis for SNP genotyping and mutation scanning at the TP53 locus. *British Medical Council Genetics*. [online]. 10, pp.1–12.
- van der Gast, C.J. (2015) Microbial biogeography: The end of the ubiquitous dispersal hypothesis? *Environmental Microbiology*. [online]. 17 (3), pp.544–546.
- Gathercole, L.A.P., Nocchi, G., Brown, N., Coker, T.L.R., Plumb, W.J., Stocks, J.J., Nichols, R.A., Denman, S. and Buggs, R.J.A. (2021) Evidence for the widespread occurrence of bacteria implicated in acute oak decline from incidental genetic sampling. *Forests*. [online]. 12 (12).
- Gause, G.F. (1934) Experimental Analysis of Vito Volterra's Mathematical Theory of the Struggle for Existence. *Science*. 79 (2036), pp.16–18.
- Geider, K., Gernold, M., Jock, S., Wensing, A., Völksch, B., Gross, J. and Spiteller, D. (2015) Unifying bacteria from decaying wood with various ubiquitous *Gibbsiella* species as *G. acetica* sp. nov. based on nucleotide sequence similarities and their acetic acid secretion. *Microbiological Research*. [online]. 181, Elsevier GmbH., pp.93–104. Available from: <http://dx.doi.org/10.1016/j.micres.2015.05.003>.
- Gevers, D., Cohan, F.M., Lawrence, J.G., Spratt, B.G., Coenye, T and Feil, E.J. (2005) Re-evaluating prokaryotic species. *Nature Reviews Microbiology*. [online]. 3 (9), pp.733–739.
- Gibbs, J.N. and Greig, B.J.W. (1997) Biotic and abiotic factors affecting the dying back of pedunculate oak *Quercus robur* L. *Forestry*. [online]. 70 (4), pp.403–406.
- Gimenez, E., Salinas, M. and Manzano-Agugliaro, F. (2018) Worldwide research on plant defense against biotic stresses as improvement for sustainable agriculture. *Sustainability (Switzerland)*. [online]. 10 (2), pp.1–19.
- Glaeser, S.P. and Kämpfer, P. (2015a) Multilocus sequence analysis (MLSA) in prokaryotic taxonomy. *Systematic and Applied Microbiology*. [online]. 38 (4), Elsevier GmbH., pp.237–245. Available from: <http://dx.doi.org/10.1016/j.syapm.2015.03.007>.
- González, A.J. and Ciordia, M. (2020) *Brenneria goodwinii* and *Gibbsiella quercinecans* isolated from weeping cankers on *Quercus robur* L. in Spain. *European Journal of Plant Pathology*. [online]. 156 (3), European Journal of Plant Pathology, pp.965–969.

- Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P. and Tiedje, J.M. (2007) DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *International Journal of Systematic and Evolutionary Microbiology*. [online]. 57 (1), pp.81–91.
- Gottel, N.R. Castro, H.F., Kerley, M., Yang, Z., Pelletier, D.A and Podar, M. (2011) Distinct microbial communities within the endosphere and rhizosphere of *Populus deltoides* roots across contrasting soil types. *Applied and Environmental Microbiology*. [online]. 77 (17), pp.5934–5944.
- Grace, E.R., Rabiey, M., Friman, V.P. and Jackson, R.W. (2021) Seeing the forest for the trees: Use of phages to treat bacterial tree diseases. *Plant Pathology*. [online]. 70 (9), pp.1987–2004.
- Gray, M.W. (2012) Mitochondrial evolution. *Cold Spring Harbor Perspectives in Biology*. [online]. 4 (9), pp.1476–1482.
- Griffiths, R.I., Whiteley, A.S., O’Donnell, A.G. and Bailey, M.J. (2003) Physiological and Community Responses of Established Grassland Bacterial Populations to Water Stress. *Applied and Environmental Microbiology*. [online]. 69 (12), pp.6961–6968.
- Grundmann, L.G. and Gourbière, F. (1999) A micro-sampling approach to improve the inventory of bacterial diversity in soil. *Applied Soil Ecology*. [online]. 13 (2), pp.123–126.
- Guasp, C., Moore, E.R.B., Lalucat, J. and Bennisar, A. (2000) Utility of internally transcribed 16S-23S rDNA spacer regions for the definition of *Pseudomonas stutzeri* genomovars and other *Pseudomonas* species. *International Journal of Systematic and Evolutionary Microbiology*. [online]. 50 (4), pp.1629–1639.
- Guindon, S. and Gascuel, O. (2003) A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Systematic Biology*. [online]. 52 (5), pp.696–704.
- Habermann, M. and Preller, J. (2003) Studies on the biology and control of two-spotted lichen buprestid (*Agrilus biguttatus* Fabr.). *Forst und Holz*. 58 (8), pp.215–220.
- Haichar, F. el Z., Santaella, C., Heulin, T. and Achouak, W. (2014) Root exudates mediated interactions belowground. *Soil Biology and Biochemistry*. [online]. 77, Elsevier Ltd, pp.69–80.

Available from: <http://dx.doi.org/10.1016/j.soilbio.2014.06.017>.

Haichar, F.E.Z., Marol, C., Berge, O., Rangel-Castro, J.I., Prosser, J.I., Balesdent, J., Heulin, T. and Achouak, W. (2008) Plant host habitat and root exudates shape soil bacterial community structure. *International Society for Microbial Ecology Journal*. [online]. 2 (12), pp.1221–1230.

Hallman, J. (2001) Plant Interactions with Endophytic Bacteria In: *Biotic interactions in plant-pathogen associations*. Bonn.

Hamaki, T., Suzuki, M., Fudou, R., Jojima, Y., Kajiura, T., Tabuchi, A., Sen, K. and Shibai, H. (2005) Isolation of novel bacteria and actinomycetes using soil-extract agar medium. *Journal of Bioscience and Bioengineering*. [online]. 99 (5), pp.485–492.

Hammond-Kosack, K. and Jones, J. (1997) Plant disease resistance genes. *Annual review of plant biology*. 48 (1), pp.575–607.

Hammond-Kosack, K.E. and Jones, J.D.G. (1996) Resistance gene-dependent plant defense responses. *Plant Cell*. [online]. 8 (10), pp.1773–1791.

Hanage, W.P., Fraser, C. and Spratt, B.G. (2005) Fuzzy species among recombinogenic bacteria. *BMC Biology*. [online]. 3, pp.1–7.

Hanson, C.A., Fuhrman, J.A., Horner-Devine, M.C. and Martiny, J.B.H. (2012) Beyond biogeographic patterns: Processes shaping the microbial landscape. *Nature Reviews Microbiology*. [online]. 10 (7), Nature Publishing Group, pp.497–506.

Hartmann, A., Rothballer, M. and Schmid, M. (2008) Lorenz Hiltner, a pioneer in rhizosphere microbial ecology and soil bacteriology research. *Plant and Soil*. [online]. 312 (1–2), pp.7–14.

He, S.Y., Huang, H.C. and Collmer, A. (1993) *Pseudomonas syringae* pv. *syringae* harpin Pss: A protein that is secreted via the hrp pathway and elicits the hypersensitive response in plants. *Cell*. [online]. 73 (7), pp.1255–1266.

Hedlund, B.P. Chuvochina, M., Hugenholtz, P., Konstantinidis, K.T., Murray, A.E and Palmer, M. (2022) SeqCode: a nomenclatural code for prokaryotes described from sequence data. *Nature Microbiology*. [online]. 7 (October), Springer US.

Hildebrand, M., Tebbe, C.C. and Geider, K. (2001) Survival studies with the fire blight pathogen *Erwinia amylovora* in soil and in a soil-inhabiting insect. *Journal of Phytopathology*. [online].

149 (11–12), pp.635–639.

Hiltner, L. (1904) Über neuere Erfahrungen und Probleme auf dem Gebiet der Bodenbakteriologie und unter besonderer Berücksichtigung der Grundungung and Brache. *Arbeiten der Deutschen Landwirtschaftlichen Gesellschaft*. 98, pp.59–78.

Hinsinger, P., Bengough, A.G., Vetterlein, D. and Young, I.M. (2009) Rhizosphere: Biophysics, biogeochemistry and ecological relevance. *Plant and Soil*. [online]. 321 (1–2), pp.117–152.

Hjort, K., Presti, I., Elväng, A., Marinelli, F. and Sjöling, S. (2014) Bacterial chitinase with phytopathogen control capacity from suppressive soil revealed by functional metagenomics. *Applied Microbiology and Biotechnology*. [online]. 98 (6), pp.2819–2828.

Hol, W.H.G., Garbeva, P., Hordijk, C., Hundscheid, M.P.J., Klein Gunnewiek, P.J.A., Van Agtmaal, M., Kuramae, E.E. and De Boer, W. (2015) Non-random species loss in bacterial communities reduces antifungal volatile production. *Ecology*. [online]. 96 (8), pp.2042–2048.

Hongxia, M., Jingfeng, F., Jiwen, L., Su jie., Zhiyi, W and Yantao, W. (2021) Full-length 16S rRNA gene sequencing reveals spatiotemporal dynamics of bacterial community in a heavily polluted estuary, China. *Environmental Pollution*. [online]. 275, Elsevier Ltd, p.116567. Available from: <https://doi.org/10.1016/j.envpol.2021.116567>.

Hu, Y., Li, Y., Yang, X., Li, C., Wang, L., Feng, J., Chen, S., Li, X. and Yang, Y. (2021) Effects of integrated biocontrol on bacterial wilt and rhizosphere bacterial community of tobacco. *Scientific Reports*. [online]. 11 (1), Nature Publishing Group UK, pp.1–11. Available from: <https://doi.org/10.1038/s41598-021-82060-3>.

Huang, J., Hammerbacher, A., Gershenzon, J., van Dam, N.M., Sala, A and McDowell, N.G. (2021) Storage of carbon reserves in spruce trees is prioritized over growth in the face of carbon limitation. *Proceedings of the National Academy of Sciences of the United States of America*. [online]. 118 (33), pp.1–7.

Huang, J. (1986) Ultrastructure of bacterial penetration in plants. *Annual Review of Phytopathology*. [online]. 24, pp.141–157.

Huang, L., Dong, B.-C., Xue, W., Peng, Y.-K. and Zhang, M.-X. (2013) Soil Particle Heterogeneity Affects the Growth of a Rhizomatous Wetland Plant. *Public Library of Science ONE*. [online].

8 (7), p.69836. Available from: [www.plosone.org](http://www.plosone.org).

Hugenholtz, P., Chuvochina, M., Oren, A., Parks, D.H. and Soo, R.M. (2021) Prokaryotic taxonomy and nomenclature in the age of big sequence data. *International Society for Microbial Ecology Journal*. [online]. 15 (7), Springer US, pp.1879–1892. Available from: <http://dx.doi.org/10.1038/s41396-021-00941-x>.

Huson, D.H., Beier, S., Flade, I., Górska, A., El-Hadidi, M. and Mitra, S. (2016) MEGAN Community Edition-Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *Public Library of Science Comput Biology*. [online]. 12 (6), p.1004957. Available from: <https://github.com/danielhuson/megan-ce>.

Ibrahim, Y.E., Saleh, A.A. and Al-Saleh, M.A. (2017) Management of asiatic citrus canker under field conditions in Saudi Arabia using bacteriophages and acibenzolar-s-methyl. *Plant Disease*. [online]. 101 (5), pp.761–765.

Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T. and Aluru, S. (2018) High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature Communications*. [online]. 9 (1), pp.1–8.

James, E.K. and Olivares, F.L. (2010) Critical Reviews in Plant Sciences Infection and Colonization of Sugar Cane and Other Graminaceous Plants by Endophytic Diazotrophs Infection and Colonization of Sugar Cane and Other Graminaceous Plants by Endophytic Diazotrophs. *Critical Reviews in Plant Sciences*. (February 2012), pp.37–41.

Janda, J.M. and Abbott, S.L. (2021) The Changing Face of the Family *Enterobacteriaceae* Order : Syndromes. *Clinical Microbiology Reviews*. 34 (2), pp.1–45.

Jeong, J. Yun, K., Mun, S., Chung, W.H., Choi, S.Y and Nam Y, d.o. (2021) The effect of taxonomic classification by full-length 16S rRNA sequencing with a synthetic long-read technology. *Scientific Reports*. [online]. 11 (1), Nature Publishing Group UK, pp.1–12. Available from: <https://doi.org/10.1038/s41598-020-80826-9>.

Johnson, J.S. Spakowicz, D.J., Hong, B.Y., Petersen, L.M., Demkowicz, P and Chen., L. (2019) Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nature Communications*. [online]. 10 (1), Springer US, pp.1–11. Available from: <http://dx.doi.org/10.1038/s41467-019-13036-1>.

- Jones, J.B., Jackson, L.E., Balogh, B., Obradovic, A., Iriarte, F.B. and Momol, M.T. (2008) Bacteriophages for plant disease control. *Annual Review of Phytopathology*. [online]. 45 (February), pp.245–262.
- Jones, J.D.G. and Dangl, J.L. (2006) The Plant Immune System. *Nature*. 444 (7117), pp.323–329.
- Jousset, A., Bienhold, C., Chatzinotas, A., Gallien, L., Gobet, A and Kurm, V. (2017) Where less may be more: How the rare biosphere pulls ecosystems strings. *International Society for Microbial Ecology Journal*. [online]. 11 (4), pp.853–862.
- Kaczmarek, M., Mullett, M.S., McDonald, J.E. and Denman, S. (2017) Multilocus sequence typing provides insights into the population structure and evolutionary potential of *Brenneria goodwinii*, associated with acute oak decline. *Public Library of Science ONE*. [online]. 12 (6), pp.1–19.
- Kämpfer, P. (2015a) *Buttiauxella*. *Bergey's Manual of Systematics of Archaea and Bacteria*. [online]. pp.1–16.
- Kämpfer, P. (2015b) *Rahnella*. *Bergey's Manual of Systematics of Archaea and Bacteria*. [online]. pp.1–16.
- Kämpfer, P., Meyer, S. and Müller, H.E. (1997) Characterization of *Buttiauxella* and *Kluyvera* species by analysis of whole cell fatty acid patterns. *Systematic and Applied Microbiology*. [online]. 20 (4), pp.566–571.
- Kandel, S., Joubert, P. and Doty, S. (2017) Bacterial Endophyte Colonization and Distribution within Plants. *Microorganisms*. [online]. 5 (4), p.77.
- Kang, S.M., Shahzad, R., Khan, M.A., Hasnain, Z., Lee, K.E., Park, H.S., Kim, L.R. and Lee, I.J. (2021) Ameliorative effect of indole-3-acetic acid- and siderophore-producing *Leclercia adecarboxylata* MO1 on cucumber plants under zinc stress. *Journal of Plant Interactions*. [online]. 16 (1), pp.30–41.
- Keren, Y., Keshet, D., Eidelman, M., Geffen, Y., Raz-Pasteur, A. and Hussein, K. (2014) Is *Leclercia adecarboxylata* a new and unfamiliar marine pathogen? *Journal of Clinical Microbiology*. [online]. 52 (5), pp.1775–1776.

- Kile, H., Arnold, D., Allainguillaume, J., Denman, S. and Brady, C. (2022) *Brenneria tiliae* sp. nov., isolated from symptomatic *Tilia × moltkei* and *Tilia × europaea* trees in the UK. *International Journal of Systematic and Evolutionary Microbiology*. [online]. 72 (10), pp.1–11.
- Kinsella, K., Schulthess, C.P., Morris, T.F. and Stuart, J.D. (2009) Rapid quantification of *Bacillus subtilis* antibiotics in the rhizosphere. *Soil Biology and Biochemistry*. [online]. 41 (2), Elsevier Ltd, pp.374–379. Available from: <http://dx.doi.org/10.1016/j.soilbio.2008.11.019>.
- Kirk, J.L., Beaudette, L.A., Hart, M., Moutoglis, P., Klironomos, J.N., Lee, H. and Trevors, J.T. (2004) Methods of studying soil microbial diversity. *Journal of Microbiological Methods*. [online]. 58 (2), pp.169–188.
- Koch, R. (1884) Die Aetiologie der Tuberkulose. *Mitt Kaiser Gesundh.* 2, pp.1–88.
- Konstantinidis, K.T., Rosselló-Móra, R. and Amann, R. (2017) Uncultivated microbes in need of their own taxonomy. *International Society for Microbial Ecology Journal*. [online]. 11 (11), Nature Publishing Group, pp.2399–2406. Available from: <http://dx.doi.org/10.1038/ismej.2017.113>.
- Konstantinidis, K.T. and Tiedje, J.M. (2005a) Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Sciences of the United States of America*. [online]. 102 (7), pp.2567–2572.
- Konstantinidis, K.T. and Tiedje, J.M. (2005b) Towards a genome-based taxonomy for prokaryotes. *Journal of Bacteriology*. [online]. 187 (18), pp.6258–6264.
- Kumar, A. and Dubey, A. (2020) Rhizosphere microbiome: Engineering bacterial competitiveness for enhancing crop production. *Journal of Advanced Research*. [online]. 24, Cairo University, pp.337–352. Available from: <https://doi.org/10.1016/j.jare.2020.04.014>.
- Kumar, P., Dubey, R.C. and Maheshwari, D.K. (2012) *Bacillus* strains isolated from rhizosphere showed plant growth promoting and antagonistic activity against phytopathogens. *Microbiological Research*. [online]. 167 (8), Elsevier GmbH, pp.493–499. Available from: <http://dx.doi.org/10.1016/j.micres.2012.05.002>.
- Kumawat, K.C., Sharma, P., Singh, I., Sirari, A. and Gill, B.S. (2019) Co-existence of *Leclercia adecarboxylata* (LSE-1) and *Bradyrhizobium* sp. (LSBR-3) in nodule niche for multifaceted

effects and profitability in soybean production. *World Journal of Microbiology and Biotechnology*. [online]. 35 (11), Springer Netherlands, pp.1–17. Available from: <https://doi.org/10.1007/s11274-019-2752-4>.

Kushalappa, A.C. and Zulfiqar, M. (2001) Effect of wet incubation time and temperature on infection, and of storage time and temperature on soft rot lesion expansion in potatoes inoculated with *Erwinia carotovora* ssp. *carotovora*. *Potato Research*. [online]. 44 (3), pp.233–242.

Lam, T.Y.C., Mei, R., Wu, Z., Lee, P.K.H., Liu, W.T. and Lee, P.H. (2020) Superior resolution characterisation of microbial diversity in anaerobic digesters using full-length 16S rRNA gene amplicon sequencing. *Water Research*. [online]. 178, Elsevier Ltd, p.115815. Available from: <https://doi.org/10.1016/j.watres.2020.115815>.

Langendries, S. and Goormachtig, S. (2021) *Paenibacillus polymyxa*, a Jack of all trades. *Environmental Microbiology*. [online]. 23 (10), pp.5659–5669.

Langille, M.G.I., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D and Reyes, J.A. (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology*. [online]. 31 (9), Nature Publishing Group, pp.814–821.

Lauber, C.L., Zhou, N., Gordon, J.I., Knight, R. and Fierer, N. (2010) Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples. *Federation of European Microbiological Societies Microbiology Letters*. [online]. 307 (1), pp.80–86.

Lefort, V., Desper, R. and Gascuel, O. (2015) FastME 2.0: A comprehensive, accurate, and fast distance-based phylogeny inference program. *Molecular Biology and Evolution*. [online]. 32 (10), pp.2798–2800.

Lefort, V., Longueville, J.E. and Gascuel, O. (2017) SMS: Smart Model Selection in PhyML. *Molecular biology and evolution*. [online]. 34 (9), pp.2422–2424.

Liang, L. and Fei, S. (2014) Divergence of the potential invasion range of emerald ash borer and its host distribution in North America under climate change. *Climatic Change*. [online]. 122 (4), pp.735–746.

- Liang, Y., Tóth, K., Cao, Y., Tanaka, K., Espinoza, C. and Stacey, G. (2014) Lipochitooligosaccharide recognition: An ancient story. *Journal of Physiology*. [online]. 204 (2), pp.289–296.
- Liebeke, M., Brözel, V.S., Hecker, M. and Lalk, M. (2009) *Chemical characterization of soil extract as growth media for the ecophysiological study of bacteria Applied Microbiology and Biotechnology*. [online]. 83 (1), 161–173.
- Liesche, J. and Patrick, J. (2017) An update on phloem transport: a simple bulk flow under complex regulation. *F1000Research*. [online]. 6, p.2096.
- Lindahl, V. (1995) Nucleic Acids in the Environment. *Nucleic Acids in the Environment*. [online]. (January 1995).
- Lindahl, V. and Bakken, L.R. (1995) Evaluation of methods for extraction of bacteria from soil. *Federation of European Microbiological Societies Microbiology Ecology*. [online]. 16 (2), pp.135–142.
- Liu, Z., Zhou, S., Zhang, W. and Wu, S. (2021) First Report of *Cedecea neteri* Causing Yellow Rot Disease in *Pleurotus pulmonarius* in China. *Plant Disease*. (April), p.1189.
- Long, T. and Or, D. (2005) Aquatic habitats and diffusion constraints affecting microbial coexistence in unsaturated porous media. *Water Resources Research*. [online]. 41 (8), pp.1–10.
- Lugtenberg, B. and Kamilova, F. (2009) Plant-Growth-Promoting Rhizobacteria. *Annual Review of Microbiology*. [online]. 63 (1), Annual Reviews, pp.541–556.
- Lugtenberg, B.J., Kravchenko, L. V. and Simons, M. (1999) Tomato seed and root exudate sugars: composition, utilization by *Pseudomonas* biocontrol strains and role in rhizosphere colonization. *Environmental microbiology*. [online]. 1 (5), pp.439–446.
- Lugtenberg, B.J.J., Dekkers, L. and Bloemberg, G. V. (2001) Molecular Determinants of Rhizosphere Colonization By *Pseudomonas*. *Annual Review of Phytopathology*. [online]. 39 (1), pp.461–490.
- Maes, M., Baeyen, S., De Croo, H., De Smet, K. and Steenackers, M. (2017) Monitoring of endophytic *Brenneria salicis* in willow and its relation to watermark disease. *Plant Protection*

*Science*. [online]. 38 (SI 2-6th Conf EFPP 2002), pp.528–530.

Maghnia, F.Z., Abbas, Y., Mahé, F., Prin, Y., El Ghachtouli, N., Duponnois, R. and Sanguin, H. (2019) The rhizosphere microbiome: A key component of sustainable cork oak forests in trouble. *Forest Ecology and Management*. [online]. 434 (December 2018), Elsevier, pp.29–39. Available from: <https://doi.org/10.1016/j.foreco.2018.12.002>.

Manion, P.. (1981) *Tree disease concepts*. Englewood Cliffs, N.J. (ed.) [online]. Prentice-Hall. Available from: <https://archive.org/details/treediseaseconce0000mani/page/n19/mode/2up>.

Mansfield, J.W. (2009) From bacterial avirulence genes to effector functions via the hrp delivery system: An overview of 25 years of progress in our understanding of plant innate immunity. *Molecular Plant Pathology*. [online]. 10 (6), pp.721–734.

Marathe, N.P., Salvà-Serra, F., Karlsson, R., Larsson, D.G.J., Moore, E.R.B., Svensson-Stadler, L. and Jakobsson, H.E. (2019) *Scandinavium goeteborgense* gen. nov., sp. nov., a New Member of the Family *Enterobacteriaceae* Isolated From a Wound Infection, Carries a Novel Quinolone Resistance Gene Variant. *Frontiers in Microbiology*. [online]. 10 (November), pp.1–13.

Mark, G.L. Dow, J.M., Kiely, P.D., Higgins, H., Haynes, J and Baysse, C. (2005) Transcriptome profiling of bacterial responses to root exudates identifies genes involved in microbe-plant interactions. *Proceedings of the National Academy of Sciences of the United States of America*. [online]. 102 (48), pp.17454–17459.

Martin, N.J. and Macdonald, R.M. (1981) Separation of Non-filamentous Micro-organisms from Soil by Density Gradient Centrifugation in Percoll. *Journal of Applied Bacteriology*. [online]. 51 (2), pp.243–251.

Martínez-Granero, F., Rivilla, R. and Martín, M. (2006) Rhizosphere selection of highly motile phenotypic variants of *Pseudomonas fluorescens* with enhanced competitive colonization ability. *Applied and Environmental Microbiology*. [online]. 72 (5), pp.3429–3434.

Martiny, J.B.H., Eisen, J.A., Penn, K., Allison, S.D. and Horner-Devine, M.C. (2011) Drivers of bacterial  $\beta$ -diversity depend on spatial scale. *Proceedings of the National Academy of Sciences of the United States of America*. [online]. 108 (19), pp.7850–7854.

- McManus, P.S. and Stockwell, V.O. (2001) Antibiotic Use for Plant Disease Management in the United States. *Plant Health Progress*. [online]. 2 (1).
- McMurdie, P.J. and Holmes, S. (2013) Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *Public Library of Science ONE*. [online]. 8 (4).
- McNear, D.H. (2013) The Rhizosphere - Roots, Soil and Everything In Between Meeting the Global Challenge of Sustainable Food, Fuel and Fiber Production. *Nature Education Knowledge*. 4 (3), pp.1–15.
- Meaden, S., Metcalf, C.J.E. and Koskella, B. (2016) The effects of host age and spatial location on bacterial community composition in the English oak tree (*Quercus robur*). *Environmental Microbiology Reports*. [online]. 8 (5), pp.649–658.
- Meier-Kolthoff, J.P., Auch, A.F., Klenk, H.P. and Göker, M. (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*. [online]. 14.
- Merbach, W., Mirus, E., Knof, G., Remus, R., Ruppel, S., Russow, R., Gransee, A. and Schulze, J. (1999) Release of carbon and nitrogen compounds by plant roots and their possible ecological importance. *Journal of Plant Nutrition and Soil Science*. [online]. 162 (4), pp.373–383.
- Mikha, M.M., Rice, C.W. and Milliken, G.A. (2005) Carbon and nitrogen mineralization as affected by drying and wetting cycles. *Soil Biology and Biochemistry*. [online]. 37 (2), pp.339–347.
- Monchgesang, S., Strehmel, N., Schmidt, S., Westphal, L., Taruttis, F., Muller, E., Herklotz, S., Neumann, S. and Scheel, D. (2016) Natural variation of root exudates in *Arabidopsis thaliana*-linking metabolomic and genomic data. *Scientific Reports*. [online]. 6 (February), Nature Publishing Group, pp.1–11. Available from: <http://dx.doi.org/10.1038/srep29033>.
- Monteiro, R.A. Balsanelli, E., Wasseem, R., Marin, A.M., Brusamarello-Santos, L.C.C and Schmidt, M.A. (2012) Herbaspirillum-plant interactions: Microscopical, histological and molecular aspects. *Plant and Soil*. [online]. 356 (1–2), pp.175–196.

Moradi-Amirabad, Y., Rahimian, H., Babaeizad, V. and Denman, S. (2019) *Brenneria* spp. and *Rahnella victoriana* associated with acute oak decline symptoms on oak and hornbeam in Iran. *Forest Pathology*. [online]. 49 (4), pp.1–14.

Morris, C.E., Monteil, C.L. and Berge, O. (2013) The life history of *pseudomonas syringae*: Linking agriculture to earth system processes. *Annual Review of Phytopathology*. [online]. 51, pp.85–104.

Mukherjee, S., Kuang, Z., Ghosh, S., Detroja, R., Carmi, G and Tripathy, S. (2022) Incipient Sympatric Speciation and Evolution of Soil Bacteria Revealed by Metagenomic and Structured Non-Coding RNAs Analysis. *Biology*. [online]. Available from: <https://doi.org/10.3390/biology11081110>.

Nagy, A., Vitásková, E., Černíková, L., Křivda, V., Jiřincová, H., Sedlák, K., Horníčková, J. and Havlíčková, M. (2017) Evaluation of TaqMan qPCR system integrating two identically labelled hydrolysis probes in single assay. *Scientific Reports*. [online]. 7, pp.1–10.

Naum, M., Brown, E.W. and Mason-Gamer, R.J. (2008) Is 16S rDNA a reliable phylogenetic marker to characterize relationships below the family level in the *enterobacteriaceae*? *Journal of Molecular Evolution*. [online]. 66 (6), pp.630–642.

Nelson, S.O. (2015) Assessment of Soil Treatment for Pest Control. *Dielectric Properties of Agricultural Materials and their Applications*. [online]. Academic Press, pp.109–122.

Nguyen, T.M., Seo, C., Ji, M., Paik, M.-J., Myung, S.-W. and Kim, J. (2018) Effective Soil Extraction Method for Cultivating Previously Uncultured Soil Bacteria. *Applied and Environmental Microbiology*. 84 (24), pp.1–14.

Nichols, D., Cahoon, N., Trakhtenberg, E.M., Pham, L., Mehta, A., Belanger, A., Kanigan, T., Lewis, K. and Epstein, S.S. (2010) Use of ichip for high-throughput in situ cultivation of "uncultivable microbial species". *Applied and Environmental Microbiology*. [online]. 76 (8), pp.2445–2450.

Nishiguchi, M.K., Hirsch, A.M., Devinney, R., Vedantam, G., Riley, M.A. and Mansky, L.M. (2008) Deciphering evolutionary mechanisms between mutualistic and pathogenic symbioses. *Vie et Milieu*. 58 (2), pp.87–106.

Niu, C., Yu, D., Wang, Y., Ren, H., Jin, Y and Zhou, W. (2013) Common and pathogen-specific virulence factors are different in function and structure. *Virulence*. [online]. 4 (6), pp.473–482.

Nurk, S., Bankevich, A., Antipov, D., Gurevich, A., Korobeynikov, A and Lapidus, A. (2013) Assembling genomes and mini-metagenomes from highly chimeric reads. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. [online]. 7821 LNBI, pp.158–170.

O’Leary, N.A., Wright, M.W., Brister, J.R., Ciufu, S., Haddad, D and McVeigh, R. (2016) Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*. [online]. 44 (D1), pp.D733–D745.

Okonechnikov, K., Golosova, O., Fursov, M., Varlamov, A., Vaskin, Y and Efremov, I. (2012) Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics*. [online]. 28 (8), pp.1166–1167.

Olanrewaju, O.S., Ayangbenro, A.S., Glick, B.R. and Babalola, O.O. (2019) Plant health: feedback effect of root exudates-rhizobiome interactions. *Applied Microbiology and Biotechnology*. [online]. 103 (3), Applied Microbiology and Biotechnology, pp.1155–1166.

Ondov, B.D., Bergman, N.H. and Phillippy, A.M. (2011) Interactive metagenomic visualization in a Web browser. *BioMed Central Bioinformatics*. [online]. 12 (1), BioMed Central Ltd, p.385. Available from: <http://www.biomedcentral.com/1471-2105/12/385>.

Osbourn, A.E. (1996) Preformed antimicrobial compounds and plant defense against fungal attack. *Plant Cell*. [online]. 8 (10), pp.1821–1831.

Ostry, M.E., Venette, R.C. and Juzwik, J. (2011) Decline as a disease category: Is it helpful. *Phytopathology*. [online]. 101 (4), pp.404–409.

Overmann, J., Huang, S., Nübel, U., Hahnke, R.L. and Tindall, B.J. (2019) Relevance of phenotypic information for the taxonomy of not-yet-cultured microorganisms. *Systematic and Applied Microbiology*. [online]. 42 (1), Elsevier GmbH., pp.22–29. Available from: <https://doi.org/10.1016/j.syapm.2018.08.009>.

Palmer, M., Steenkamp, E.T., Blom, J., Hedlund, B.P. and Venter, S.N. (2020) All anis are not

created equal: Implications for prokaryotic species boundaries and integration of anis into polyphasic taxonomy. *International Journal of Systematic and Evolutionary Microbiology*. [online]. 70 (4), pp.2937–2948.

Paradis, S., Boissinot, M., Paquette, N., Bélanger, S.D., Martel, E.A and Boudreau, D.K. (2005) Phylogeny of the *Enterobacteriaceae* based on genes encoding elongation factor Tu and F-ATPase  $\beta$ -subunit. *International Journal of Systematic and Evolutionary Microbiology*. [online]. 55 (5), pp.2013–2025.

Patel, J.K. and Archana, G. (2017) Diverse culturable diazotrophic endophytic bacteria from *Poaceae* plants show cross-colonization and plant growth promotion in wheat. *Plant and Soil*. [online]. 417 (1–2), pp.99–116.

Patra, N., Prakash, M.R., Patil, S. and Rao, M.B. (2018) First Case Report of Surgical Site Infection Due to *Buttiauxella agrestis* in a Neurocare Center in India. [online]. pp.117–118.

Pausch, J. and Kuzyakov, Y. (2018) Carbon input by roots into the soil: Quantification of rhizodeposition from root to ecosystem scale. *Global Change Biology*. [online]. 24 (1), pp.1–12.

Pereira, S.I.A., Monteiro, C., Vega, A.L. and Castro, P.M.L. (2016) Endophytic culturable bacteria colonizing *Lavandula dentata* L. plants: Isolation, characterization and evaluation of their plant growth-promoting activities. *Ecological Engineering*. [online]. 87, Elsevier B.V., pp.91–97. Available from: <http://dx.doi.org/10.1016/j.ecoleng.2015.11.033>.

Perombelon, M.C.M. and Hyman, L.J. (1989) Survival of soft rot coliforms, *Erwinia carotovora* subsp. *carotovora* and *E. carotovora* subsp. *atroseptica* in soil in Scotland. *Journal of Applied Bacteriology*. 66 (2), pp.95–106.

Pertile, M., Sousa, R.M.S., Mendes, L.W., Antunes, J.E.L., Oliveira, L.M. de S., de Araujo, F.F., Melo, V.M.M. and Araujo, A.S.F. (2021) Response of soil bacterial communities to the application of the herbicides imazethapyr and flumyazin. *European Journal of Soil Biology*. [online]. 102 (November 2020), Elsevier Masson SAS, p.103252. Available from: <https://doi.org/10.1016/j.ejsobi.2020.103252>.

Pettifor, B.J., Doonan, J., Denman, S. and McDonald, J.E. (2020) Survival of *Brenneria goodwinii* and *Gibbsiella quercinecans*, Associated with Acute Oak Decline, in Rainwater and

Forest Soil. *Systematic and Applied Microbiology*. Elsevier GmbH., pp.9–10.

Pii, Y., Mimmo, T., Tomasi, N., Terzano, R., Cesco, S. and Crecchio, C. (2015) Microbial interactions in the rhizosphere: beneficial influences of plant growth-promoting rhizobacteria on nutrient acquisition process. A review. *Biology and Fertility of Soils*. [online]. 51 (4), pp.403–415.

Pinho, D., Barroso, C., Froufe, H., Brown, N., Vanguelova, E., Egas, C. and Denman, S. (2020) Linking tree health, rhizosphere physicochemical properties, and microbiome in acute oak decline. *Forests*. [online]. 11 (11), pp.1–19.

Popescu, L. and Cao, Z.P. (2018) From microscopy to genomic approach in soil biodiversity assessment. *Current Issues in Molecular Biology*. [online]. 27, pp.195–198. Available from: <https://dx.doi.org/10.21775/cimb.027.195>.

Portillo, M.C., Leff, J.W., Lauber, C.L. and Fierer, N. (2013) Cell size distributions of soil bacterial and archaeal taxa. *Applied and Environmental Microbiology*. [online]. 79 (24), pp.7610–7617.

Poza-Carrión, C., Aguilar, I., Gallego, F.J., Nuñez-Moreno, Y., Biosca, E.G., González, R., López, M.M. and Rodríguez-Palenzuela, P. (2008) *Brenneria quercina* and *Serratia* spp. isolated from Spanish oak trees: Molecular characterization and development of PCR primers. *Plant Pathology*. [online]. 57 (2), pp.308–319.

Prabhu, D., Vidhyavathi, R. and Jeyakanthan, J. (2017) Computational identification of potent inhibitors for Streptomycin 3'-adenyltransferase of *Serratia marcescens*. *Microbial Pathogenesis*. [online]. 103, pp.94–106.

Pritchard, L., Glover, R.H., Humphris, S., Elphinstone, J.G. and Toth, I.K. (2016) Genomics and taxonomy in diagnostics for food security: Soft-rotting enterobacterial plant pathogens. *Analytical Methods*. [online]. 8 (1), pp.12–24.

Prosser, J.I. (2015) Dispersing misconceptions and identifying opportunities for the use of 'omics' in soil microbial ecology. *Nature Reviews Microbiology*. [online]. 13 (7), Nature Publishing Group, pp.439–446. Available from: <http://dx.doi.org/10.1038/nrmicro3468>.

Purcell, A.H. (2009) Plant Diseases and Insects. *Encyclopedia of Insects*. [online]. Academic

Press, pp.802–806.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckner, F.O. (2013) The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*. [online]. 41 (D1), pp.590–596.

Raaijmakers, J.M., Paulitz, T.C., Steinberg, C., Alabouvette, C. and Moëgne-Loccoz, Y. (2009) The rhizosphere: A playground and battlefield for soilborne pathogens and beneficial microorganisms. *Plant and Soil*. [online]. 321 (1–2), pp.341–361.

Raaijmakers, J.M., Vlami, M. and de Souza, J.T. (2002) Antibiotic production by bacterial biocontrol agents. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*. [online]. 81 (1–4), pp.537–547.

Rahman, M.M., Flory, E., Koyro, H.W., Abideen, Z., Schikora, A., Suarez, C., Schnell, S. and Cardinale, M. (2018) Consistent associations with beneficial bacteria in the seed endosphere of barley (*Hordeum vulgare* L.). *Systematic and Applied Microbiology*. [online]. 41 (4), Elsevier GmbH, pp.386–398. Available from: <https://doi.org/10.1016/j.syapm.2018.02.003>.

Rahn, O. (1937) New Principles for the Classification of Bacteria. *Zentralblatt fur Bakteriologie, Parasitenkunde, Infektionskrankheiten und Hygiene*. 96, pp.273–286.

Raina, V., Nayak, T., Ray, L., Kumari, K. and Suar, M. (2019) A Polyphasic Taxonomic Approach for Designation and Description of Novel Microbial Species. *Microbial Diversity in the Genomic Era*. [online]. Academic Press, pp.137–152.

Ramírez-Bahena, M.H., Peix, A., Rivas, R., Camacho, M., Rodríguez-Navarro, D.N., Mateos, P.F., Martínez-Molina, E., Willems, A. and Velázquez, E. (2009) *Bradyrhizobium pachyrhizi* sp. nov. and *Bradyrhizobium jicamae* sp. nov., isolated from effective nodules of *Pachyrhizus erosus*. *International Journal of Systematic and Evolutionary Microbiology*. [online]. 59 (8), pp.1929–1934.

Reed, K., Denman, S., Leather, S.R., Forster, J. and Inward, D.J.G. (2018) The lifecycle of *Agrilus biguttatus*: the role of temperature in its development and distribution, and implications for Acute Oak Decline. *Agricultural and Forest Entomology*. [online]. 20 (3), pp.334–346.

Rhoads, A. and Au, K.F. (2015) PacBio Sequencing and Its Applications. *Genomics, Proteomics*

*and Bioinformatics*. [online]. 13 (5), Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China, pp.278–289. Available from: <http://dx.doi.org/10.1016/j.gpb.2015.08.002>.

Richter-Heitmann, T., Eickhorst, T., Knauth, S., Friedrich, M.W. and Schmidt, H. (2016) Evaluation of strategies to separate root-associated microbial communities: A crucial choice in rhizobiome research. *Frontiers in Microbiology*. [online]. 7 (MAY), pp.1–11.

Richter, M., Rosselló-Móra, R., Oliver Glöckner, F. and Peplies, J. (2016) JSpeciesWS: A web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics*. [online]. 32 (6), pp.929–931.

Ritika, B. and Utpal, D. (2014) An overview of fungal and bacterial biopesticides to control plant pathogens/diseases. *African Journal of Microbiology Research*. [online]. 8 (17), pp.1749–1762.

Rochelle, P.A., Cragg, B.A., Fry, J.C., John Parkes, R. and Weightman, A.J. (1994) Effect of sample handling on estimation of bacterial diversity in marine sediments by 16S rRNA gene sequence analysis. *Federation of European Microbiological Societies Microbiology Ecology*. [online]. 15 (1–2), pp.215–225.

Rodriguez-R, L.M. and Konstantinidis, K.T. (2016) The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ Preprints*. 4, p.e1900v1.

Rong, X., Liu, N., Ruan, J. and Huang, Y. (2010) Multilocus sequence analysis of *Streptomyces griseus* isolates delineating intraspecific diversity in terms of both taxonomy and biosynthetic potential. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*. [online]. 98 (2), pp.237–248.

Rosselló-Móra, R. and Amann, R. (2015) Past and future species definitions for Bacteria and Archaea. *Systematic and Applied Microbiology*. [online]. 38 (4), Elsevier GmbH., pp.209–216. Available from: <http://dx.doi.org/10.1016/j.syapm.2015.02.001>.

Ruffner, B., Schneider, S., Meyer, J., Queloz, V. and Rigling, D. (2020) First report of acute oak decline disease of native and non-native oaks in Switzerland. *New Disease Reports*. [online]. 41 (1), pp.18–18.

- Saito, M., Shinozaki-Kuwahara, N. and Takada, K. (2012) *Gibbsiella dentisursi* sp. nov., isolated from the bear oral cavity. *Microbiology and Immunology*. [online]. 56 (8), pp.506–512.
- Sallé, A., Nageleisen, L.M. and Lieutier, F. (2014) Bark and wood boring insects involved in oak declines in Europe: Current knowledge and future prospects in a context of climate change. *Forest Ecology and Management*. [online]. 328, pp.79–93.
- Sandler, S.R. and Karo, W. (1992) SULFITES. *Sourcebook of Advanced Organic Laboratory Preparations*. [online]. Academic Press, pp.202–207.
- Santander, R.D., Català-Senent, J.F., Figàs-Segura, À. and Biosca, E.G. (2020) From the roots to the stem: Unveiling pear root colonization and infection pathways by *Erwinia amylovora*. *Federation of European Microbiological Societies Microbiology Ecology*. [online]. 96 (12), pp.1–12.
- Sapp, M., Lewis, E., Moss, S., Barrett, B., Kirk, S., Elphinstone, J.G. and Denman, S. (2016) Metabarcoding of bacteria associated with the acute oak decline syndrome in England. *Forests*. [online]. 7 (5).
- Sarkar, S.F. and Guttman, D.S. (2004) Evolution of the Core Genome of *Pseudomonas syringae*, a Highly Clonal, Endemic Plant Pathogen. *Society*. [online]. 70 (4), pp.1999–2012.
- Sasse, J., Martinoia, E. and Northen, T. (2018) Feed Your Friends: Do Plant Exudates Shape the Root Microbiome? *Trends in Plant Science*. [online]. 23 (1), Elsevier Ltd, pp.25–41. Available from: <http://dx.doi.org/10.1016/j.tplants.2017.09.003>.
- Saunders, D.S. (2002) *Insect Clocks* Steel, X.V.C.G.H. and Lewis, R.D. (eds.) 3rd edition. Amsterdam, Elsevier Science.
- Scarlett, K., Denman, S., Clark, D.R., Forster, J., Vanguelova, E., Brown, N. and Whitby, C. (2021) Relationships between nitrogen cycling microbial community abundance and composition reveal the indirect effect of soil pH on oak decline. *International Society for Microbial Ecology Journal*. [online]. 15 (3), Springer US, pp.623–635. Available from: <http://dx.doi.org/10.1038/s41396-020-00801-0>.
- Scheinert, P., Krausse, R., Ullmann, U., Söller, R. and Krupp, G. (1996) Molecular differentiation of bacteria by PCR amplification of the 6S-23S rRNA spacer. *Journal of*

*Microbiological Methods*. [online]. 26 (1–2), pp.103–117.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M and Hollister, E.B. (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*. [online]. 75 (23), pp.7537–7541.

Schulz, H.N. and Jørgensen, B.B. (2001) Big Bacteria. *Molecules*. 55, pp.105–137.

Schulz, S., Brankatschk, R., Dümig, A., Kögel-Knabner, I., Schloter, M. and Zeyer, J. (2013) The role of microorganisms at different stages of ecosystem development for soil formation. *Biogeosciences*. [online]. 10 (6), pp.3983–3996.

Seemann, T. (2014) Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*. [online]. 30 (14), pp.2068–2069.

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S. and Huttenhower, C. (2011) Metagenomic biomarker discovery and explanation. *Genome Biology*. 12, p.R60.

Semenov, M. V. (2021) Metabarcoding and Metagenomics in Soil Ecology Research: Achievements, Challenges, and Prospects. *Biology Bulletin Reviews*. [online]. 11 (1), pp.40–53.

Sharma, A., Gautam, S. and Wadhawan, S. (2014) *Xanthomonas*. *Encyclopedia of Food Microbiology: Second Edition*. [online]. Academic Press, pp.811–817.

Sharma, A., Sharma, N.K. and Kardile, H.B. (2021) Soil Borne Fungal Diseases and Their Control in Below Ground CropsIn: Kaushal, M. and Prasad, R. (eds.) *Microbial Biotechnology in Crop Protection*. Singapore, Springer, 251–268.

Sharma, S., Mehta, R., Gupta, R. and Schloter, M. (2012) Improved protocol for the extraction of bacterial mRNA from soils. *Journal of Microbiological Methods*. [online]. 91 (1), Elsevier B.V., pp.62–64. Available from: <http://dx.doi.org/10.1016/j.mimet.2012.07.016>.

Shigo, A.L. (1986) *A new tree biology dictionary* 3rd edition. Durham, N.H.

De Silva, N.I., Brooks, S., Lumyong, S. and Hyde, K.D. (2019) Use of endophytes as biocontrol agents. *Fungal Biology Reviews*. [online]. 33 (2), Elsevier Ltd, pp.133–148. Available from: <https://doi.org/10.1016/j.fbr.2018.10.001>.

- Simon, C. and Daniel, R. (2011) Metagenomic analyses: Past and future trends. *Applied and Environmental Microbiology*. [online]. 77 (4), pp.1153–1161.
- Sinclair, W.A. (1965) Comparisons of Recent Declines of White Ash, Oaks and Sugar Maple in Northeastern Woodlands. *The Cornell Plantations*. 20 (4), pp.62–67.
- Sinclair, W.A. (1967) Decline of Hardwoods: Possible Causes. *International Shade Tree Conference*. [online]. pp.17–31.
- Sinclair, W.A. and Hudler, G.W. (1988) Tree Delinence: Four Concepts of Causality. *Journal of Arboriculture*. [online]. 14 (2), pp.29–35.
- Sirichoat, A., Sankuntaw, N., Engchanil, C., Buppasiri, P., Faksri, K., Namwat, W., Chantratita, W. and Lulitanond, V. (2021) Comparison of different hypervariable regions of 16S rRNA for taxonomic profiling of vaginal microbiota using next-generation sequencing. *Archives of Microbiology*. [online]. 203 (3), Springer Berlin Heidelberg, pp.1159–1166. Available from: <https://doi.org/10.1007/s00203-020-02114-4>.
- Słomka, M., Sobalska-Kwapis, M., Wachulec, M., Bartosz, G. and Strapagiel, D. (2017) High resolution melting (HRM) for high-throughput genotyping-limitations and caveats in practical case studies. *International Journal of Molecular Sciences*. [online]. 18 (11).
- Sokolov, M.S., Semenov, A.M., Spiridonov, Y.Y., Toropova, E.Y. and Glinushkin, A.P. (2020) Healthy Soil—Condition for Sustainability and Development of the Argo- and Sociospheres (Problem-Analytical Review). *Biology Bulletin*. [online]. 47 (1), pp.18–26.
- Soosaar, J.L.M., Burch-Smith, T.M. and Dinesh-Kumar, S.P. (2005) Mechanisms of plant resistance to viruses. *Nature Reviews Microbiology*. [online]. 3 (10), pp.789–798.
- Soutar, C.D. and Stavrinides, J. (2020) Phylogenetic analysis supporting the taxonomic revision of eight genera within the bacterial order enterobacterales. *International Journal of Systematic and Evolutionary Microbiology*. [online]. 70 (12), pp.6524–6530.
- Steen, A.D., Crits-Christoph, A., Carini, P., DeAngelis, K.M., Fierer, N., Lloyd, K.G. and Cameron Thrash, J. (2019) High proportions of bacteria and archaea across most biomes remain uncultured. *International Society for Microbial Ecology Journal*. [online]. 13 (12), Springer US, pp.3126–3130. Available from: <http://dx.doi.org/10.1038/s41396-019-0484-y>.

Stevenson, J., Hymas, W. and Hillyard, D. (2005) Effect of sequence polymorphisms on performance of two real-time PCR assays for detection of herpes simplex virus. *Journal of Clinical Microbiology*. [online]. 43 (5), pp.2391–2398.

Sutcliffe, I.C., Dijkshoorn, L. and Whitman, W.B. (2020) Minutes of the international committee on systematics of prokaryotes online discussion on the proposed use of gene sequences as type for naming of prokaryotes, and outcome of vote. *International Journal of Systematic and Evolutionary Microbiology*. [online]. 70 (7), pp.4416–4417.

Sutton, R. and Sposito, G. (2005) Molecular structure in soil humic substances: The new view. *Environmental Science and Technology*. [online]. 39 (23), pp.9009–9015.

Taktikos, J., Stark, H. and Zaburdaev, V. (2013) How the motility pattern of bacteria affects their dispersal and chemotaxis. *Public Library of Science ONE*. [online]. 8 (12).

Tamura, K., Stecher, G. and Kumar, S. (2021) MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*. [online]. 38 (7), pp.3022–3027.

Tariq, M., Yasmin, S. and Hafeez, F.Y. (2010) Biological Control of Potato Black Scurf by Rhizosphere Associated Bacteria. *Brazilian Journal of Microbiology*. 41, pp.439–451.

Tatusova, T., Dicuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E.P and Zaslavsky, L. (2016) NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Research*. [online]. 44 (14), pp.6614–6624.

The Met office (2019) *Microclimates* [online]. Library, National Meteorological, 24. Available from: [www.metoffice.gov.uk/research/library-and-archive/publications/factsheets%0AMet](http://www.metoffice.gov.uk/research/library-and-archive/publications/factsheets%0AMet).

Thomas, F. and Hartmann, G. (1996) Soil and tree water relations in mature oak stands of Northern Germany differing in the degree of decline. *Annals of Forest Science*. [online]. 53 (January 1996), pp.697–720.

Thomas, F.M. (2008) Recent advances in cause-effect research on oak decline in Europe. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*. [online]. 3 (July 2008).

Thomas, F.M., Blank, R. and Hartmann, G. (2002) Abiotic and biotic factors and their interactions as causes of oak decline in Central Europe. *Forest Pathology*. [online]. 32 (4–5),

pp.277–307.

Thompson, D.K. and Sharkady, S.M. (2020) Expanding spectrum of opportunistic *Cedecea* infections: Current clinical status and multidrug resistance. *International Journal of Infectious Diseases*. [online]. 100, International Society for Infectious Diseases, pp.461–469. Available from: <https://doi.org/10.1016/j.ijid.2020.09.036>.

Thomson, B.C., Ostle, N.J., McNamara, N.P., Whiteley, A.S. and Griffiths, R.I. (2010) Effects of sieving, drying and rewetting upon soil bacterial community structure and respiration rates. *Journal of Microbiological Methods*. [online]. 83 (1), pp.69–73.

Tian, B., Zhang, C., Ye, Y., Wen, J., Wu, Y and Wang, H. (2017) Beneficial traits of bacterial endophytes belonging to the core communities of the tomato root microbiome. *Agriculture, Ecosystems and Environment*. [online]. 247 (2), Elsevier, pp.149–156. Available from: <http://dx.doi.org/10.1016/j.agee.2017.06.041>.

Tilman, D. (1994) Competition and Biodiversity in Spatially Structured Habitats Author ( s ): David Tilman Stable URL : <http://www.jstor.org/stable/1939377> 75 (1), pp.2–16.

Torsvik, V., Øvreås, L. and Thingstad, T.F. (2002) Prokaryotic diversity - Magnitude, dynamics, and controlling factors. *Science*. [online]. 296 (5570), pp.1064–1066.

Toth, I.K., Pritchard, L. and Birch, P.R.J. (2006) Comparative Genomics Reveals What Makes An Enterobacterial Plant Pathogen. *Annual Review of Phytopathology*. [online]. 44 (1), pp.305–336.

Totsche, K.U., Rennert, T., Gerzabek, M.H., Kögel-Knabner, I., Smalla, K., Spiteller, M. and Vogel, H.J. (2010) Biogeochemical interfaces in soil: The interdisciplinary challenge for soil science. *Journal of Plant Nutrition and Soil Science*. [online]. 173 (1), pp.88–99.

Turner, T.R., James, E.K. and Poole, P.S. (2013) The plant microbiome. *Genome Biology*. [online]. 14 (6), pp.1–10.

Vančura, V. (1964) Root exudates of plants: analysis of root exudates of barley and wheat in their initial phases of growth. *Plant and Soil*. [online]. 21 (2), pp.231–248.

Vandamme, P., Pot, B., Gillis, M., De Vos, P., Kersters, K. and Swings, J. (1996) Polyphasic Taxonomy, a Consensus Approach to Bacterial Systematics. *Microbiological reviews*. [online].

60 (2), pp.98407–438.

Vansteenkiste, D., Tirry, L., Acker, J. van and Stevens, M. (2004) Predispositions and symptoms of *Agrilus* borer attack in declining oak trees. *Annals of Forest Science*. [online]. 61 (8), pp.815–823.

Varnam, A. and Evans, M. (1995) *Environmental microbiology* London, Manson Publishing.

Versalovic, J., M, S., FJ, de B. and JR, L. (1994) Genomic Fingerprinting of Bacteria Using Repetitive Sequence-Based Polymerase Chain Reaction. *Methods in Molecular and Cellular Biology*. [online]. 25 (1), pp.25–40.

Vetterlein, D., Carminati, A., Kögel-Knabner, I., Bienert, G.P., Smalla, K and Oburger, E. (2020) Rhizosphere Spatiotemporal Organization—A Key to Rhizosphere Functions. *Frontiers in Agronomy*. [online]. 2 (July), pp.1–22.

Vieira-Silva, S., Touchon, M., Abby, S.S. and Rocha, E.P.C. (2011) Investment in rapid growth shapes the evolutionary rates of essential proteins. *Proceedings of the National Academy of Sciences of the United States of America*. [online]. 108 (50), pp.20030–20035.

Vives-Peris, V., de Ollas, C., Gómez-Cadenas, A. and Pérez-Clemente, R.M. (2020) Root exudates: from plant to rhizosphere and beyond. *Plant Cell Reports*. [online]. 39 (1), Springer Berlin Heidelberg, pp.3–17. Available from: <https://doi.org/10.1007/s00299-019-02447-5>.

Voroney, R.P. (2007) *Chapter 2 THE SOIL HABITAT Soil Microbiology, Ecology and Biochemistry*. Third Edit. [online]. Elsevier Inc. Available from: <https://www.sciencedirect.com/science/article/pii/B9780080475141500068>.

Vos, M., Wolf, A.B., Jennings, S.J. and Kowalchuk, G.A. (2013) Micro-scale determinants of bacterial diversity in soil. *Federation of European Microbiological Societies Microbiology Reviews*. [online]. 37 (6), pp.936–954.

Vu, N.T. and Oh, C.S. (2020) Bacteriophage usage for bacterial disease management and diagnosis in plants. *Plant Pathology Journal*. [online]. 36 (3), pp.204–217.

Vuts, J., Woodcock, C.M., Sumner, M.E., Caulfield, J.C., Reed, K and Inward, D.J. (2016) Responses of the two-spotted oak buprestid, *Agrilus biguttatus* (Coleoptera: Buprestidae), to host tree volatiles. *Pest Management Science*. [online]. 72 (4), pp.845–851.

- Walker, T.S., Bais, H.P., Grotewold, E. and Vivanco, J.M. (2003) Update on Root Exudation and Rhizosphere Biology. *Root Exudation and Rhizosphere Biology*. [online]. 132 (May), pp.44–51.
- Walters, K.E. and Martiny, J.B.H. (2020) Alpha-, beta-, and gamma-diversity of bacteria varies across habitats. *Public Library of Science ONE*. [online]. 15 (9 September). Available from: <https://doi.org/10.1371/journal.pone.0233872>.
- Wargo, P.M. (1996) Consequences of environmental stress on oak: predisposition to pathogens. *Annales des Sciences Forestieres*. [online]. 53 (2–3), pp.359–368.
- Weig, A.R., Peršoh, D., Werner, S., Betzlbacher, A. and Rambold, G. (2013) Diagnostic assessment of mycodiversity in environmental samples by fungal ITS1 rDNA length polymorphism. *Mycological Progress*. [online]. 12 (4), pp.719–725.
- West, J.S. (2014) Plant Pathogen Dispersal. *Encyclopedia of Life Sciences*. [online]. pp.1–7.
- Wiley, E. (2020) Do Carbon Reserves Increase Tree Survival during Stress and Following Disturbance? *Current Forestry Reports*. 6 (1), Current Forestry Reports, pp.14–25.
- Wisplinghoff, H. (2017) *Pseudomonas* spp., *Acinetobacter* spp. and Miscellaneous Gram-Negative *Bacilli*. *Infectious Diseases, 2-Volume Set*. [online]. Elsevier, pp.1579-1599.e2.
- De Wit, R. and Bouvier, T. (2006) ‘Everything is everywhere, but, the environment selects’; what did Baas Becking and Beijerinck really say? *Environmental Microbiology*. [online]. 8 (4), pp.755–758.
- Wu, K., Luo, J., Li, J., An, Q., Yang, X., Liang, Y. and Li, T. (2018) Endophytic bacterium *Buttiauxella* sp. SaSR13 improves plant growth and cadmium accumulation of hyperaccumulator *Sedum alfredii*. *Environmental Science and Pollution Research*. [online]. 25 (22), pp.21844–21854.
- Xin, X.F., Kvitko, B. and He, S.Y. (2018) *Pseudomonas syringae*: What it takes to be a pathogen. *Nature Reviews Microbiology*. [online]. 16 (5), Nature Publishing Group, pp.316–328. Available from: <http://dx.doi.org/10.1038/nrmicro.2018.17>.
- Xu, J., Zhang, Y., Zhang, P., Trivedi, P., Riera, N and Wang, Y. (2018) The structure and function of the global citrus rhizosphere microbiome. *Nature Communications*. [online]. 9 (1).
- Xu, Z., Hansen, M.A., Hansen, L.H., Jacquiod, S. and Sørensen, S.J. (2014) Bioinformatic

approaches reveal metagenomic characterization of soil microbial community. *Public Library of Science ONE*. [online]. 9 (4).

Xue, W., Huang, L. and Yu, F.H. (2016) Spatial heterogeneity in soil particle size: Does it affect the yield of plant communities with different species richness? *Journal of Plant Ecology*. [online]. 9 (5), pp.608–615.

Yan, J.J., Liu, Y.Y., Wang, R.Q., Mukhtar, I., Liu, F., Lin, Z.Y., Jiang, Y.J. and Xie, B.G. (2019) First report of *Cedecea neteri* causing yellow sticky disease in *Flammulina velutipes* in China. *Plant Disease*. [online]. 103 (5), pp.6–11.

Yemshanov, D., Koch, F.H., Ben-Haim, Y., Downing, M., Sapio, F. and Siltanen, M. (2013) A new multicriteria risk mapping approach based on a multiattribute frontier concept. *Risk Analysis*. [online]. 33 (9), pp.1694–1709.

Yoon, S.H., Ha, S.M., Kwon, S., Lim, J., Kim, Y., Seo, H. and Chun, J. (2017) Introducing EzBioCloud: A taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *International Journal of Systematic and Evolutionary Microbiology*. [online]. 67 (5), pp.1613–1617.

Yu, K., Pieterse, C.M.J., Bakker, P.A.H.M. and Berendsen, R.L. (2019) Beneficial microbes going underground of root immunity. *Plant Cell and Environment*. [online]. 42 (10), pp.2860–2870.

Zalkalns, O. and Celma, L. (2021) The distribution of bacteria *Gibbsiella quercinecans* and *Brenneria goodwinii* in oak (*Quercus robur* L.) stands in Latvia. *IOP Conference Series: Earth and Environmental Science*. [online]. 875 (1).

Zhao, Y., Blumer, S.E. and Sundin, G.W. (2005) Identification of *Erwinia amylovora* genes induced during infection of immature pear tissue. *Journal of Bacteriology*. [online]. 187 (23), pp.8088–8103.

Zhao, Y., Damicone, J.P. and Bender, C.L. (2002) Detection, survival, and sources of inoculum for bacterial diseases of leafy crucifers in Oklahoma. *Plant Disease*. [online]. 86 (8), pp.883–888.

Zytynska, S.E., Eicher, M., Rothballer, M. and Weisser, W.W. (2020) Microbial-Mediated Plant Growth Promotion and Pest Suppression Varies Under Climate Change. *Frontiers in Plant*

*Science*. [online]. 11 (September), pp.1–9.

## Supplementary data:

**Suppl. Table. S1.** Comparison of the sequencing identifications obtained from different recovery methods on the Hatchlands samples H7 and H8, H13 and H14. Isolates were taken from the samples and recovery performed using suspension and disruption in EE broth and ¼ Ringers. The left hand column shown the unique isolate number and ID for EE isolation in comparison to the right hand column which show the same information for Ringers isolation allowing for the comparison of isolates from each method.

EE Isolation Method		Ringers Isolation Method	
Isolate number	Isolate ID	Isolate number	Isolate ID
H7 E1	<i>Buttiauxella brennerae</i> (99.48)	H7 EA	<i>Stenotrophomonas</i> sp. R-41388 (99.21)
H7 E2	uncultured bacteia (99.89)	H7 EB	<i>Stenotrophomonas maltophilia</i> (94.78)
H7 E3	<i>Serratia proteamaculans</i> (99.74)	H7 EC	<i>Pseudomonas</i> sp. ADAK22 (99.24)
H7 E4	<i>Serratia proteamaculans</i> (99.87)	H7 ED	<i>Paenibacillus amylolyticus</i> (99.68)
H7 E5	<i>Serratia quinivorans</i> (100)	H7 EE	<i>Streptomyces</i> sp. (99.79)
H7 E6	<i>Buttiauxella brennerae</i> (99.49)	H7 EF	<i>Streptomyces</i> sp. 1H-TWYE2 (99.69)
H7 E7	<i>Serratia</i> sp. W2Dec25 (99.87)	H7 EG	<i>Stenotrophomonas maltophilia</i> (99.59)
H7 E8	<i>Serratia quinivorans</i> (100)	H7 EH	<i>Stenotrophomonas</i> sp. R-41388 (99.69)
H7 E9	uncultured bacterium (99.37)	H7 EI	<i>Chryseobacterium</i> sp. A1-652 (99.90)
H7 E10	<i>Serratia proteamaculans</i> (100)	H7 EJ	<i>Bacillus mycoides</i> (99.79)
H7 E11	<i>Serratia proteamaculans</i> (100)	H7 EK	<i>Pseudomonas tolaasii</i> (99.48)

H7 E12	<i>Serratia proteamaculans</i> (100)	H7 EL	<i>Serratia proteamaculans</i> (99.62)
H7 E13	<i>Enterobacteriaceae</i> bacterium SZMC-H2579B (99.48)	H7 EM	<i>Buttiauxella brennerae</i> (99.58)
H7 E14	<i>Buttiauxella</i> sp. 3AFRM03 (99.58)	H7 EN	<i>Pseudomonas psychrophila</i> (98.11)
H7 E15	<i>Serratia proteamaculans</i> (100)	H7 EO	<i>Serratia</i> sp. W2Dec25 (99.87)
H7 E16	uncultured bacterium (99.68)		
H7 E18	<i>Enterobacteriaceae</i> bacterium SZMC-H2579B (99.58)		
H7 E19	<i>Serratia</i> sp. 3ACOL1 (99.69)		
H7 E20	uncultured bacterium (99.79)		
H7 N1	<i>Serratia quinivorans</i> (99.60)	H7 NA	<i>Acetobacter ascendens</i> (99.71)
H7 N2	<i>Serratia quinivorans</i> (99.87)	H7 NB	<i>Bacillus</i> sp. JG-B35 (99.79)
H7 N3	<i>Buttiauxella</i> sp. 3AFRM03 (95.82)	H7 NC	<i>Bacillus</i> sp. JG-B35 (99.32)
H7 N4	<i>Serratia quinivorans</i> (99.60)	H7 ND	<i>Peribacillus butanolivorans</i> (99.44)
H7 N6	<i>Serratia quinivorans</i> (99.87)	H7 NE	<i>Arthrobacter ginkgonis</i> (99.89)
H7 N7	<i>Serratia quinivorans</i> (99.87)	H7 NF	<i>Pseudomonas</i> sp. CFSAN084952 (99.79)

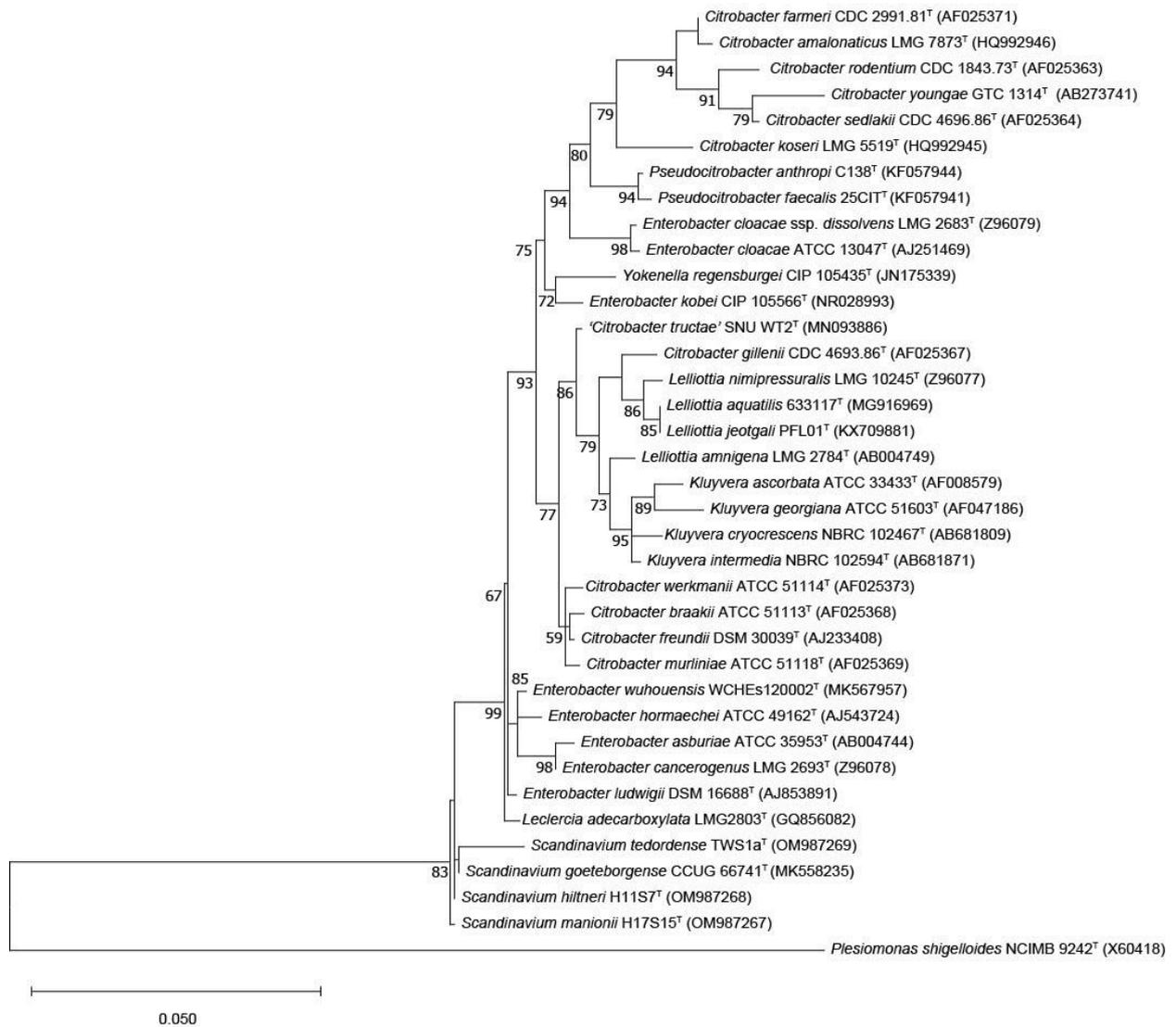
H7 N8	<i>Serratia quinivorans</i> (99.87)	H7 NG	<i>Stenotrophomonas rhizophila</i> (99.90)
H7 N9	<i>Scandinavium goeteborgense</i> (99.06)	H7 NH	<i>Paenibacillus amylolyticus</i> (99.89)
		H7 NI	<i>Stenotrophomonas maltophilia</i>
		H7 NJ	<i>Paenibacillus</i> sp. (99.47)
H8 N1	<i>Serratia quinivornas</i> (99.19)	H8 NA	<i>Bacillus mycoides</i> (99.78)
H8 N2	<i>Serratia quinivornas</i> (99.19)	H8 NB	<i>Lysinibacillus parviboronicapiens</i> (99.78)
H8 N3	<i>Citrobacter gillengii</i> (100)	H8 NC	<i>Peribacillus butanolivorans</i> (99.69)
H8 N4	<i>Serratia quinivornas</i> (99.19)	H8 ND	<i>Bacillus mycoides</i> (99.59)
H8 N5	<i>Citrobacter gillengii</i> (100)	H8 NE	<i>Paenibacillus</i> sp. (99.46)
H8 N6	<i>Citrobacter freundii</i> (99.87)	H8 NF	<i>Pseudomonas migulae</i> (98.97)
H8 N7	<i>Rhodococcus erythropolis</i> (99.73)	H8 NG	<i>Streptomyces cirratus</i> (99.48)
H8 N9	<i>Serratia quinivornas</i> (99.19)	H8 NH	<i>Lysinibacillus sphaericus</i> (99.46)
H8 N10	<i>Serratia</i> sp. (98.11)	H8 NI	<i>Streptomyces subrutilus</i> (99.69)
		H8 NJ	<i>Rahnella</i> sp. ERMR1:05 (99.59)
		H8 NK	<i>Curtobacterium</i> sp. (99.89)
		H8 NL	<i>Pseudomonas chlororaphis</i> (99.59)
		H8 NM	<i>Bacillus</i> sp. HD-2011-F1 (99.49)
H8 W1	<i>Enterobacter</i> sp. IK-2016 (99.28)	H8 WA	<i>Bacillus</i> sp. BYMS05 (99.46)
H8 W2	<i>Raoultella terrigena</i> (99.79)	H8 WB	<i>Bacillus mycoides</i> (99.90)
H8 W3	<i>Serratia</i> sp. W2Dec25 (99.87)	H8 WC	<i>Bacillus</i> sp. (in: Bacteria) (99.59)

H8 W4	<i>Buttiauxella agrestis</i> (99.60)	H8 WD	<i>Bacillus mycoides</i> (99.79)
H8 W5	uncultured bacterium (99.37)	H8 WE	<i>Bacillus licheniformis</i> (99.49)
H8 W6	uncultured bacterium (99.69)	H8 WF	<i>Streptomyces</i> sp. MM108 (99.19)
H8 W7	uncultured bacterium (99.58)	H8 WG	<i>Psychrobacillus psychrodurans</i> (99.79)
H8 W8	uncultured bacterium (99.47)	H8 WH	bacterium (99.14)
H8 W10	<i>Enterobacter cloacae</i> (97.93)	H8 WI	<i>Paenibacillus</i> sp. 37 (98.88)
H8 W11	<i>Serratia</i> sp. (100)	H8 WJ	uncultured <i>Bacillus</i> sp. (99.69)
H8 W12	<i>Serratia</i> sp. 3ACOL1 (99.90)	H8 WK	<i>Bacillus idriensis</i> (99.68)
H8 W13	<i>Serratia fonticola</i> (99.79)	H8 WL	<i>Paenibacillus</i> sp. FSL H7-0737 (99.59)
H13 N1	<i>Rahnella victoriana</i> (99.73)	H13 NA	<i>Peribacillus butanolivorans</i> (99.59)
H13 N2	<i>Rahnella aquatilis</i> (92.45)	H13 NB	uncultured bacterium (99.05)
H13 N3	<i>Rhodococcus erythropolis</i> (95.96)	H13 NC	<i>Bacillus mycoides</i> (99.62)
H13 N4	<i>Rhodococcus erythropolis</i> (96.06)	H13 ND	<i>Bacillus mycoides</i> (98.0)
H13 N5	<i>Rhodococcus erythropolis</i> (95.82)	H13 NE	<i>Bacillus</i> sp. CDC-c (99.68)
H13 N6	<i>Serratia quinivorans</i> (96.77)	H13 NF	<i>Paenibacillus sinopodophylli</i> (99.57)
H13 N7	<i>Buttiauxella</i> sp. 3AFRM03 (95.82)	H13 NG	<i>Streptomyces</i> sp. MM87 (99.80)
H13 N8	<i>Buttiauxella</i> sp. 3AFRM03 (95.55)	H13 NH	<i>Microbacterium</i> sp. VKM Ac-1808 (99.79)

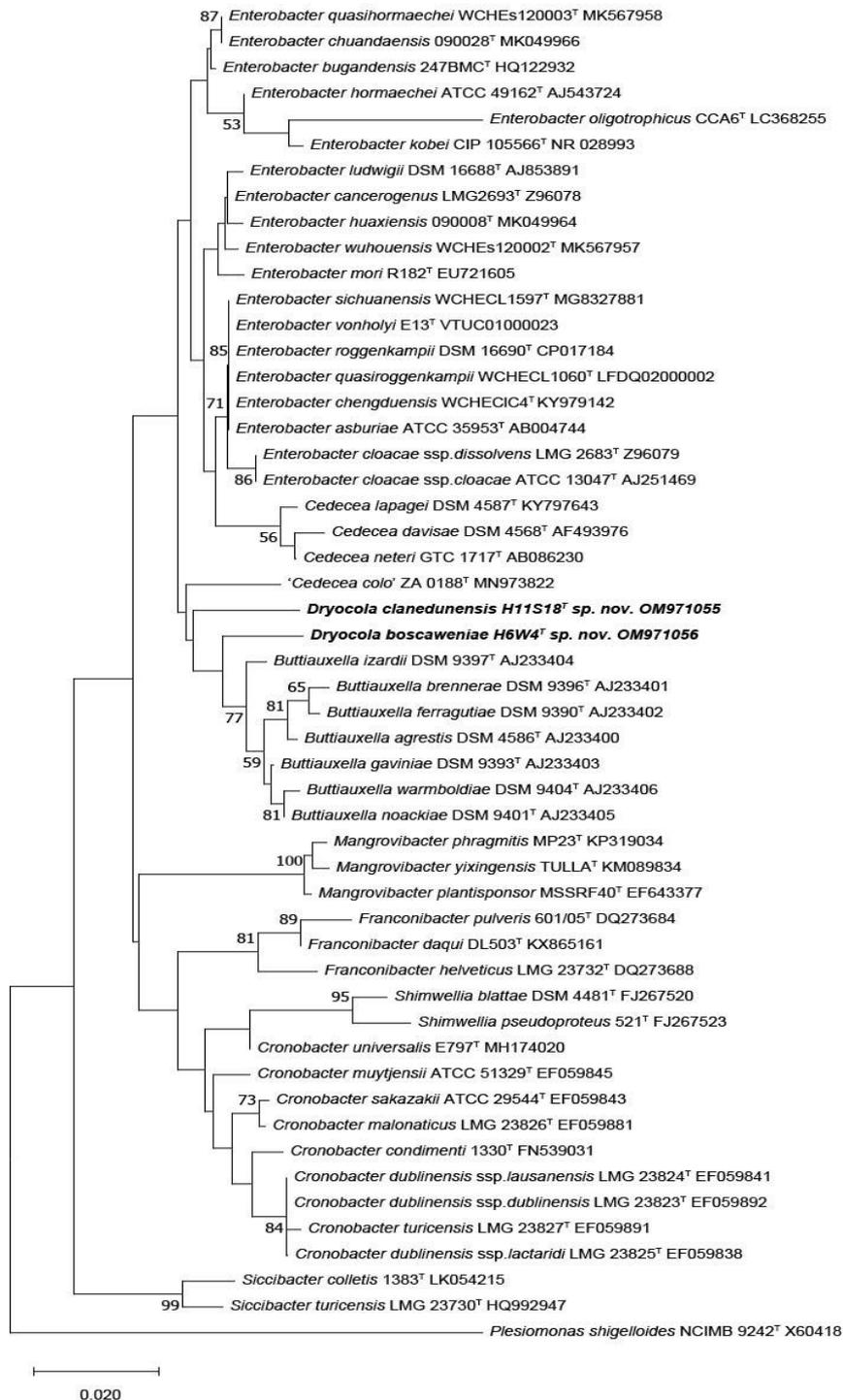
H13 N9	<i>Rahnella victoriana</i> (99.73)	H13 NI	<i>Neobacillus novalis</i> (99.68)
H13 N10	<i>Buttiauxella ferragutiae</i> (95.96)	H13 NJ	<i>Stenotrophomonas rhizophila</i> (99.69)
H13 N11	<i>Buttiauxella ferragutiae</i> (96.63)	H13 NK	<i>Kurthia</i> sp. NRRL (99.78)
H13 N12	<i>Rhodococcus erythopolis</i> (95.82)	H13 NL	<i>Microbacterium testaceum</i> (90.34)
H13 N13	<i>Rhodococcus erythopolis</i> (95.82)	H13 NM	<i>Stenotrophomonas</i> sp. (99.18)
H13 N14	<i>Buttiauxella agrestis</i> (95.69)	H13 NN	<i>Viridibacillus</i> sp. (99.47)
H13 W1	<i>Lelliottia aqualitilis</i> (100)	H13 WA	<i>Serratia quinivorans</i> (99.90)
H13 W2	<i>Escherichia fergusonii</i> (100)	H13 WB	<i>Paraburkholderia ginsengisoli</i> (99)
H13 W3	<i>Escherichia coli</i> (99.87)	H13 WC	<i>Streptomyces inaequalis</i> (99.59)
H13 W4	<i>Escherichia coli</i> (100)	H13 WD	<i>Viridibacillus</i> sp. (99.67)
H13 W5	<i>Escherichia coli</i> (100)	H13 WE	<i>Peribacillus muralis</i> (99.39)
H13 W6	<i>Lelliottia aqualitilis</i> (100)	H13 WF	<i>Bacillus velezensis</i> (100)
H13 W7	<i>Serratia quinivornas</i> (97.04)	H13 WG	<i>Paenibacillus</i> sp. (99.77)
H13 W8	<i>Shigella dysenteriae</i> (99.33)	H13 WH	<i>Micrococcus</i> sp. (99.49)
H13 W9	<i>Escherichia coli</i> (100)	H13 WI	<i>Chryseobacterium piscium</i> (99.79)
H13 W10	<i>Escherichia coli</i> (100)	H13 WJ	<i>Bacillus</i> sp. (99.79)
H13 W11	<i>Escherichia marmotae</i> (99.87)	H13 WK	<i>Bacillus mycoides</i> (99.49)

H13 W12	<i>Escherichia coli</i> (100)	H13 WL	uncultured bacterium (99.48)
H13 W13	<i>Escherichia marmotae</i> (100)	H13 WM	<i>Bacillus cereus</i> (99.68)
		H13 WN	<i>Peribacillus butanolivorans</i> (100%)
		H13 WO	<i>Streptomyces celluloflavus</i> (99.79)
		H13 WP	<i>Silvimonas</i> sp. (98.97)
		H13 WQ	<i>Viridibacillus</i> sp. Or <i>Kurthia</i> sp.(both 100)
		H13 WR	<i>Streptoverticillium verticillium</i> (100)
H14 E1	<i>Rahnella aquatilis</i> (99.60)	H14 EA	<i>Bacillus</i> sp. SP2601 (99.68)
H14 E2	<i>Rahnella bruchi</i> (100)	H14 EB	<i>Bacillus mycoides</i> (99.38)
H14 E3	<i>Serratia liquefaciens</i> (99.73)	H14 EC	<i>Viridibacillus</i> sp. (99.67)
H14 E4	<i>Rahnella victoriana</i> (100)	H14 ED	<i>Paenibacillus</i> sp. NA10131 (98.56)
H14 E5	<i>Escherichia marmotae</i> (100)	H14 EE	<i>Bacillus</i> sp. (in: Bacteria) (98.97)
H14 E6	<i>Rahnella victoriana</i> (100)	H14 EF	<i>Bacillus</i> sp. JG-B3 (99.58)
H14 E7	<i>Escherichia marmotae</i> (100)	H14 EG	<i>Bacillus</i> sp. (in: bacteria) (99.47)
H14 E8	<i>Escherichia marmotae</i> (100)	H14 E8	<i>Pseudomonas</i> sp. ADAK18 (99.69)
H14 E9	<i>Escherichia marmotae</i> (100)	H14 EH	<i>Bacillus</i> sp. (in: Bacteria) (99.39)
H14 E10	<i>Serratia proteamaculans</i> (99.60)	H14 EI	<i>Bacillus circulans</i> (99.68)
H14 E11	<i>Serratia proteamaculans</i> (99.60)	H14 EJ	<i>Bacillus circulans</i> (99.27)
H14 E12	<i>Rahnella victoriana</i> (99.73)	H14 EK	<i>Bacillus</i> sp. (in: bacteria) (99.42)
H14 W1	<i>Rahnella variigena</i> (100)	H14 WA	<i>Neobacillus novalis</i> (99.48)
H14 W2	<i>Rahnella variigena</i> (100)	H14 WB	<i>Neobacillus novalis</i> (99.57)

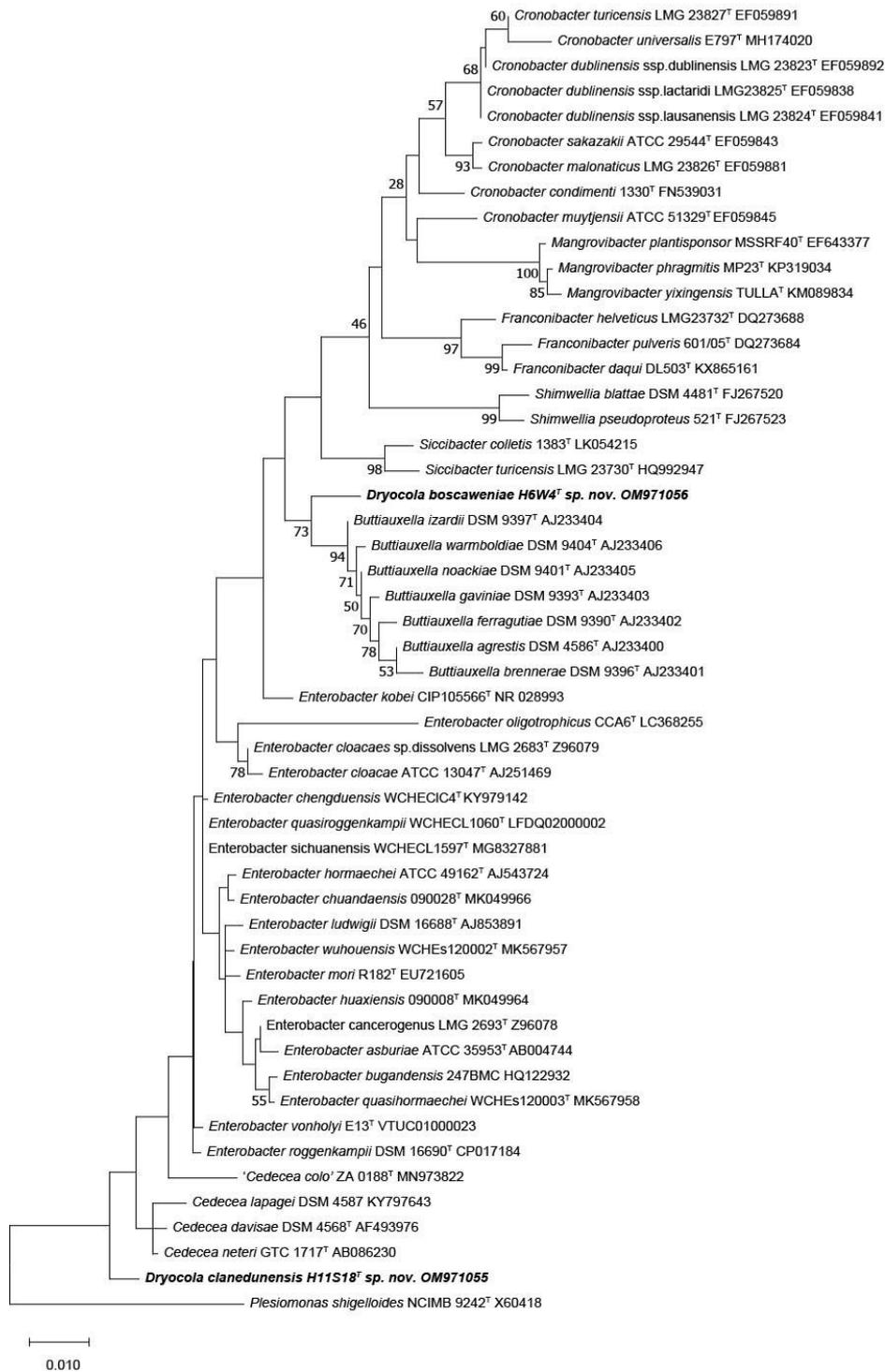
H14 W3	<i>Serratia quinivorans</i> (96.60)	H14 WC	<i>Rhodococcus koreensis</i> (99.88)
H14 W4	<i>Rahnella victoriana</i> (100)	H14 WD	uncultured bacterium (98.09)
H14 W5	<i>Rahnella variigena</i> (100)	H14 WE	<i>Neobacillus novalis</i> (99.57)
H14 W6	<i>Rahnella variigena</i> (100)	H14 WF	<i>Serratia proteamaculans</i> (99.38)
H14 W7	<i>Serratia liquefaciens</i> (98.65)	H14 WG	<i>Bacillus mycoides</i> (99.38)
H14 W8	<i>Serratia liquefaciens</i> (100)	H14 WH	<i>Streptomyces</i> sp. (99.79)
H14 W9	<i>Buttiauxella</i> sp. 3AFRM03 (99.67)	H14 WI	<i>Viridibacillus</i> sp./ <i>Kurthia</i> sp. (both 99.79)
H14 W10	<i>Rahnella variigena</i> (100)	H14 WJ	<i>Streptoverticillium verticillium</i> subsp. <i>quintum</i> (99.79)



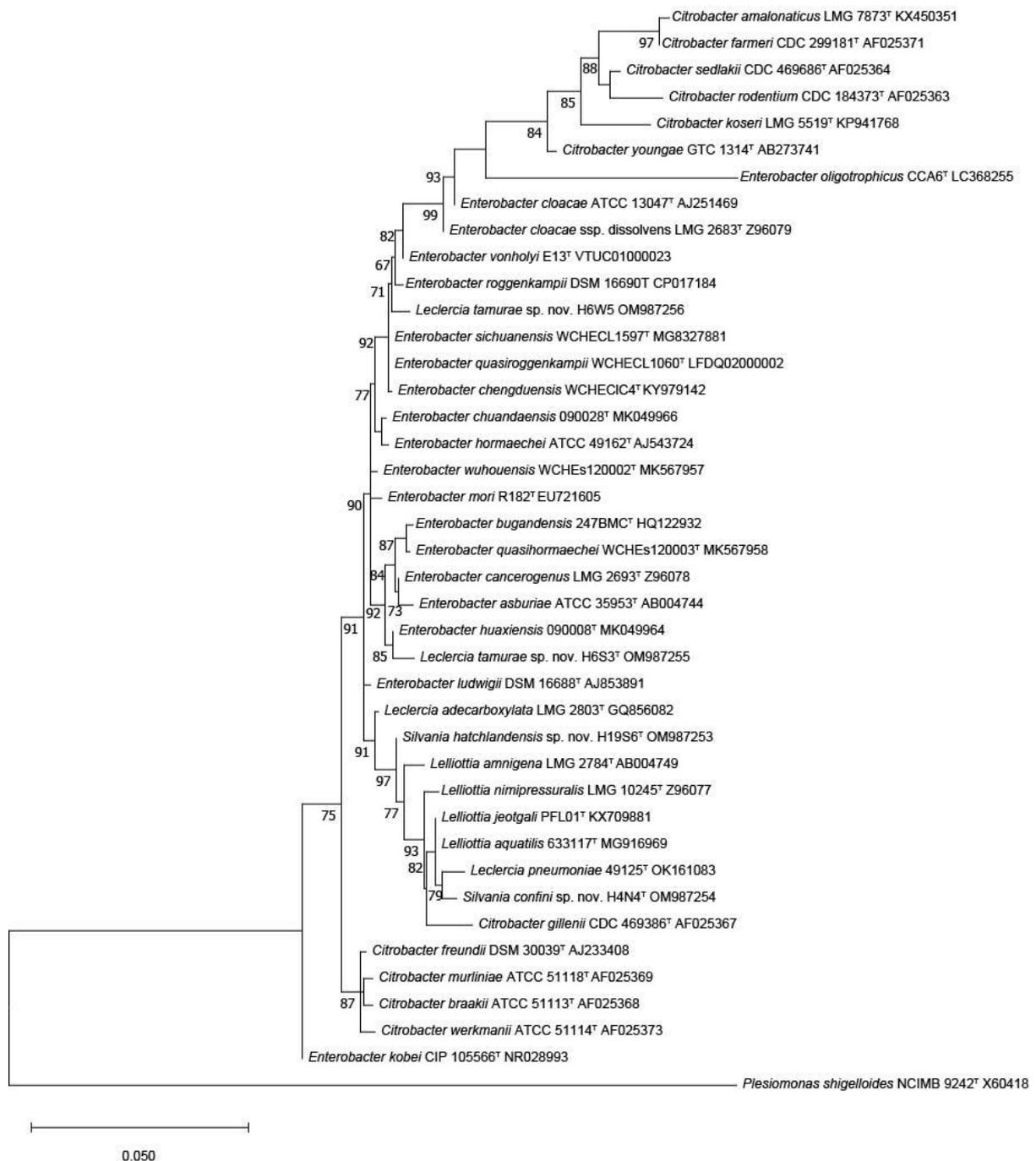
**Suppl. Figure. S1.** 16S rRNA gene maximum likelihood phylogenetic tree for *Scandinavium* species, the proposed novel species and their closest phylogenetic neighbours. The near complete (1,346 bp) 16S rRNA gene sequences were used, 1000 bootstrap replicate percentage values (> 50 %) are shown at the nodes and the scale bar indicates the number of nucleotide substitutions per site. The outgroup is *Plesiomonas shigelloides* NCIMB 9242<sup>T</sup>. GenBank accession numbers shown in parentheses and <sup>T</sup> = type strain.



**Suppl. Figure. S2.** Neighbour Joining 16S rRNA phylogenetic tree for the novel genus *Dryocola* and closest phylogenetic neighbours. Near complete (1,346 bp) 16S rRNA gene sequences were used, with the scale showing the nucleotide substitutions per site and bootstrap values exceeding 50 % from 1000 replicates shown at nodes. Species names are followed by the strain number and GenBank accession number, with <sup>T</sup> indicating the type strain. The outgroup is *Plesiomonas shigelloides* NCIMB 9242<sup>T</sup>.



**Supple. Figure. S3.** Maximum Likelihood 16S rRNA gene phylogenetic tree for the novel genus *Dryocola* and closest phylogenetic neighbours. Near complete (1,346 bp) 16S rRNA gene sequences were used, with the scale showing the nucleotide substitutions per site and bootstrap values exceeding 50 % from 1000 replicates shown at nodes. Species names are followed by the strain number and GenBank accession number, with <sup>T</sup> indicating the type strain. The outgroup is *Plesiomonas shigelloides* NCIMB 9242<sup>T</sup>.



**Supple. Figure. S4.** Maximum Likelihood phylogenetic tree based on 16S rRNA gene sequences for species of the novel genus *Sylvania* gen. nov., *Leclercia*, the novel species *Leclercia tamurae* sp. nov. and several closest phylogenetic neighbours. Near complete (1,346 bp) 16S rRNA gene sequences were used, with the scale showing the nucleotide substitutions per site and bootstrap values exceeding 50 % from 1000 replicates shown at nodes. Species names are followed by the strain number and GenBank accession number, with T indicating the type strain. The outgroup is *Plesiomonas shigelloides* NCIMB 9242<sup>T</sup>.

**Suppl. Table. S2.** The whole genome sequence information for strains investigated in this study.

Strain	GenBank accession	Biosample number	Size (Mbp)	Number of contigs	N50	Mean coverage	Number of coding sequences	Numbers of RNAs	GC content (mol %)
<i>Scandinavium goeteborgense</i>									
H5W5	JALIGB000000000	SAMN27163994	4.77	21	652013	90.7	4,439	103	54.5
<i>Scandinavium hiltneri</i>									
H11S7 <sup>T</sup>	JALIGE000000000	SAMN27163997	4.84	79	309159	169.2	4,565	90	53.9
BAC 14-01-01	JALIGF000000000	SAMN27163998	4.61	63	295467	143.8	4,299	96	54.2
<i>Scandinavium manionii</i>									
H17S15 <sup>T</sup>	JALIGC000000000	SAMN27163995	4.64	40	394704	74.1	4,314	90	54.2
SB 3.3	JALIGD000000000	SAMN27163996	4.39	47	439894	82.6	4,834	91	53.9

<i>Scandinaviu tedordense</i>									
TWS1a <sup>T</sup>	JALIGG000000000	SAMN27163999	4.75	61	433482	152.6	4,457	94	53.9

**Suppl. Table. S3.** Whole genome sequence information of strains investigated in this study.

Strain	GenBank accession	Biosample number	Size (Mbp)	Number of contigs	N50	Number of coding sequences	Numbers of RNAs	G + C content (mol %)
<i>Dryocola boscaweniae</i>								
H6W4 <sup>T</sup>	JALHAP000000000	SAMN26554629	4.41	85	279 318	4290	90	53.0
H20N1	JALHAN000000000	SAMN26554631	4.40	69	279 318	4275	91	53.0
H18W14	JALHAO000000000	SAMN26554630	4.45	96	346 750	4309	84	53.1
<i>Dryocola clanedunesis</i>								
H11S18 <sup>T</sup>	JALHAM000000000	SAMN26554632	5.23	124	190 839	5134	86	53.8

H16N7	JALHAL000000000	SAMN26554633	4.82	132	431 659	4610	91	53.9
-------	-----------------	--------------	------	-----	---------	------	----	------

**Suppl. Table. S4.** Whole genome sequence information of strains investigated in this study.

Strain	GenBank accession	Biosample number	Size (Mbp)	Number of contigs	N50	Number of coding sequences	Numbers of RNAs	G + C content (mol %)
<i>Leclercia adecarboxylata</i>								
H10E4	JAMHKT000000000	SAMN28207097	4.83	61	336 719	4 585	91	55.6
<i>Leclercia tamurae</i>								
H6S3 <sup>T</sup>	JAMHKS000000000	SAMN28207096	4.71	80	268 955	4 465	96	56.4
H6W5	JAMHKR000000000	SAMN28207095	4.86	94	323 472	4 647	94	56.4
<i>Silvania hatchlandensis</i>								
H19S6 <sup>T</sup>	JAMGZK000000000	SAMN28207118	4.78	67	207 460	4 493	88	55.9

<b><i>Silvania confinis</i></b>								
H4N4 <sup>T</sup>	JAMGZJ0000000000	SAMN28207119	4.87	82	225 593	4 675	92	55.7

*Scandinaviium* protologs:

### **Emendation to the Genus *Scandinaviium***

*Scandinaviium* (Scan.di.na'vi.um. N.L. neut. n. *Scandinaviium* genus named after Scandinavia: the European peninsula where the type strain of the type species was isolated and characterised).

This description is based on the data from Marathe *et al.*, 2019 and this study.

Gram-negative straight rods which are 1 - 1.3 x 1.9 – 2.7  $\mu\text{m}$ , motile by peritrichous flagella and possess fimbriate. Facultative anaerobic, oxidase negative and catalase positive. The colonies appear as moist, white circles with clear rims on CBA averaging 1 – 3 mm in size. Growth is observed at 4 – 37 °C with an optimum growth temperature of 30 °C; the salt and pH range are 1 – 7 % and 6 – 8 respectively. Outside of the salt range, coagulated masses of growth can be observed in broth. Positive for  $\beta$ -galactosidase. Negative for arginine dihydrolase, ornithine decarboxylase, H<sub>2</sub>S production, urease, tryptophan deaminase, indole production, acetoin production and gelatinase. Nitrate is reduced to nitrite. Acid is produced from: glucose, mannitol, amygdalin, L-arabinose, glycerol, D-ribose, D-xylose, D-galactose, D-fructose, D-mannose, methyl- $\alpha$ D-glucopyranoside, N-acetylglucosamine, arbutin, esculin ferric citrate, salicin, D-cellobiose, D-maltose, D-lactose, D-trehalose, gentiobiose and potassium gluconate (API 20 and 50 CHB/E). Acid is produced from galacturonate and  $\beta$ -glucosidase is utilised as a carbon source (ID32). Utilises the following carbon sources: dextrin, D-salicin, N-acetyl-D-glucosamine, N-acetyl- $\beta$ -D-mannosamine, N-acetyl neuraminic acid, 3-methyl glucose, inosine, D-glucose-6-phosphate, D-fructose-6-phosphate, glycyl-L-proline, L-alanine, L-arginine, L-aspartic acid, L-glutamic acid, L-histidine, L-serine, D-galaturonic acid, L-galactonic acid lactone, D-gluconic acid, D-glucuronic acid, glucuronamide, quinic acid, methyl pyruvate, L-lactic acid, citric acid, L-malic acid, bromo-succinic acid and acetic acid. Resistant to 1% sodium lactate, D-serine, troleandomycin, rifamycin, lincomycin, guanidine hydrochloric acid, niaproof 4, vancomycin, tetrazolium violet, tetrazolium blue, nalidixic acid, lithium chloride, aztreonam and sodium butyrate (Biolog Gen III). Variable for lysine decarboxylase and citrate utilization; fermentation of sorbitol, rhamnose, melibiose, D-adonitol, dulcitol, D-raffinose, D-turanose, D-fucose, L-fucose, D-arabitol and potassium 2-

ketogluconate; acidification of phenol red, palatinose and production of malonate,  $\alpha$ -glucosidase,  $\alpha$ -galactosidase. The following carbon sources are variable: sucrose, stachyose,  $\beta$ -methyl-D-glucoside, N-acetyl-D-galactosamine, myo-inositol, D-aspartic acid, L-pyroglutamic acid, mucic acid, D-saccharic acid, p-hydroxy-phenylacetic acid, D-lactic acid methyl ester, tween 40,  $\gamma$ -amino-butyric acid,  $\alpha$ -hydroxy-butyric acid,  $\beta$ -hydroxy-D,L-butyric acid and acetoacetic acid, and susceptibility to fusidic acid, D-serine and minocycline also varies. The major classes of fatty acids are C<sub>12:0</sub>, C<sub>14:0</sub>, C<sub>16:0</sub>, C<sub>18:1</sub>  $\omega$ 7c, C<sub>17:0</sub> cyclo, summed features 2 (C<sub>14:0</sub> 3-OH and/or iso-C<sub>16:1</sub>) and summed features 3 (C<sub>16:1</sub>  $\omega$ 7c and/or C<sub>16:1</sub>  $\omega$ 6c). The DNA G + C content ranges from 53.9 – 54.5 %.

Strains have been isolated from a range of sources, including human wound infection, rhizosphere soil and bleeding lesions of broadleaf hosts.

The type species is *Scandinavium goeteborgense* (CCUG 66741<sup>T</sup> = CECT 9823<sup>T</sup> = NCTC 14286<sup>T</sup>).

#### **Description of *Scandinavium hiltneri* sp. nov.**

*Scandinavium hiltneri* (hilt'ne.ri. N.L. gen. n. *hiltneri*, named in honour of Lorenz Hiltner, the scientist who coined the term 'rhizosphere' in 1904, where the majority of isolates originated from).

The species shares the major characteristics of the genus. Gram negative motile rods (1.08 – 1.19 x 1.96 – 2.41  $\mu$ m) that occur singly. Colonies appear as moist, raised, white circles with clear smooth margins on CBA averaging 1 – 2 mm in size. Positive for lysine decarboxylase and citrate utilisation. Acid is produced from: sorbitol, rhamnose, D-adonitol, D-raffinose, D-arabitol, potassium 2-ketogluconate (API 20/ 50 CHB/E). Positive for production of  $\alpha$ -galactosidase and acidification of phenol red (ID32). The following carbon sources are utilised: sucrose, stachyose,  $\beta$ -methyl-D-glucoside, N-acetyl-D-galactosamine, myo-inositol, D-aspartic acid, D-serine, L-pyroglutamic acid, p-hydroxy-phenylacetic acid, D-lactic acid methyl ester, tween 40 and acetoacetic acid. Resistant to fusidic acid and minocycline (Biolog Gen III). Variable for the fermentation of melibiose, dulcitol, D-turanose, L-fucose and utilisation of the carbon source mucic acid.

The DNA G + C content of the type strain is 54.2 mol %.

The type strain is H11S7<sup>T</sup> (= LMG 32612<sup>T</sup> = CCUG 76179<sup>T</sup>) and was isolated from *Quercus robur* rhizosphere soil in Hatchlands, Guildford, UK.

#### **Description of *Scandinavium manionii* sp. nov.**

*Scandinavium manionii* (ma.ni.o'ni.i. N.L. gen. n. *manionii*, named after Paul Manion, who defined the decline disease spiral furthering our understanding of the range of influences on forest diseases).

The species shares the major characteristics of the genus. Gram negative motile rods (1.12 - 1.18 x 2.11 – 2.45 µm) which occur singly. Colonies appear moist, raised, white circles with clear smooth margins on CBA averaging 2 – 3 mm in size. On TSA strains have a dried-out brittle appearance which allows visual differentiation from other species of *Scandinavium*. Negative for lysine decarboxylase and citrate utilisation. Acid is produced from D-adonitol, D-arabitol and potassium 2-ketogluconate (API 20 and 50 CHB/E). Resistant to fusidic acid (Biolog Gen III). Variable for fermentation of rhamnose, dulcitol, acidification of phenol red, palatinose and production of malonate and α-glucosidase. Utilisation of the following carbon sources is variable: N-acetyl-D-galactosamine, mucic acid, D-saccharic acid and p-hydroxy-phenylacetic acid.

The DNA G + C content of the type strain is 53.9 mol %.

The type strain is H17S15<sup>T</sup> (= LMG 32613<sup>T</sup> = CCUG 76183<sup>T</sup>) and was isolated from *Quercus robur* rhizosphere soil in Hatchlands, Guildford, UK.

#### **Description of *Scandinavium tedordense* sp. nov.**

*Scandinavium tedordense* (te.dor.den'se. M.L. neut. n. *tedordense*, pertaining to Tedorde, the medieval name of Tidworth where the type strain was isolated).

The species shares the major characteristics of the genus. Gram negative motile rods (1 x 2.3 -2.5 µm) which occur singly. Colonies appear moist, raised, white circles with clear smooth margins on CBA averaging 1 – 2 mm in size. Positive for lysine decarboxylase. Acid is produced from rhamnose, D-adonitol, D-raffinose, D-turanose, D-fucose, L-fucose, D-arabitol and potassium 2-ketogluconate (API 20 and 50 CHB/E). Utilises the following carbon sources: β-

methyl-D-glucoside, N-acetyl-D-galactosamine, D-aspartic acid, D-serine, mucic acid, D-saccharic acid, D-lactic acid methyl ester and tween 40 (Biolog Gen III).

The DNA C + G content of the type strain is 53.9 %.

The type strain is TWS1a<sup>T</sup> (= LMG 32614<sup>T</sup>= CCUG 76188<sup>T</sup>) and was isolated from the bleeding lesion on a *Tilia x europaea* in Tidworth, Wiltshire, UK.

### **Emendation to the species *Scandinavium goeteborgense***

The description follows that of Marathe *et al.*, 2019 with the following additions.

Positive for lysine decarboxylase. Acid is produced from sorbitol and D-turanose (API 20/ 50 CHB/E). Acidifies phenol red (ID32). Utilises the following carbon sources:  $\beta$ -methyl-D-glucoside, N-acetyl-D-galactosamine, D-aspartic acid, D-serine, L-pyroglutamic acid, D-saccharic acid, tween 40,  $\gamma$ -amino-butyric acid,  $\alpha$ -hydroxy-butyric acid,  $\beta$ -hydroxy-D,L-butyric acid and acetoacetic acid. Resistant to fusidic acid. Variable for the fermentation of potassium 2-ketogluconate (Biolog Gen III).

The DNA C + G content of the type strain is 54.3 %.

The type strain is CCUG 66741<sup>T</sup> (= CECT 9823<sup>T</sup>= NCTC 14286<sup>T</sup>) and was isolated from a wound infection in Kungälv, Sweden.

### *Dryocola* protologs:

#### **Description of *Dryocola* gen. nov.**

*Dryocola* (Dry.o'co.la. Gr. fem. n. *drys*, an oak; L. suff. *-cola* (from L. masc. n. *incola*), inhabitant; N.L. masc. n. *Dryocola*, an inhabitant of oaks).

Gram-negative rods (0.96 – 1.34 x 1.87 – 2.49  $\mu$ m), facultatively anaerobic, oxidase negative and catalase positive. Cells occur singly, in pairs, groups of 3 and occasionally form chains, are motile by peritrichous flagella and can produce fimbriae. Colonies are cream coloured, with a darker convex centre and uneven margins on TSA. Growth is observed between 4 and 37 °C, optimum temperature is 30 °C but some strains are capable of growth at 41 °C. Strains can grow in 1 – 6 % supplemented salt and at a pH concentration of 6 – 8, with some growth at

pH 9. Positive for  $\beta$ -galactosidase and arginine dihydrolase. Negative for lysine decarboxylase, citrate utilisation, H<sub>2</sub>S, urease, tryptophan deaminase, indole production, acetoin production and gelatinase. Nitrate is reduced to nitrite. Acid is produced from: glucose, mannitol, rhamnose, amygdalin, L-arabinose, D-ribose, D-xylose, D-galactose, D-fructose, D-mannose, methyl- $\alpha$ D-glucopyranoside, N-acetylglucosamine, esculin ferric citrate, D-cellobiose, D-maltose, D-trehalose, gentiobiose and potassium gluconate (API 20 and 50CHB/E). Production of  $\beta$ -glucosidase, malonate and  $\alpha$ -glucosidase; acid is produced from phenol red (ID32). Utilise the following carbon sources: dextrin, D-salicin, N-acetyl-D-glucosamine, N-acetyl- $\beta$ -D-mannosamine, N-acetyl neuraminic acid, 3-methyl glucose, inosine, D-glucose-6-phosphate, D-fructose-6-phosphate, D-serine, glycyl-L-proline, L-alanine, L-aspartic acid, L-glutamic acid, L-histidine, L-serine, D-galacturonic acid, L-galactonic acid lactone, D-gluconic acid, D-glucuronic acid, glucuronamide, mucic acid, D-saccharic acid, methyl pyruvate, L-lactic acid, L-malic acid, acetoacetic acid and acetic acid. Resistant to 1% sodium lactate, fusidic acid, D-serine, troleandomycin, rifamycin, lincomycin, guanidine hydrochloric acid, niaproof 4, vancomycin, tetrazolium violet, tetrazolium blue, lithium chloride, aztreonam and sodium butyrate (Biolog Gen III). Variable for ornithine decarboxylase; fermentation of melibiose, glycerol, arbutin, salicin, D-lactose, D-raffinose, D-turanose, D-lyxose, L-fucose, D-arabitol, potassium 2-ketogluconate and potassium 5-ketogluconate; acidification of palatinose and production of  $\alpha$ -galactosidase. Utilisation of the following carbon sources are variable: sucrose,  $\beta$ -methyl-D-glucoside, D-aspartic acid, L-arginine, L-pyroglutamic acid, D-lactic acid methyl ester, citric acid, D-malic acid, bromo-succinic acid, tween 40 and  $\alpha$ -hydroxy-butyric acid and susceptibility to nalidixic acid. Major fatty acids are C<sub>12:0</sub>, C<sub>14:0</sub>, C<sub>16:0</sub>, C<sub>18:1</sub>  $\omega$ 7c, summed feature 2 (C<sub>14:0</sub> 3-OH and/or iso-C<sub>16:1</sub>) and summed feature 3 (C<sub>16:1</sub>  $\omega$ 7c and/or C<sub>16:1</sub>  $\omega$ 6c). The DNA G + C content ranges from 53.0 – 53.9 mol %.

The type species is *Dryocola boscaweniae*.

#### **Description of *Dryocola boscaweniae* sp. nov.**

*Dryocola boscaweniae* (bos.ca.we'ni.ae. N.L. gen. n. *boscaweniae*, of Boscawen, named to honour Lady Frances Boscawen, the first lady of the Hatchlands Park estate, Surrey, UK).

The description is as given for the genus with the following additions. Cells are short rods (0.8 – 1.3 x 1.8 – 2.4  $\mu\text{m}$ ), that occur singly, as pairs and in chains. Strains grow well between 4 – 41 °C and exhibit strong growth from pH 6 – 9. Acid is produced from melibiose, glycerol, D-adonitol, arbutin, salicin, D-lactose, D-raffinose, D-lyxose, D-arabitol, potassium 2-ketogluconate and potassium 5-ketogluconate (API 20 and 50 CHB/E). Positive for acidification of palatinose and production of  $\alpha$ -galactosidase (ID32). Utilises the additional carbon sources:  $\beta$ -methyl-D-glucoside, D-lactic acid methyl ester, tween 40 and  $\alpha$ -hydroxybutyric acid. Variable for acid production from D-turanose and galacturonate and utilisation of bromo-succinic acid.

The DNA G + C content of the type strain is 53.0 mol %.

The type strain is H6W4<sup>T</sup> (= CCUG 76177<sup>T</sup> = LMG 32610<sup>T</sup>) and was isolated from the rhizosphere of healthy oak from Hatchlands Park, Surrey, UK.

#### **Description of *Dryocola clanedunensis* sp. nov.**

*Dryocola clanedunensis* (*cla.ne.dun.en'sis*. M.L. masc. adj. *clanedunensis*, pertaining of Clanedun, the medieval name of Clandon where Hatchlands Park, the origin of isolation for the original strains, is located).

The description is as given for the genus with the following additions. Cells are longer rods (1 – 1.4 x 2.3 – 2.7  $\mu\text{m}$ ) that occur singly, in clusters of three and in chains. Strains grow well between 4 – 37 °C and exhibit strong growth at pH 6 – 8. Utilises the additional carbon sources: L-arginine, citric acid and bromo-succinic acid. Variable for ornithine decarboxylase, acid production from glycerol, arbutin, salicin, L-fucose and utilisation of sucrose,  $\beta$ -methyl-D-glucoside, D-aspartic acid, D-lactic acid methyl ester, D-malic acid and tween 40.

The DNA G + C content of the type strain is 53.8 mol %.

The type strain is H11S18<sup>T</sup> (= CCUG 76181<sup>T</sup> = LMG 32611<sup>T</sup>) and was isolated from the rhizosphere of oak suffering from AOD from Hatchlands Park, Surrey, UK.

*Leclercia* and *Silvania* protologs:

#### **Emendation description of the genus *Leclercia***

*Leclercia* (Leclerc' i.a. M.L. fem. n. *Leclercia* was named to honour H. Leclerc, a French bacteriologist, who first described and named this organism *Escherichia adecarboxylata* in 1962, and who made many other contributions to enteric bacteriology).

Gram-negative rods, ranging from 1.39 -1.54  $\mu\text{m}$  wide and 2.01 – 3.06  $\mu\text{m}$  long. All strains possess fimbriae and are motile by peritrichous flagella, and are oxidase negative, catalase positive, facultative anaerobes. After 48 h on TSA, all species appear as cream-coloured, circular, convex colonies between 2 - 3 mm in diameter with entire, slightly undulate margins. After longer periods of incubation some strains may develop a yellow diffusible pigment, although the conditions required are not consistent. Growth is observed from 10 – 41 °C for all strains, although some strains can grow at 4 °C, with optimal growth observed between 30 – 35 °C. The majority of strains grow at pH 6 – 9 and at supplemented salt concentrations of 1 – 8 %, with some strains only able to grow up to 7 %. Positive for  $\beta$ -galactosidase and indole production. Negative for arginine dihydrolase, lysine decarboxylase, ornithine decarboxylase, citrate utilization, H<sub>2</sub>S production, urease, tryptophan deaminase, acetoin production (VP) and gelatinase. Nitrite is reduced to nitrate. Production of  $\beta$ -glucosidase and  $\alpha$ -galactosidase, acidification of galacturonate and phenol red, (ID 32). Resistant to 1% sodium lactate, fusidic acid, D-serine, rifamycin, lincomycin, guanidine HCl, niaproof 4, vancomycin, tetrazolium violet, tetrazolium blue, lithium chloride, aztreonam and sodium butyrate (Biolog Gen III).

The major fatty acids are C<sub>16:0</sub>, C<sub>18:1</sub>  $\omega$ 7c) and summed feature 3 (C<sub>16:1</sub>  $\omega$ 7c and/or C<sub>16:1</sub>  $\omega$ 6c).

The DNA G + C content ranges from 55.8 – 56.4 mol %.

The type species is *Leclercia adecarboxylata*.

#### **Emendation description of *Leclercia adecarboxylata***

The description is as given above for the genus with the following additional characteristics.

In addition to the carbon sources listed in Table S4, acid is produced from D-adonitol, D-arabitol and potassium 2-ketogluconate; and D-salicin, D-aspartic acid and tween 40 are utilised. Variable for the fermentation of saccharose, dulcitol, D-raffinose and D-lyxose; the acidification of palatinose and the production of malonate. Utilisation of the following carbon

sources is variable: stachyose, L-pyroglutamic acid, pectin, D-malic acid and  $\alpha$ -hydroxy-butyric acid. Variable resistance to nalidixic acid is observed.

The DNA G + C content of the type strain is 55.8 mol %.

The type strain is *Leclercia adecarboxylata* (ATCC 23216; CIP 82.92; DSM 30081; DSM 5077; HAMBI 1696; JCM 1667; LMG 2803; NBRC 102595; NCTC 13032).

#### **Description of *Leclercia tamurae* sp. nov.**

*Leclercia tamurae* (ta.mu'rae. N.L. gen. masc. n. *tamurae*, of Tamura, named in honour of Kazumichi Tamura for his role in defining the genus *Leclercia*).

The description is as given above for the genus with the following additional characteristics.

After 48 h on TSA, colonies are circular, matte, brittle and cream-coloured with slightly undulate margins with an average diameter of 3 mm. All strains are capable of forming the yellow pigmentation associated with *Leclercia*, although not within a set timeframe.

In addition to the carbon sources listed in Table S4, acid is produced from sorbitol and dulcitol and acidification of malonate is observed. Variable features include the fermentation of methyl- $\alpha$ -D-glucopyranoside, D-trehalose, D-lyxose and potassium 5-ketogluconate, and the production of *N*-acetyl- $\beta$ -glucosaminidase. Utilisation of the following carbon sources is variable: sucrose, D-salicin, 3-methyl glucose, D-serine, L-pyroglutamic acid, citric acid, D-malic acid, tween 40,  $\alpha$ -hydroxy-butyric acid,  $\beta$ -hydroxy-D, L-butyric acid and formic acid. Variable resistance to troleandomycin and potassium tellurite is observed.

The DNA G + C content of the type strain is 56.4 mol %.

The type strain is H6S3<sup>T</sup> (= LMG 32609<sup>T</sup> = CCUG 76176<sup>T</sup>) and was isolated from healthy *Quercus robur* rhizosphere soil in Hatchlands, Guildford, UK.

#### **Description of *Silvania* gen. nov**

*Silvania* (Sil.va'ni.a. N.L. fem. n. *Silvania*, named after Silvanus the Roman deity of woodlands).

Gram-negative, straight rods (1.2 – 1.4 x 1.6 – 2.0 µm) and motile by peritrichous flagella. Cells appear singly or in pairs. Oxidase negative, catalase positive facultative anaerobes. Colonies appear as cream-coloured, convex circles with raised entire margins and a diameter of 3 – 4 mm on TSA. Growth is observed between 4 – 37 °C with an optimum growth temperature of 30 °C. Positive for β-galactosidase, negative for arginine dihydrolase, lysine decarboxylase, ornithine decarboxylase, citrate utilization, H<sub>2</sub>S production, urease, tryptophan deaminase, acetoin production and gelatinase. Nitrite is reduced to nitrate. Positive for the acidification of galacturonate and production of β-glucosidase and α-galactosidase (ID 32). Resistance to 1% sodium lactate, rifamycin, lincomycin, guanidine HCl, niaproof 4, vancomycin, tetrazolium violet, tetrazolium blue, lithium chloride, aztreonam and sodium butyrate is observed. Variable features of the genus include indole production; fermentation of rhamnose, saccharose, methyl-α-D-mannopyranoside, methyl-α-D-glucopyranoside, D-lyxose, D-tagatose; acidification of phenol red and production of β-glucuronidase and malonate. Utilisation of the following carbon sources is variable: stachyose, N-acetyl-D-galactosamine, fusidic acid, D-serine, L-pyroglutamic acid, pectin, quinic acid, α-keto-glutaric acid and D-malic acid. Variable resistance to troleandomycin, nalidixic acid and potassium tellurite is observed. The major fatty acids are C<sub>16:0</sub>, C<sub>18:1</sub> ω7c and summed feature 3 (C<sub>16:1</sub> ω7c and/or C<sub>16:1</sub> ω6c).

The DNA G + C content ranges from 55.7 to 55.9 mol %.

The type species is *Silvania hatchlandensis*.

### **Description of *Silvania hatchlandensis***

*Silvania hatchlandensis* (hatch.lan.den'sis. N.L. fem. adj. *hatchlandensis*, pertaining to Hatchlands the national park in Guildford, UK where the strains were isolated from).

The description is as given above for the genus with the following additional characteristics.

Cells are on average 1.25 x 1.94 µm in size. After 48 h on TSA, the colonies appear as slightly raised circles with raised entire margins and an average diameter of 4 mm. Positive for indole production (API 20 and API 50 CHB/E), the acidification of phenol red and the production of

$\beta$ -glucuronidase (ID 32). Variable features of the species include the fermentation of methyl- $\alpha$ -D-mannopyranoside and methyl- $\alpha$ -D-glucopyranoside; the production of malonate. In addition to the carbon sources listed in Table S4, *N*-acetyl-D-galactosamine, L-pyroglutamic acid, quinic acid,  $\alpha$ -keto-glutaric acid and D-malic acid are utilised.

The DNA G + C content of the type strain is 55.9 mol %

The type strain is H19S6<sup>T</sup> (= LMG 32608<sup>T</sup> = CCUG 76185<sup>T</sup>) and was isolated from diseased *Quercus robur* rhizosphere soil in Hatchlands, Guildford, UK.

### **Description of *Silvania confinis***

*Silvania confinis* (con.fi'nis. L. fem. adj. *confinis*, adjoining/akin, referring to the close phylogenetic relationship to the type species of the genus).

The description is as given above for the genus with the following additional characteristics.

Cells are on average 1.37 x 1.68  $\mu$ m in size. After 48 h on TSA, the colonies appear as slightly raised circles with raised entire margins and an average diameter of 3 mm. In addition to the carbon sources listed in Table S4, acid is produced from D-lyxose and D-tagatose. Resistance to D-serine, troleandomycin, nalidixic acid and potassium tellurite is observed.

The DNA G + C content of the type strain is 55.7 mol %.

The type strain is H4N4<sup>T</sup> (= LMG 32607<sup>T</sup> = CCUG 76175<sup>T</sup>) and was isolated from healthy *Quercus robur* rhizosphere soil in Hatchlands, Guildford, UK.

**Suppl. Table. S5.** The hyperlink to the interactive krona plots for the 16S rRNA data collected for each sample from Hatchlands park.

Samples	Hyperlink	Repository citation
<p>All samples divided by, location, health status and individual tree number have been placed in the depository with links to each krona plot.</p>	<p><a href="http://researchdata.uwe.ac.uk/702">http://researchdata.uwe.ac.uk/702</a></p>	<p>Maddock, D. TAXONOMIC AND GENOMIC INVESTIGATION OF SOIL MICROORGANISMS ASSOCIATED WITH ACUTE OAK DECLINE. <i>UWE data repository</i> [online]. Available from:  <a href="http://researchdata.uwe.ac.uk/702">http://researchdata.uwe.ac.uk/702</a>                      [Accessed 08 August 2023].</p>