Evolution of host specificity of *Pseudomonas syringae* on *Prunus*



Michelle Teresa Hulin
School of Biological Sciences
University of Reading

A thesis submitted for the degree of Doctor of Philosophy

March 2017

Abstract

Pseudomonas syringae is a globally important plant pathogen, that includes pathovars that infect over 180 plant species. Individual pathovars are highly specialised and only infect one or a few hosts. *Pseudomonas syringae* uses a range of type III effector proteins to cause disease. However, plants have evolved to be able to detect these effectors and trigger immunity. Therefore, it is believed that a strain's repertoire of effectors dictates its host range and genetic alteration of these repertoires enables host range expansion. This topic was explored using comparative genomics of three divergent clades that have convergently evolved to cause bacterial canker on cherry (*Prunus avium*). The clades include *P. syringae* pv. *morsprunorum* (*Psm*) (which is differentiated into two races based upon host response) and *P. syringae* pv. *syringae* (*Pss*). Three reference isolates of *Psm* R1, R2 and *Pss* were sequenced with PacBio and the genomes of a further fifteen isolates were sequenced using the Illumina MiSeq system. The virulence of strains was critically assessed in both lab- and field-based pathogenicity tests. Comparative genomic analysis of the *Prunus* strains revealed highly divergent effector and toxin repertoires within and between the different clades, indicating that they may each utilise distinct mechanisms to cause the same disease. Horizontal gene transfer of effector genes has occurred frequently between clades, often through the movement of plasmids. Several effectors, whose evolution is significantly associated with pathogenicity on *Prunus* were identified. Effectors that have been lost or putatively pseudogenised in cherry-infecting clades were cloned and expressed in pathogenic strains to determine if they are avirulence factors in cherry. The expression of two effectors (HopC1 and HopAB) was found to trigger a hypersensitive response on cherry leaves and reduce pathogen population growth. This work provides an insight into the convergent evolution of pathogenicity and mechanisms controlling the host specificity of bacteria.

Declaration

I confirm that this is my own work and the use of all materials from other sources has been properly and fully acknowledged.

Michelle Teresa Hulin

Acknowledgements

I am extremely grateful to many people for the help and support provided during my PhD studies. Firstly, I would like to thank my three supervisors Richard Harrison, John Mansfield and Robert Jackson for excellent guidance throughout my PhD. It has been a brilliant experience to work with and be inspired by these three scientists. In addition, thank you to Laura Lewis for early supervision, helping kick start my PhD.

Thank you to Phil Brain, Simon Hodge, Richard Everitt and Andrew Armitage for their excellent support with statistical analysis and bioinformatics. Although a complete beginner of bioinformatics at the start of PhD, I have grown to extremely enjoy this aspect of research. Many thanks to Helen Bates for sequencing several bacterial genomes on the MinIon. Finally, thank you to Ian Brown for providing the opportunity to see the bacteria in action using electron microscopy.

I am grateful to Steve Roberts, David Guttman, Helen Neale and Mateo San José for providing bacterial strains that have been used in the thesis.

The time spent at East Malling during my PhD has been wonderful. I would like to thank all members of the Harrison Lab and everyone at East Malling for their friendship. A special mention to Emma Cascant López who has been great company, on the drive to work, in the office, laboratory and at conferences.

I would like to thank my loving family for all their encouragement during my PhD. Thank you most of all to my partner Luke, who has provided constant support through the ups and downs of PhD life. He could probably now write his own thesis on *Pseudomonas syringae* from spending so much time in my company!

Finally, I am grateful to the University of Reading and East Malling Trust for funding my PhD studies.

Contents page

Abstract ……i

Chapter 1: General Introduction

1.1 Introduction to plant disease and immunity 1

1.2 Evolution of pathogenicity and host specificity 5

1.3 The plant pathogen *Pseudomonas syringae* 10

1.4 Virulence factors in addition to proteinaceous effectors 11

1.5 *P. syringae* as a specialist 13

1.6 The *Pseudomonas syringae* – cherry pathosystem 16

1.7 Thesis Aims 20

Chapter 2: General materials and methods

2.1 Maintenance of bacterial cultures 21

2.2 Maintenance of plant material 21

2.3 DNA manipulations 22

Chapter 3: Characterisation of the *P. syringae* – *Prunus* interaction with a range of field and laboratory-based assays

3.1 Abstract 23

3.2 Introduction 24

3.3 Methods 27

3.3.1 Bacterial strains 27

3.3.2 Genome sequencing 31

3.3.3 Phylogenetic analysis 31

3.3.4 Plant material 32

3.3.5 Pathogenicity assays 32

3.3.5.1 Whole-tree inoculations 32

3.3.5.2 Experimental design of whole-tree inoculations 33

3.3.5.3 Cut shoot inoculations 33

3.3.5.4 Cherry fruit inoculations 34

3.3.5.5 Leaf inoculations 34
3.3.6 ROS and callose assays 35
3.3.7 Electron microscopy 36
3.3.8 Statistical analysis 37
3.4 Results 38
3.4.1 Phylogenetics 38
3.4.2 Whole-tree glasshouse experiment 40
3.4.3 Whole-tree field experiment 42
3.4.4 Laboratory-based pathogenicity assays 47
3.4.4.1 Cut shoot inoculations 47
3.4.4.2 Inoculation of detached immature cherry fruits 50
3.4.4.3 Inoculation of detached leaves 52
3.4.4.5 Suitability of leaves for resistance screening 63
3.4.5 Callose and ROS detection 64
3.4.6 Electron microscopy 69
3.4.7 Comparison of pathogenicity tests 71
3.5 Discussion 73
3.5.1 Whole-tree tests revealed differences in pathogenicity and host susceptibility to *P. syringae* 74
3.5.2 Comparison of laboratory-based assays 76
3.5.3 Timing of symptom development on cherry leaves is indicative of pathogenicity or the HR 78
3.5.4 Conclusions 80

Chapter 3 supplementary data 81

Chapter 4: Comparative genomics of *P. syringae* to explore the evolution of host specificity for cherry (*Prunus avium*)

4.1 Abstract 88
4.2 Introduction 89
4.3 Materials and methods 92
4.3.1 Bacterial strains 92

4.3.2 Plasmid profiling 97
4.3.3 Genome sequencing, assembly and annotation 97
4.3.3.1 Illumina 97
4.3.3.2 Pacbio and MinIon sequencing 98
4.3.4 Phylogenetic analysis of *P. syringae* 99
4.3.5 Whole-genome alignment 100
4.3.6 Effector identification 100
4.3.7 Gain and loss analysis 101
4.3.8 BayesTraits 101
4.3.9 Phylogenetic analysis of individual effector genes 101
4.3.10 Effector genome region alignments 102
4.3.11 Effector domain identification 102
4.3.12 Search for novel effector genes 102
4.3.13 Identification of genomic islands 103
4.3.14 Orthology analysis 103
4.3.15 Visualisation of genomes using Circos 104
4.4 Results 104
4.4.1 Genome statistics and plasmid profiles 105
4.4.2 Phylogenetics 109
4.4.2.1 MLST phylogeny 109
4.4.2.2 Core-genome phylogeny 111
4.4.2.3 Reference sequence alignment phylogeny 112
4.4.3 Search for virulence factors 114
4.4.3.1 The *hrp* pathogenicity island 114
4.4.3.2 Type III effectors 117
4.4.3.3 Number of effectors 120
4.4.3.4 Lack of a redundant effector group is linked to host specificity 121
4.4.4 Gain and loss analysis of effector genes 126
4.4.5 Association of effector families with cherry pathogenicity 133
4.4.6 Origins of key effectors in cherry pathogens 137
4.4.7 Search for novel effectors 141
4.4.8 Other virulence factors 143

4.4.9 Orthology analysis 146
4.4.9.1 Identification of orthogroups on branches leading to cherry pathogens 146
4.4.10 Mobility of virulence components 152
4.4.11 Genomic architecture of virulence-associated genes 165
4.4.12 Integrating results into an evolutionary model of pathogenicity ...169
4.5 Discussion 171
4.5.1 Phylogenetic analysis 171
4.5.2 Search for candidate effectors involved in cherry pathogenicity 173
4.5.3 Horizontal transfer has been important in the acquisition of key effectors 177
4.5.4 Other virulence factors important for cherry pathogenicity 178
4.5.5 Linking genomics to host specialisation 181
4.5.6 Conclusions 183

Chapter 4 supplementary data 185

Chapter 5: Cloning of candidate virulence and avirulence effector genes involved in host specificity of *P. syringae* on cherry

5.1 Abstract 239
5.2 Introduction 240
5.3 Methods 245
5.3.1 Bacterial strains 245
5.3.2 Candidate gene identification 250
5.3.3 General DNA manipulations 250
5.3.4 *P. syringae* transformations 255
5.3.4.1 Conjugation 255
5.3.4.2 Electroporation 256
5.3.5 PCR validation 256
5.3.6 Genomic libraries 257
5.3.6.1 Creation of genomic libraries 257
5.3.6.2 Conjugation of genomic libraries into *P. syringae* 257

5.3.7 Pathogenicity assays ……258
5.3.7.1 Plant material ……258
5.3.7.2 Population counts ……258
5.3.7.3 Symptom scoring ……259
5.3.7.4 Screening of genomic libraries for virulence phenotypes ……259
5.3.7.5 Tobacco HR assay ……260
5.3.8 Statistical analysis ……260
5.4 Results ……260
5.4.1 Identification of candidate avirulence effectors ……261
5.4.2 Two effector genes act as avirulence factors in cherry ……267
5.4.3 Identification of candidate virulence genes ……282
5.4.4 Single effector deletions ……284
5.4.5 Genomic library screening ……287
5.5 Discussion ……288
5.5.1 Identification of two avirulence effectors on cherry ……288
5.5.2 Alleles of HopAB vary in ability to trigger immune responses in cherry ..292
5.5.3 Deletion studies of candidate virulence genes failed to alter the virulence phenotype ……294
5.5.4 Genomic library screening ……295
5.5.5 Conclusions ……296

Chapter 5 supplementary data ……298

Chapter 6: General Discussion ……309

Bibliography ……321

Appendix ……340

List of Tables

Table 1.1 Mobile elements that influence the evolution of *P. syringae* 8
Table 1.2 Known plant cell targets of *P. syringae* type III effectors 12
Table 3.1 Bacterial strains used in this study with host of isolation and source 28
Table 3.2: Heatmap constructed using ANOVA/REML means for each strain x inoculation method combination on four different cherry cultivars 71
Table 4.1: List of bacterial strains used in this study 91
Table 4.2: Assembly statistics for all sequenced strains 104
Table 4.3: List of effectors in PacBio/Minion genomes 117
Table 4: Effector presence comparisons between cherry pathogenic and non-pathogen Psm R1 strains 119
Table 4.5: List of putative horizontal gene transfer events that have occurred between *Prunus*-infecting clades within *P. syringae* 136
Table 4.6: Novel candidate effector genes in sequenced *Prunus* strains 140
Table 4.7: Genomic islands predicted in the *Psm* R1 5244 151
Table 4.8: Genomic islands predicted in the *Psm* R2 leaf 154
Table 4.9: Genomic islands predicted in the *Pss* 9097 158
Table S4.1: Effectors identified in the genome assemblies of *Prunus*-isolates and RMA1 sequenced in this study 181
Table 5.1: All *P. syringae* strains and mutants used in this study 242
Table 5.2: All Plasmids generated during this study 244
Table 5.3: Primer sequences used in this study for cloning genes into the pBBR1MCS-5 broad-host range expression vector 247
Table 5.4: Primer sequences used in this study to create pK18-mobsacB-based deletion vectors 248

List of Figures

Figure 1.1 Overview of the plant immune system 4
Figure 1.2 Bayesian phylogenetic trees of *P. syringae* pathovars and host plant species 15
Figure 1.3 The perennial lifecycle of bacterial canker disease on *Prunus* 19

Figure 3.1: Bayesian phylogenetic tree of *P. syringae* 39
Figure 3.2: Evaluation of symptoms caused by different *P. syringae* strains on cherry cv. Van trees 41
Figure 3.3: Evaluation of symptoms caused by *P. syringae* on different cherry cultivars in the field 44
Figure 3.4: Evaluation of symptoms caused by *P. syringae* on different plum cultivars in the field 46
Figure 3.5: A cutshoot assay using different inoculated strains of *P. syringae* on cherry and plum cultivars revealed quantitative symptom differences 49
Figure 3.6: Diameter of necrosis caused by different *P. syringae* strains on immature green cherry fruits 51
Figure 3.7: Diameter of necrosis caused by cherry pathogens on four cherry cultivars using immature green cherry fruits 52
Figure 3.8: Symptoms observed in detached cherry leaves using different leaf inoculation methods 53
Figure 3.9: Population counts of different strains over time on cherry cv. Van (A) and plum cv. Victoria (B) leaves 55
Figure 3.10: Day 10 population count of all strains used in this study on cherry cv. Van leaves 57
Figure 3.11: Symptom development over time after inoculation of various *P. syringae* strains in cherry cv. Van at different concentrations 59
Figure 3.12: Pathogenicity of different strains, assessed by population counts and symptom scores, on cherry and plum leaves 61

Figure 3.13: Images of symptom development overtime on cherry and plum. A: Cherry cv. Van, B: Plum cv. Victoria 62
Figure 3.14: Day 10 population counts of three pathogenic *P. syringae* strains on different cherry cultivars 64
Figure 3.15: Representative images from callose staining experiments of cherry and tobacco leaves 66
Figure 3.16: ROS staining of leaf inoculation sites 68
Figure 3.17: TEM images of *Psm* R2-leaf in a detached cherry leaf one week after inoculation 70
Figure 3.18: Comparison of different inoculation methods.....72

Figure S3.1: Diagram of randomisation of *P. syringae* strains across two trees in field experiment 79
Figure S3.2: Diagram depicting experimental design of population counts 81
Figure S3.3: Population counts of different strains 0 dpi on cherry cv. Van (A) and plum cv. Victoria (B) leaves (for Figure 3.9) 82
Figure S3.4: Day 0 population count of all strains used in this study on cherry cv. Van leaves (for Figure 3.10) 83
Figure S3.5: Day 0 population counts on cherry cv. Van and plum cv. Victoria (for Figure 3.12) 84
Figure S3.6: Day 0 population counts of three pathogenic *P. syringae* strains on different cherry cultivars (for Figure 3.14) 85

Figure 4.1: Plasmid profiles of all *Prunus*-infecting sequenced isolates and some out-groups for comparison 106
Figure 4.2: Bayesian phylogeny based on seven MLST house-keeping genes 108
Figure 4.3: ML phylogeny based on 902 genes which represent the core genome of *P. syringae* 110
Figure 4.4: ML phylogeny based on the alignment generated using RealPhy 111
Figure 4.5: Organisation of the structural T3SS and conserved effector locus (CEL) of the three PacBio sequenced pathogens 113
Figure 4.6: Organisation of the CEL in *Psm* R1 114

Figure 4.7: Effector presence and absence heatmaps for *Prunus*-infecting strains and their close out-groups 116
Figure 4.8: Barplot showing the number of putative effector genes identified in each *Prunus*-infecting strain 118
Figure 4.9: Alignment of the DNA regions surrounding the *avrPto1* (A) and *hopAB1* (B) 120
Figure 4.10: Truncation of *hopAB3* in *Psm* R2 122
Figure 4.11: Some pathovars possess a truncated version of the *hopAB1* gene 123
Figure 4.12: Heatmap showing the presence and absence of the *avrPto*/*hopAB* REG across the *P. syringae* complex 124
Figure 4.13: Phylogenetic trees with the gain and loss of effector genes occurring on each branch leading to cherry pathogenic strains 126
Figure 4.14: Phylogenetic trees with the gain and loss of effector genes occurring on each branch leading to cherry pathogenic *Pss* strains 127
Figure 4.15: Phylogenetic trees with the gain and loss of effector genes occurring on each branch leading to cherry pathogenic strains 129
Figure 4.16: Phylogenetic trees with the gain and loss of effector genes occurring on each branch leading to cherry pathogenic *Pss* strains 130
Figure 4.17: Barplot showing the output from BayesTraits analysis of effector gene and pathogenicity evolution using the core genome phylogeny 132
Figure 4.18: Barplot showing the output from BayesTraits analysis of effector gene and pathogenicity evolution using the RealPhy phylogeny 134
Figure 4.19: Genomic locations of the *hopAR1* gene in the three PacBio sequenced pathogens 138
Figure 4.20: Heatmap showing the presence and absence of other virulence factors in the *Prunus*-infecting strains 141
Figure 4.21: Heatmap showing the presence and absence of several phytotoxin biosynthesis clusters in phylogroup 2 strains 142
Figure 4.22: Scatterplot with line of correlation of the number of putative effector genes and toxin biosynthesis clusters for each *Prunus*-infecting strain 143

Figure 4.23: Phylogenetic trees with the gain and loss of particular orthogroups occurring on each branch leading to cherry pathogenic strains. This was based on the GLOOME analysis of the core genome phylogeny 145
Figure 4.24: Phylogenetic trees with the gain and loss of particular orthogroups occurring on each branch leading to cherry pathogenic *Pss* strains, using the GLOOME analysis of the core genome phylogeny 146
Figure 4.25: Phylogenetic trees with the gain and loss of particular orthogroups occurring on each branch leading to cherry pathogenic strains, using the GLOOME analysis of the RealPhy phylogeny 148
Figure 4.26: Phylogenetic trees with the gain and loss of particular orthogroups occurring on each branch leading to cherry pathogenic *Pss* strains. This was based on the GLOOME analysis of the Realphy phylogeny 149
Figure 4.27: Heatmap showing the presence and absence of the genomic islands identified in the PacBio sequenced pathogens of (A) *Psm* R1, (B) *Psm* R2 and (C) *Pss* across the *P. syringae* complex 162
Figure 4.28: Circular visualisation of the genome of *Psm* R1-5244 produced using Circos 164
Figure 4.29: Circular visualisation of the genome of *Psm* R2-leaf produced using Circos 165
Figure 4.30: Circular visualisation of the genome of *Pss* 9097 produced using Circos 166
Figure 4.31: Gain and loss of genes statistically associated with cherry pathogenicity on the core genome phylogeny.....170

Figure S4.1: Whole-genome alignment of R1-5244 and R1-5300 using ProgressiveMauve 199
Figure S4.2: ML phylogeny based on 846 genes from 104 strains, which represent the core genome of *P. syringae* 200
Figure S4.3: Effector presence and absence heatmaps for all analysed *P. syringae* strains 201
Figure S4.4: Heatmap of average nucleotide identity (ANI) in core genome alignment of 104 *P. syringae* strains 202

Figure S4.5: ML phylogenetic tree based on the *avrD1* gene generated using FastTree 203
Figure S4.6: Alignment of the DNA region surrounding the *avrD1* gene in *Prunus* strains and out-groups 204
Figure S4.7: ML phylogenetic tree based on the *avrRps4* gene generated using FastTree 205
Figure S4.8: Alignment of the DNA region surrounding the *avrRps4* gene in *Prunus* strains and out-groups 206
Figure S4.9: ML phylogenetic tree based on the *hopAF1* gene generated using FastTree 207
Figure S4.10: Alignment of the DNA region surrounding the *hopAF1* gene in *Prunus* strains 208
Figure S4.11: ML phylogenetic tree based on the *hopAO1* gene family generated using FastTree 209
Figure S4.12: Alignment of the DNA region surrounding the *hopAO1* gene in two *Prunus* strains 210
Figure S4.13: ML phylogenetic tree based on the *hopAR1* gene generated using FastTree 211
Figure S4.14: Alignment of the DNA region surrounding the *hopAR1* gene in *Prunus* strains and out-groups 212
Figure S4.15: ML phylogenetic tree based on the *hopAT1* gene generated using FastTree 213
Figure S4.16: Alignment of the DNA region surrounding the *hopAT1* gene in *Prunus* strains and out-groups 214
Figure S4.17: ML phylogenetic tree based on the *hopAU1* gene generated using FastTree 215
Figure S4.18: Alignment of the DNA region surrounding the *hopAU1* gene in *Prunus* strains and an out-group 216
Figure S4.19: ML phylogenetic tree based on the *hopAV1* gene generated using FastTree 217
Figure S4.20: Alignment of the DNA region surrounding the *hopAV1* gene in *Prunus* strains 218

Figure S4.21: ML phylogenetic tree based on the *hopAY1* gene generated using FastTree 219
Figure S4.22: Alignment of the DNA region surrounding the *hopAY1* gene in *Prunus* strains 220
Figure S4.23: ML phylogenetic tree based on the *hopBB1* gene generated using FastTree 221
Figure S4.24: Alignment of the DNA region surrounding the *hopBB1* gene in *Prunus* strains 222
Figure S4.25: ML phylogenetic tree based on the *hopBD1* gene generated using FastTree 223
Figure S4.26: Alignment of the DNA region surrounding the *hopBD1* gene in *Prunus* strains and some out-groups 224
Figure S4.27: ML phylogenetic tree based on the *hopBF1* gene generated using FastTree 225
Figure S4.28: Alignment of the DNA region surrounding the *hopBF1* gene in *Prunus* strains and some out-groups 226
Figure S4.29: ML phylogenetic tree based on the *hopD1* gene generated using FastTree 227
Figure S4.30: Alignment of the DNA region surrounding the *hopD1* gene in *Prunus* strains and an out-group 228
Figure S4.31: ML phylogenetic tree based on the *hopE1* gene generated using FastTree 229
Figure S4.32: Alignment of the DNA region surrounding the *hopE1* gene in *Prunus* strains and an out-group 230
Figure S4.33: ML phylogenetic tree based on the *hopO* gene family generated using FastTree 231
Figure S4.34: ML phylogenetic tree based on the *hopT* gene family generated using FastTree 232
Figure S4.35: Alignment of the DNA region surrounding the *hopO1* and *hopT1* gene in Prunus strains isolated from cherry and apricot 233
Figure S4.36: ML phylogenetic trees of convergently acquired *Pss* effector genes. A: *avrRpm1*, B: *hopAW1*, C: *hopAR1* 234

Figure 5.1: Heatmap of effector gene presence across *P. syringae* strains with candidate avirulence genes highlighted 258
Figure 5.2: AvrRps4 alignment 260
Figure 5.3: HopAW1 alignment 262
Figure 5.4: Time-course of *in planta* population growth of Psm R2-leaf wildtype, Δ*hrpA* mutant and R2-leaf expressing candidate avirulence effector genes 264
Figure 5.5: Ten-day population counts of cherry pathogens (R1-5244, R2-leaf and syr9644) expressing candidate avirulence genes 266
Figure 5.6: Ten-day population counts of cherry pathogens (R1-5244, R2-leaf and syr9644) expressing different alleles of *hopAB* 268
Figure 5.7: Protein alignment of members of the HopAB effector family 269
Figure 5.8: Symptom development in cherry leaves of *P. syringae* strains expressing different effectors 271
Figure 5.9: Symptom development of pathogens expressing *hopAB* alleles and *hopC1* in cherry leaves 274
Figure 5.10: Ten-day population counts of avirulence gene deletion mutants and wildtype strains in cherry leaves for the three non-pathogens of cherry (A: R1-5300, B: R1-9657 and C: avellBPIC631) 276
Figure 5.11: Ten-day population counts of *hopAB3* gene deletion *Psm* R2-leaf mutants 277
Figure 5.12: Heatmap of effector presence across *P. syringae* with candidate virulence genes highlighted. The heatmap was generated using the R gplots library 279
Figure 5.13: Ten-day population counts of *Psm* R2-leaf and syr9644 virulence gene deletion mutants compared to the wildtype strains 280
Figure 5.14: Ten-day population counts of all *Psm* R2-leaf single, double and triple virulence gene deletion mutants and the wildtype 282

Figure S5.1: Representative images of the tobacco hypersensitive response assay for R1-5244 and R2-leaf Δ*hrpA* mutants 294
Figure S5.2: PCR products of an M13-PCR to check the insert size of all pBBR1MCS-5 vectors generated in this study 295

Figure S5.3: PCR products of an M13-PCR to check the insert size of all pK18mobsacB vectors generated in this study 296
Figure S5.4: PCR products to validate all deletions performed in this study 298
Figure S5.5: Restriction digest of 10 genomic library clones from the RMA1 (A) and R1-5300 (B) libraries 299
Figure S5.6: 0-day population counts of cherry pathogens (R1-5244, R2-leaf and syr9644) expressing candidate avirulence genes (for Figure 5.5) 300
Figure S5.7: 0-day population counts of cherry pathogens (R1-5244, R2-leaf and syr9644) expressing different alleles of *hopAB* (for Figure 5.6) 301
Figure S5.8: 0-day population counts of avirulence gene deletion mutants and wildtype strains in cherry leaves for the three non-pathogens of cherry (A: R1-5300, B: R1-9657 and C: avellBPIC631) (for Figure 5.10) 302
Figure S5.9: 0-day population counts of *hopAB3* gene deletion *Psm* R2-leaf mutants (for Figure 5.11) 303
Figure S5.10: 0-day population counts of *Psm* R2-leaf and syr9644 virulence gene deletion mutants compared to the wildtype strains (for Figure 5.13) 304
Figure S5.11: 0-day population counts of all *Psm* R2-leaf single, double and triple virulence gene deletion mutants and the wildtype (for Figure 5.14) 305

List of abbreviations

ADP-RT	ADP-Ribosyl Transferase
ALVR	Agar Low Viscosity Resin
ANI	Average Nucleotide Identity
ANOVA	Analysis Of Variance
Avr	Avirulence
BLAST	Basic Local Alignment Search Tool
CAB	Sodium Cacodylate Buffer
CEL	Conserved Effector Locus
CFU	Colony-Forming Units
CTAB	Cetyl Trimethylammonium Bromide
cv.	Cultivar
DAB	3,3'-Diaminobenzidine
DAMP	Damage-Associated Molecular Pattern
DNA	Deoxyribonucleic Acid
ETI	Effector-Triggered Immunity
ETS	Effector-Triggered Susceptibility

FASTA	FAST-All
GI	Genomic Island
GLOOME	Gain Loss Mapping Engine
GTR	Generalised Time Reversible
HGT	Horizontal Gene Transfer
Hop	Hrp Outer Protein
HR	Hypersensitive Response
Hrp	Hypersensitive Response and Pathogenicity Protein
ICE	Integrative Conjugative Element
KBA	King's B Agar
LB	Luria Broth
LBA	Luria Agar
MAPK	Mitogen-Activated Protein Kinase
MCS	Multiple Cloning Site
ME	Mobile Element
ML	Maximum Likelihood
MLST	Multi-Locus Sequence Typing
NB-LRR	Nucleotide-Binding Leucine-Rich Repeat
OD	Optical Density
orthoMCL	Ortho-Markov Cluster Algorithm
PacBio	Pacific Biosciences
PAMP	Pathogen-Associated Molecular Pattern
PCR	Polymerase Chain Reaction
PHAST	Phage Search Tool
PRR	Pattern-Recognition Receptor
Psm	*P. syringae* pv. *morsprunorum*
Pss	*P. syringae* pv. *syringae*
PTI	PAMP-Triggered Immunity
pv.	Pathovar
Quake	Quality Aware Detection and Correction of Sequencing Errors
Quast	Quality Assessment Tool for Genome Assemblies
r(SAP)	Shrimp-Alkaline Phosphorylase
RAxML	Randomized Axelerated Maximum Likelihood
RealPHY	Reference Alignment Based Phylogeny Builder
REG	Redundant Effector Group
REML	Restricted Maximum Likelihood
REP	Repetitive Element Sequence-Based
RLK	Receptor-Like Kinase
RNA	Ribonucleic Acid
ROS	Reactive Oxygen Species
SMRT	Single Molecule Real-Time
SPAdes	St. Petersburg Genome Assembler
T3E	Type Three Effector

T3SS	Type Three Secretion System
Tukey-HSD	Tukey-Honest Significant Difference
UV	Ultra-Violet
Vir	Virulence
X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside

Chapter 1: General introduction

1.1 Introduction to plant disease and immunity

Plant disease accounts for an estimated 10-16% of global crop losses annually, amounting to a cost of US$220 billion (Chakraborty & Newton 2011). In recent decades, the increased global movement of plant material has led to devastating emerging diseases, particularly of woody plant species (Ghelardini *et al.* 2016). The emergence of major diseases such as ash dieback, kiwifruit canker and the current *Xylella* outbreaks occuring in Europe, have highlighted the need for fundamental studies of wood-infecting pathogens. The study of plant pathogen biology has revolutionised agriculture. By understanding the ecological, evolutionary and molecular processes that underlie plant-pathogen interactions, resistant plant varieties can now be bred and control strategies optimised.

All plants play host to diverse communities of microorganisms, which greatly influence plant health and development (Knief *et al.* 2011). Microbes form a variety of symbiotic relationships with their host. These can range from mutualistic, where both parties benefit, to parasitic where the microorganism benefits at the expense of the host. When a microorganism causes a detrimental affect on plant fitness it is referred to as a pathogen (Van Baarlen *et al.* 2007). Pathogenic strains may vary in virulence, which is a quantitative measure of ability to cause disease. What drives the evolution of pathogenicity in microbial communities within plants? Pathogenicity is thought to be an unavoidable consequence of parasite reproduction within its host (Sacristán & GarcÍa-Arenal 2008). As microbial populations expand they must spread and acquire new sources of nutrients to survive. They eventually cause damage to their host either directly using toxins and other virulence factors or indirectly due to the withdrawal of plant nutrients . Damage may also result from the activation of plant immune responses.

Pathogens can be broadly classified as biotrophic, whereby they keep host cells alive to obtain nutrition, or necrotrophic where they actively kill plant cells. Hemi-biotrophs are thought to be biotrophic during the critical early stages of colonisation, until a change in environmental conditions or lifecycle requirements, results in a switch towards necrotrophy (Glazebrook 2005). Pathogens can also be defined by their specificity for particular plant hosts. A pathogen's host range encompasses the plants that it can successfully infect and colonise. Pathogens can be broadly defined as generalists if they are capable of infecting a large number of hosts or as specialists, which form a more intimate relationship with one or a few host species (Pan *et al.* 2014). The study of host specificity in plant-pathogen interactions and how pathogens expand their host range is an increasingly important research area as many emerging diseases result from pathogens shifting onto new hosts.

Plants are thought to be resistant to the majority of plant pathogens (Wood 1972). Plant-pathogen interactions are described as compatible, where pathogens proliferate and may cause disease, or incompatible, where the pathogen fails to infect or proliferate (Tao *et al.* 2003). However, many interactions could be characterised on a scale somewhere between these two states. Recent studies indicate that although these definitions provide a basis for study, the biotic interactions that occur in the natural environment are highly complex. Some diseases, such as acute oak decline, involve a complement of several interacting microorganisms (Denman *et al.* 2012) and microbial communities already residing within plants may strongly influence the onset of disease (Bulgarelli *et al.* 2013). Pathogenic populations within plants may also form distinct genetic lineages which vary in virulence levels, with low and high virulence lineages fluctating within populations (Rufián *et al.* 2016).

Plant hosts counter pathogens with both preformed and induced defences to prevent disease. The plant immune system lacks the specialised circulating cells found in the mammals. Specific induced reactions occur within individual plant cells following the detection of non-self or altered self molecules. Immune signals can also spread to non-infected cells and induce systemic immunity (Spoel & Dong 2012). The reactions that occur in each cell are believed to consist of two overlapping layers of defence,

which are detailed in Figure 1.1. The first layer occurs when microbial products come in contact with plant cells. Plasma membrane-localised pattern recognition receptors (PRRs) detect the presence of conserved and indispensable pathogen molecules (pathogen -associated molecular patterns, PAMPs) or damage-associated host molecules (DAMPs). PRR activation induces an immune response referred to as PAMP-triggered immunity (PTI). PTI includes an array of defensive responses such as reactive oxygen species (ROS) production, activation of mitogen-activated protein kinase (MAPK) signalling, callose deposition and cell-wall restructuring and the activation of defence-associated gene expression. Successful pathogens are thought to overcome PTI by using virulence effector proteins, which act within plant cells to suppress PTI.

Knowledge of the other branch of the plant immune system arose from early work looking at gene-for-gene interactions between plants and their pathogens (Flor 1941; Flor 1971). Harold Flor discovered that when pairs of pathogen and host factors interact they confer strong resistance to the pathogen. These pathogen factors are now known to be effector proteins. These effectors are detected by plant resistance (R) proteins, leading to effector-triggered immunity (ETI). When an effector is recognised and induces immunity it is called an avirulence effector as it activates resistance thereby rendering the pathogen avirulent. R proteins are usually nucleotide-binding leucine-rich repeat (NB-LRR) proteins which either directly interact with effectors, or more often guard the effector's immune target and detect effector-triggered modification of the target protein. NB-LRR-associated immunity has been widely studied. Proteins that mimic the true targets of effectors have been identified and some R proteins even have integrated effector target protein domains (decoy domains) to capture effector proteins (van der Hoorn & Kamoun 2008). ETI is usually associated with a programmed cell death mechanism called the hypersensitive response (HR). The HR prevents the spread of the pathogen to further tissues and therefore limits disease. Pathogen effectors that suppress ETI have been discovered and therefore this model of plant immunity has been termed the zig-zag model. This model predicts that a co-evolutionary arms race between plants and their pathogens has led to the widespread diversification of plant immune receptors and pathogen effectors (Jones & Dangl 2006).

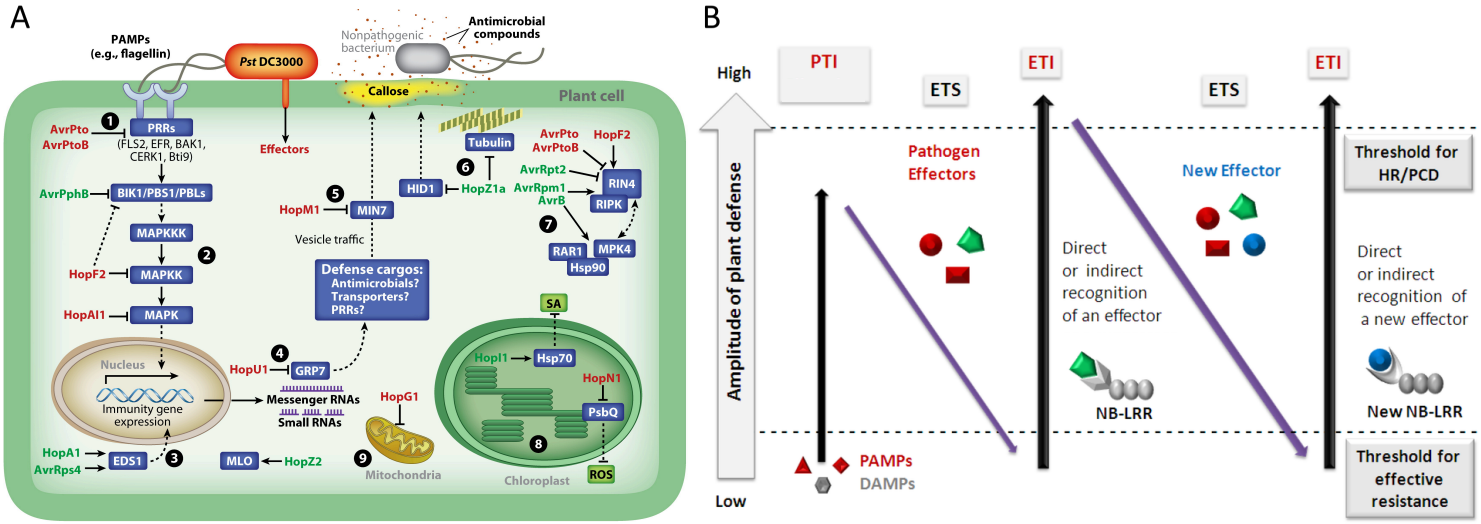


Figure 1.1: Overview of the plant immune system. ***A: Diagram of*** **P. syringae** ***interacting with a plant cell.*** *The effectors deployed by the pathogenic strain and their targets within host cells are shown. Effectors identified in* P.s *pv.* tomato *DC3000 are red, whilst those from other pathovars are green. Numbers indicate processes targeted. 1: MAMP perception, 2: MAPK cascade, 3: EDS1 complex, 4: RNA metabolism, 5: vesicle traffic, 6: microtubule function, 7: RIN4 complex, 8: chloroplast function, and 9: mitochondria function. This Figure was taken from (Xin & He 2013). Further explanation of this diagram can be found in the original paper.* ***B: The zig-zag model of plant pathogen evolution.*** *The amplitude of disease resistance varies between PTI and ETI responses. First PAMPs or DAMPs trigger the basal immune response PTI. Pathogens have evolved to overcome this response and trigger ETS (effector-triggered susceptibility). When effectors are detected by Resistance proteins an ETI response occurs which is an amplified version of PTI and can cause the HR. The zig-zag then continues with pathogen and plant coevolving by the pathogen losing detected effectors or gaining new ones or the plant acquiring new mechanisms of resistance. This Figure was taken from (Zvereva & Pooggin 2012).*

The two-layer model of plant immunity has provided an excellent hypothesis for the study of plant-pathogen interactions, however there are exceptions to its rules. There is significant overlap and cross-talk between PTI and ETI. Also, not all pathogen molecules that trigger immunity can be strictly classed as PAMPs or effectors. Some PAMPs appear to be fast evolving, such as the flg22 epitope of flagellin, which shows substantial sequence variation between *Pseudomonas syringae* strains on different plant hosts. Whilst some effectors are highly conserved, such as AvrE which has homologues in *Pseudomonas*, *Ralstonia*, *Pantoea* and various other plant pathogenic bacteria species (Cook *et al.* 2014; Ham *et al.* 2010). This general introduction will introduce the concepts explored in this thesis. Specific research of bacterial pathogenicity, genomics and the cloning of effectors is explored in more detail in the introductions to each results chapter.

1.2 Evolution of pathogenicity and host specificity

The molecular mechanisms underlying plant pathogen interactions have been well characterised. Genetic variation in both plant immune receptors and pathogen virulence factors can dictate the success or failure of a microbe to cause disease. Specialised plant pathogens are thought to have adapted to maximise fitness on particular hosts and therefore the complement of virulence factors may reflect host range. A major determinant of bacterial pathogenicity is the type III secretion system (T3SS). The role of the T3SS is to inject effector proteins into eukaryotic cells. This macro-molecular complex includes a basal body (a multi-ring complex spanning both bacterial membranes), a pilus for transporting effectors and a translocon, which creates a pore within the host cell membrane. In *P. syringae*, the T3SS is encoded by the *hrp* pathogenicity island in the majority of strains, although some lineages have been shown to contain atypical systems (Baltrus *et al.* 2017). Similar protein delivery systems are found in other plant pathogens such as *Xanthomonas*, *Ralstonia* and *Pantoea*. The acquisition of the T3SS has therefore been crucial during the evolution of plant pathogenicity of various bacterial species. It is believed either to have evolved from the flagellum machinery or that both the T3SS and flagellum constructs diverged from a common ancestor (McCann & Guttman 2008). Each pathogenic strain

possesses a repertoire of effector proteins that act within plant cells to suppress the host's immune system. As a pathogen specialises to cause disease on a particular host species, it will fine-tune its effector repertoire to maximise disease and lose any effectors that are detected by the host immune system. Other non-effector bacterial virulence factors that contribute to pathogenicity have also been identified. These include phytotoxins, ice nucleating proteins and auxins that directly damage the host or manipulate immune responses.

What genetic mechanisms underlie the ability of bacterial species to evolve pathogenicity towards plants? The plant immune system represents a considerable challenge to bacterial proliferation and survival. Bacteria populations must therefore have the capacity to continuously adapt to be able to survive on new hosts or improve virulence on their current host. The bacterial genome can be divided into core and accessory regions, which together make up the whole 'pan' genome. The core genome is shared between all strains within a taxonomic group, whereas the accessory genome consists of genes present in only a subset of strains (Vos *et al.* 2015). Bacterial genomes, particularly their accessory regions, are highly dynamic allowing rapid gene gain, loss, rearrangement and mutation which may produce phenotypic flexibility and thus a basis for the operation of natural selection (Nowell *et al.* 2014).

A major mechanism for genomic change is lateral or horizontal gene transfer (HGT).This process involves the acquisition of foreign DNA through direct transformation of cells, during phage transfection or by conjugation (mating) with other bacteria. Regions that have undergone HGT often contain mobile elements. A mobile element is a broad term, collectively used to refer to plasmids, integrative conjugative elements (ICEs), phage sequences and transposons (Frost *et al.* 2005). Table 1.1 gives a comprehensive description of each mobile element type. When mobilised, these elements may acquire DNA from the host genome, leading to the transfer of DNA regions between bacteria. Mobile elements may form part of a genomic island, which is a region usually located near tRNA loci, which contains genes whose products are involved in adaptation to particular niches, such as pathogenicity-related or antibiotic resistance genes (Guttman 2009). If the genomic island provides a

fitness benefit, it will be retained in the bacterial population. The presence of virulence genes on genomic islands in *P. syringae* suggests that these genes are often mobilised between strains via HGT. Horizontal gene transfer of virulence factors has played a major role in the evolution of *P. syringae* and other bacterial plant pathogens (Rohmer *et al.* 2004; Ma *et al.* 2006; Jacques *et al.* 2016; Aritua *et al.* 2015). For example, the gain of a plasmid, containing a T3SS and effectors by the usually commensal bacterium *Pantoea agglomerans*, led to the evolution of host-specific pathogenicity on *Gypsophila* and beet (Barash & Manulis-Sasson 2007). Phylogenetic trees based solely on *P. syringae* effector genes often conflict with phylogenies built using house-keeping genes, indicating that the effectors have evolved through horizontal transfer (Baltrus *et al.* 2011). Lovell *et al.* (2009), experimentally demonstrated that horizontal transfer of a genomic island containing the effector *hopAR1* occurs *in planta* via transformation. The genomic island is excised from the chromosome and stored in a self-replicating DNA molecule called an episome. This is released from the donor cells, subsequently taken up by recipient cells and integrated into the chromosome by recombination.

***Table 1.1 (overleaf): Mobile elements that influence the evolution of* P. syringae.** *Types of mobile genetic element are listed with a brief description. Examples of how they are influence* P. syringae *virulence are given with references. These mobile elements may be associated with integrons which do not contain the genes required to mobilise on their own. Integrons are gene cassettes that contain an integrase, a recombination site (*attI*) and a promotor to direct gene transcription. They incorporate open reading frames by site-specific recombination and drive their expression, turning them into functional genes (Mazel 2006). Definitions in the table were adapted from reviews* (Gyles & Boerlin 2014; Frost et al. 2005; Jackson et al. 2011a).

Mobile element	Key features	*P. syringae* virulence effects
Insertion sequence	Small DNA region (<2.5kb) capable of integrating into genomic DNA, it contains a transposase enzyme and inverted repeats.	IS elements have been found to insert into effector genes e.g. in *P.s* pv. *tomato* DC3000 IS*Pssy* has inserted into *hopAG1* (Greenberg 2003).
Transposon	Mobile DNA element flanked by two identical IS elements in direct or inverted orientation. May carry genes involved in adaptation to different environmental conditions.	Effectors are often present in transposons. For example, *hopX1* in *P.s* pv. *tomato* DC3000 is part of an active transposon (Landgraf *et al.* 2006).
Prophage	DNA originating from a bacteriophage that is integrated into the genome or more rarely maintained as a circular or linear plasmid. Phages may acquire and transmit host DNA which has been accidentally packaged along with the phage DNA.	Prophages often carry type III effector genes, e.g. in both *P.s* pv. *tomato* DC3000 the effector gene *avrPto1* is adjacent to a prophage (Kim *et al.* 1998; Yan *et al.* 2008).
Integrative conjugative element (ICE)	Mobile elements that are integrated into the bacterial genome. They share features with transposons and plasmids. They possess conjugation machinery allowing transfer to other cells, however they are not self-replicating like plasmids	There are multiple examples of ICEs containing effectors in *P. syringae*. For example, the effectors *hopD1*, *hopQ1-1* and *hopR1* are found on an ICE in *P.s* pv. *tomato* DC3000 (Cunnac *et al.* 2009).
Genomic/Pathogenicity island	Regions that contain evidence of past horizontal transfer such as other mobile elements, unusual GC content, variable presence in closely related strains and nearby tRNA genes. They often contain genes whose products are involved in adaptation to environmental conditions. They are sometimes mobile depending on the complement of mobile elements they contain.	The T3SS of *P. syringae* and many other plant pathogenic bacteria is present on a genomic island, suggesting past acquisition via horizontal transfer (Alfano *et al.* 2000).
Plasmid	Extra-chromosomal element that is able to self-replicate and usually contains conjugation machinery allowing transfer to other cells via conjugation.	A 154kb plasmid containing the effector *hopAB1* is crucial for virulence of *P.s* pv. *phaseolicola* on bean (Jackson *et al.* 1999).

Recombination occuring within genomes can lead to dramatic changes in virulence genes. There are frequent examples of insertion of mobile elements into genes encoding effectors. If these insertions greatly disrupt the gene, they may become pseudogenes that are no longer expressed or function differently (Jackson et al. 2011b). For example, in *X. campestris* pv. *vesicatoria*, the insertion of IS476 inactivates the effector gene *avrBs1*, allowing the evasion of host immunity due to the Bs1 R protein (Kearney & Staskawicz 1990). A study of *Ralstonia solanacearum* found that geographically distinct bacterial populations varied in the number of insertion elements within the effector gene *avrA*, which affected their ability to induce HR on tobacco (Robertson *et al.* 2004). Effector genes may also be deleted due to recombination of mobile elements. The effector gene *hopF1* was found to be deleted in the bean pathogen *P. syringae* pv. *phaseolicola* due to the mobilisation of a chimeric transposable element. This has led to the emergence of a race that lacks HopF1. This race is able to evade ETI due to recognition of HopF1 in bean cultivars expressing the corresponding R1 R protein (Rivas *et al.* 2005). Similarily in bean, *P.s* pv. *phaseolicola* has been shown to overcome host resistance to the effector HopAR1, by excision of the genomic island containing this effector from the chromosome (Pitman *et al.* 2005; Lovell *et al.* 2011).

Recombination between effector genes has been shown to occur in a process called terminal reassortment. Effectors have been described as modular proteins as they consist of several different domains, with an N terminus containing a secretion signal and functional domains in the centre and C terminus. In terminal reassortment, the termini of effector genes are mobilised and reassort with other DNA regions to create chimeric genes, whose products may have new functions *in planta* (Stavrinides *et al.* 2006). This mechanism is believed to have made a major contribution to the expansion of effector repertoires. The expansion of the *hopF* effector family in *P. syringae* has produced effectors such as HopBB1 and AvrRpm2. These two effectors contain the myristoylation motifs of HopF that are involved in plant membrane localisation, but possess novel domains that may function differently to increase virulence (Lo *et al.* 2016).

Pathoadaptation is another mechanism of bacterial evolution, whereby small, random mutations, usually due to mistakes during replication, occur in genes whose products are involved in virulence (Bartoli *et al.* 2016). If these mutations confer fitness benefits they will be selected for and increase in frequency within bacterial populations. Many effector genes have been shown to evolve at a faster rate than genes in the core genome. Particular effector domains may be under stronger selective pressures to mutate when they are being detected *in planta* (Hogenhout *et al.* 2009). Several studies have shown the importance of pathoadaptation in plant pathogen evolution. Pathoadaptation has allowed pathogens to retain effectors that are important for virulence but trigger immune responses, due to specific mutations within the detected protein domains. A study of HopZ1 in *P. syringae* showed that this effector has diverged into different allelic classes during host specialisation due to the selective pressures exerted by different host immune systems. The ancestral effector HopZ1a induces a HR in multiple plant species. However, the HopZ1b allele produced by the soybean pathogen *P.s* pv. *glycinea* has undergone a mutation to prevent recognition by the soybean immune system, whilst still retaining its virulence functions *in planta* (Ma *et al.* 2006; Lewis *et al.* 2011). Single point mutations may also significantly affect the expression of effectors. The hazelnut pathogen, *P. s* pv. *avellanae avrE1* effector gene was found to have an atypical start codon due to a single base pair change (GTG to GTA) and therefore expression of this gene may be drastically reduced (O'Brien *et al.* 2012).

1.3 The plant pathogen *Pseudomonas syringae*

This thesis is focused on the *Pseudomonas syringae* species complex, which consists of strains associated with plants and the abiotic environment (Morris *et al.* 2013). *Pseudomonas syringae* is often referred to as a species complex due to the high level of divergence between individual clades (Berge *et al.* 2014). Currently, nine genomospecies, based on DNA-DNA hybridisation, and thirteen phylogroups, based on Multi-locus sequence typing (MLST), have been described (Gardan *et al.* 1999; Parkinson *et al.* 2011). Genomic comparisons based on pairwise average nucleotide identity (ANI) show that these groups do indeed constitute separate species, based on

an ANI boundary of 95% (Marcelletti & Scortichini 2014). The definition of a bacterial species is, however, widely debated and arguably the occurrence of recombination between phylogroups confuses the matter considerably (Baltrus 2016).

Most *P. syringae* research has focused on the role of the bacterium as a plant pathogen. The overall species complex has a broad host range, infecting over 180 different plant species, however individual strains are usually restricted to one or a few hosts. Traditionally, strains were phenotypically classified without phylogenetic context into over 60 pathovars. Pathovars (pathogen variants) consist of strains that can infect and cause disease on one or a few host species (Misaghi & Grogan 1969; Young *et al.* 1978). These pathovars can be further divided into races due to specificity for particular host cultivars (Joardar *et al.* 2005). The diseases caused by *P. syringae* are remarkably diverse. *P. syringae* is responsible for both recurring chronic diseases in perennial crops such as olive knot and cherry canker (Lamichhane *et al.* 2014) and sporadic outbreaks on annual crops such as bacterial speck of tomato (Şahin 2001). Large epidemics have been reported when pathogens invade new territories, for example the recent outbreaks of kiwi fruit canker in New Zealand (McCann *et al.* 2013) and horse-chestnut canker in Northern Europe (Green *et al.* 2010).

1.4 Virulence factors in addition to proteinaceous effectors

As mentioned above, the T3SS and its effectors are a major virulence determinant of *P. syringae*. *P. syringae* has been used as a model to study plant-pathogen interactions for decades (Mansfield *et al.* 2012). Therefore, its virulence factors are probably the most well characterised of any plant pathogen. The *in planta* functions of many effectors have been elucidated in *Arabidopsis*. Table 1.2 lists the known targets of various effectors that have been elucidated through functional molecular studies. Deletion studies have shown many effectors to be functionally redundant, allowing *P. syringae* to rapidly adapt its repertoire to match the immune response of varying hosts with little impact on virulence (Cunnac *et al.* 2011). The T3SS, effectors and various other virulence factors have all been shown to be under the control of the HrpL alternative sigma factor (Xiao *et al.* 1994). HrpL activates promoters containing a

specific motif called the *hrp* box and therefore regulates gene expression, ensuring virulence factors are produced *in planta* (Vencato *et al.* 2006).

T3E target	Example	Reference
PRRs (pattern recognition receptors)	AvrPto1, HopAB1, HopAR1, HopAO1	Xiang *et al.* 2008; Göhre *et al.* 2008; Zhang *et al.* 2010; Macho *et al.* 2014
MAPK cascade	HopF2, HopAI1	Wang *et al.* 2010; Zhang *et al.* 2012
Chloroplast	HopI1	Jelenska *et al.* 2010
Mitochondrion	HopG1	Block *et al.* 2010
Vesicular trafficking of antimicrobials	AvrE1, HopM1, HopR1	Ham *et al.* 2009; Nomura *et al.* 2006; Kvitko *et al.* 2009
Gene expression	HopD1	Block *et al.* 2014
Protein translation	HopU1	Nicaise *et al.* 2013
Programmed cell death	HopAB1, HopN1	Abramovitch *et al.* 2003; López-Solanilla *et al.* 2004
Host cytoskeleton	HopZ1a, HopW1, HopG1	Lee *et al.* 2012; Kang *et al.* 2014; Shimono *et al.* 2016
Cell wall mediated defences	HopZ1a, HopE1	Lee *et al.* 2012; Guo *et al.* 2016
Proteosome	HopZ4, HopM1, HopAO1, HopA1, HopG1	Üstün *et al.* 2014; Üstün *et al.* 2016
Hormone signalling and stomatal immunity	HopX1, AvrB, HopAF1, HopBB1, HopF2	Gimenez-Ibanez *et al.* 2014; Cui *et al.* 2010; Washington *et al.* 2016; Hurley *et al.* 2014; Yang *et al.* 2017

Table 1.2: Known plant cell targets of* P. syringae *type III effectors. *The plant cell targets of each effector that have been identified through functional molecular studies are listed. References of each study are given.*

Effectors in plant pathology can be broadly defined as any pathogen protein or small molecule that alters host cell function (Alfano 2009). Although type III effectors are key to pathogenicity, other virulence effectors, some of which are part of the HrpL regulon, can greatly influence bacterial virulence (Fouts *et al.* 2002). New research based on findings from fungal plant pathogens has revealed that *P. syringae* may produce apoplastic effectors (Shindo *et al.* 2016). These apoplastic effectors are

usually protease inhibitors that inhibit plant proteases that are involved in defence. These under studied effectors may be just as important in *P. syringae* virulence as those secreted through the T3SS. Another group of molecules, the non-host specific phytotoxins have been widely studied. They are secondary metabolites produced by a range of *P. syringae* strains infecting various hosts. Unlike effectors which are translocated into plant cells, toxins are secreted into the extracellular space. They have diverse functions, including interference with plant amino acid metabolism, subversion of hormone signalling and inhibition of the proteasome, therefore indirectly suppressing defence responses reliant on these pathways (Geng *et al.* 2012; Schellenberg *et al.* 2010) . Other toxins may directly kill plant cells by the formation of pores in plant cell membranes (Bender *et al.* 1999). Some strains of *P. syringae* also produce auxins, which manipulate plant growth regulation and can interfere with hormone signalling (Fu & Wang 2011). Auxin production has been shown to be crucial for virulence of the olive knot pathogen *P.s* pv. *savastanoi* (Cerboneschi *et al.* 2016). Another common virulence factor, particularly on woody hosts, is the production of ice nucleating proteins. These proteins are found on the outer membrane of bacterial cells and act as a nucleation sites, allowing the assembly of ice crystals at temperatures as high as -2°C. This ability has been shown to increase bacterial infectivity, as frost damage of the plant can provide entry points for epiphytic bacteria (Lindow *et al.* 1989; Lamichhane *et al.* 2014).

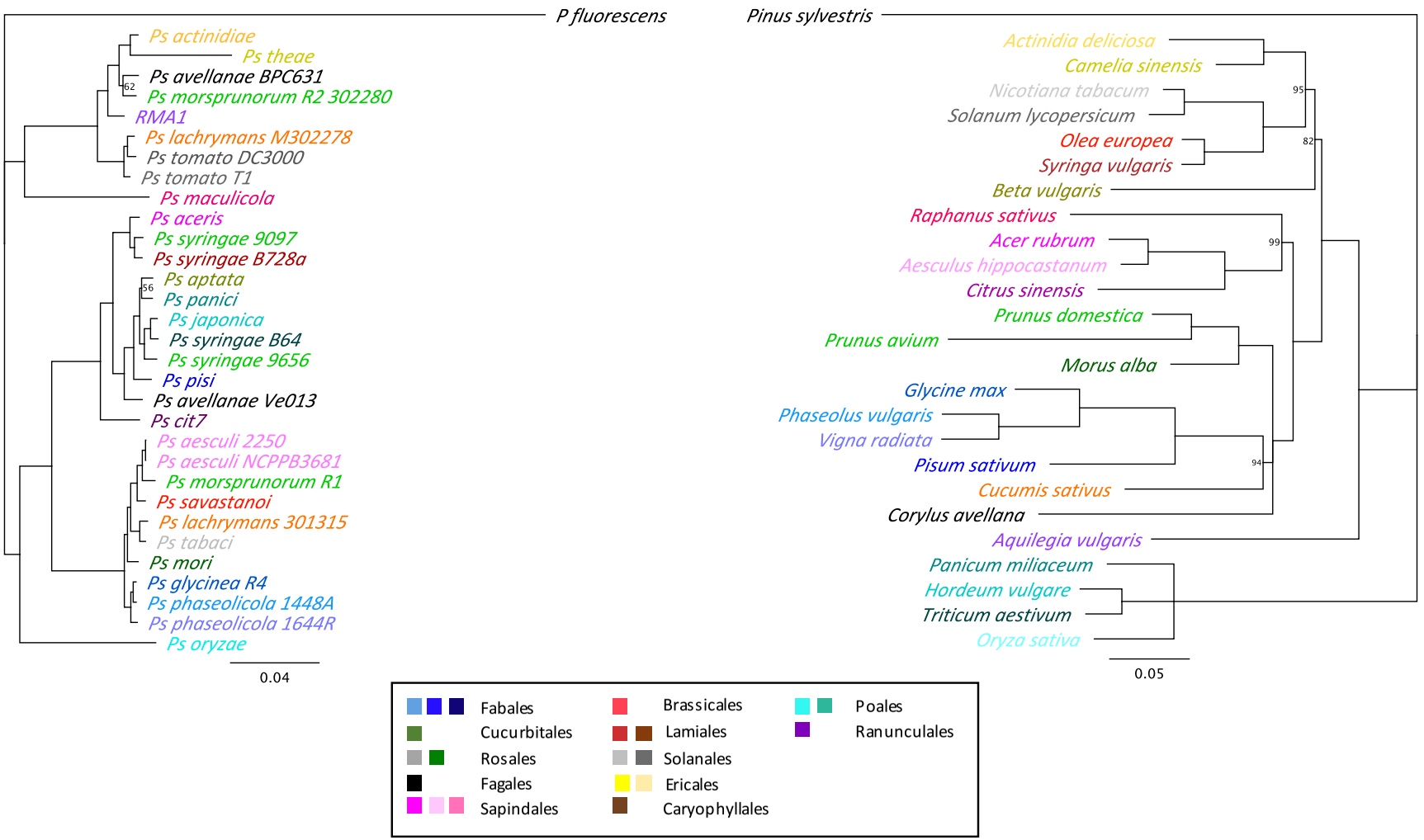
It may appear that the wealth of study on *P. syringae* has identified the majority of virulence mechanisms deployed by this pathogen. However, the field of plant pathology is a fast-evolving, with innovations in sequencing making it easier to analyse population-size datasets. With new project focusing on non-model strains, studies are identifying novel mechanisms of virulence (Matas *et al.* 2014; Shindo *et al.* 2016)

1.5 *P. syringae* as a specialist?

As most *P. syringae* strains are only capable of infecting one or a few hosts, *P. syringae* is usually referred to as a specialist plant pathogen. Figure 1.2 presents two phylogenetic trees, one of *P. syringae* and the other of its plant hosts. Incongruence between the two trees indicates that host jumps, whereby *P. syringae* moves onto a

distantly related host, have occurred frequently. There are clear examples of the molecular mechanisms that restrict host range (Lin & Martin 2007; Baltrus *et al.* 2012). Specialisation occurs on various scales, between host plant species, specific host cultivars and between plant tissue types. Several observations in recent years have contradicted the traditional host-specific, pathovar classification of *P. syringae*. The finding that *P. syringae* is frequently found in aquatic environments led to the realisation that the lifecycle of this pathogen is closely linked to the water cycle. Sequencing of environmental strains revealed that they are phylogenetically interspersed with pathogenic strains, strongly suggesting that recombination has occurred frequently between plant infecting and aquatic strains. Sampling of populations from non-agricultural environments revealed that these strains possess the T3SS and effectors, and can successfully infect tomato and several other plant species (Morris *et al.* 2007; Monteil *et al.* 2013). Therefore, although crop-associated strains may be highly clonal and host-specific, a reservoir of diverse, recombining populations also exists in the abiotic environment (Monteil *et al.* 2016). The finding that *P. syringae* and other members of the *Pseudomonas* genus exist in the wider environment is likely the reason for their large genome sizes (~6Mb) compared to other genera, as they must retain gene sets required for multiple ecological niches (Spiers *et al.* 2000).

Figure 1.2 (Overleaf): Bayesian phylogenetic trees of* P. syringae *pathovars and host plant species*. The*** P. syringae *tree is based on concatenated data from seven genes (*acnB, fruK, gapA, gltA, gyrB, pgi *and* rpoD*). The host tree is based on concatenated data from three genes (*matK, rbcL *and* trnl*). Both were made using MrBayes (Huelsenbeck & Ronquist 2001). The method of tree construction can be found in Chapter 4: Section 4.4.2. Different colours are used for each host based on plant order.* Pseudomonas fluorescens *is used as an out-group in the pathogen tree, whilst* Pinus sylvestris *for the host tree*


P fluorescens
Ps actinidiae
Ps theae
Ps avellanae BPC631
Ps morsprunorum R2 302280
RMA1
Ps lachrymans M302278
Ps tomato DC3000
Ps tomato T1
Ps maculicola
Ps aceris
Ps syringae 9097
Ps syringae B728a
Ps aptata
Ps panici
Ps japonica
Ps syringae B64
Ps syringae 9656
Ps pisi
Ps avellanae Ve013
Ps cit7
Ps aesculi 2250
Ps aesculi NCPPB3681
Ps morsprunorum R1
Ps savastanoi
Ps lachrymans 301315
Ps tabaci
Ps mori
Ps glycinea R4
Ps phaseolicola 1448A
Ps phaseolicola 1644R
Ps oryzae
62
56
0.04
Pinus sylvestris
Actinidia deliciosa
Camelia sinensis
Nicotiana tabacum
Solanum lycopersicum
Olea europea
Syringa vulgaris
Beta vulgaris
Raphanus sativus
Acer rubrum
Aesculus hippocastanum
Citrus sinensis
Prunus domestica
Prunus avium
Morus alba
Glycine max
Phaseolus vulgaris
Vigna radiata
Pisum sativum
Cucumis sativus
Corylus avellana
Aquilegia vulgaris
Panicum miliaceum
Hordeum vulgare
Triticum aestivum
Oryza sativa
95
82
99
94
0.05
Fabales
Cucurbitales
Rosales
Fagales
Sapindales
Brassicales
Lamiales
Solanales
Ericales
Caryophyllales
Poales
Ranunculales

Detailed characterisation of host specificity is often overlooked when new diseases emerge and pathovars are defined (Lamichhane *et al.* 2014). Some pathovars in particular *P.s* pv. *syringae*, which was named due to pathogenicity for lilac, include pathogens of multiple plant species. Studies suggest that *P.s* pv. *syringae* strains have undergone some specialisation, however, when tested in the laboratory, they are often able to infect many other supposed non-host plant species (Young 1991; Pour & Taghavi 2011). The model strain *P.s* pv. *tomato* DC3000 can infect both tomato and *Arabidopsis*, and when a particular avirulence effector (HopQ1-1) is deleted it becomes pathogenic to *N. benthamiana* (Preston 2000; Wei *et al.* 2007). The degree of host specialisation within *P. syringae* may therefore vary between phylogenetic clades. Exploring this diversity of host specialisation within *P. syringae* has important applications. Particular clades may possess a greater capacity to shift onto new plant hosts and cause disease epidemics if their new host lacks sources of resistance. With genomics becoming both easier and cheaper to perform, future technology could allow the prediction of a particular strain's adaptive potential based on its complement of virulence factors.

1.6 The *Pseudomonas syringae* – cherry pathosystem

This thesis will explore the molecular mechanisms involved in virulence and host specificity of *P. syringae* strains that cause bacterial canker on cherry. The term cherry encompasses several species within the genus *Prunus* and includes sweet, wild, sour and ornamental species (Hummer & Janick 2009). The stone-fruit genus *Prunus* contains several other economically important species such as plum, peach, almond and apricot. Cherry is grown commercially around the world for both fruit and timber production. The production of sweet cherry has grown drastically in the last few decades, aided by early and late season varieties, innovations in rootstocks and higher-density, covered growing systems (Lang 2000; Børve *et al.* 2003). The largest producers of cherry are Turkey, the United States and Iran (FAO 2014). Cherry is a high-value crop, but there are major challenges faced by growers - frost, disease and pests may lower fruit yields, whilst fruit cracking can make produce unsellable (Jackson & Looney 2011). Bacterial canker is a major disease of cherry and other stone-fruits, that reoccurs annually in most orchards. The disease is particularly

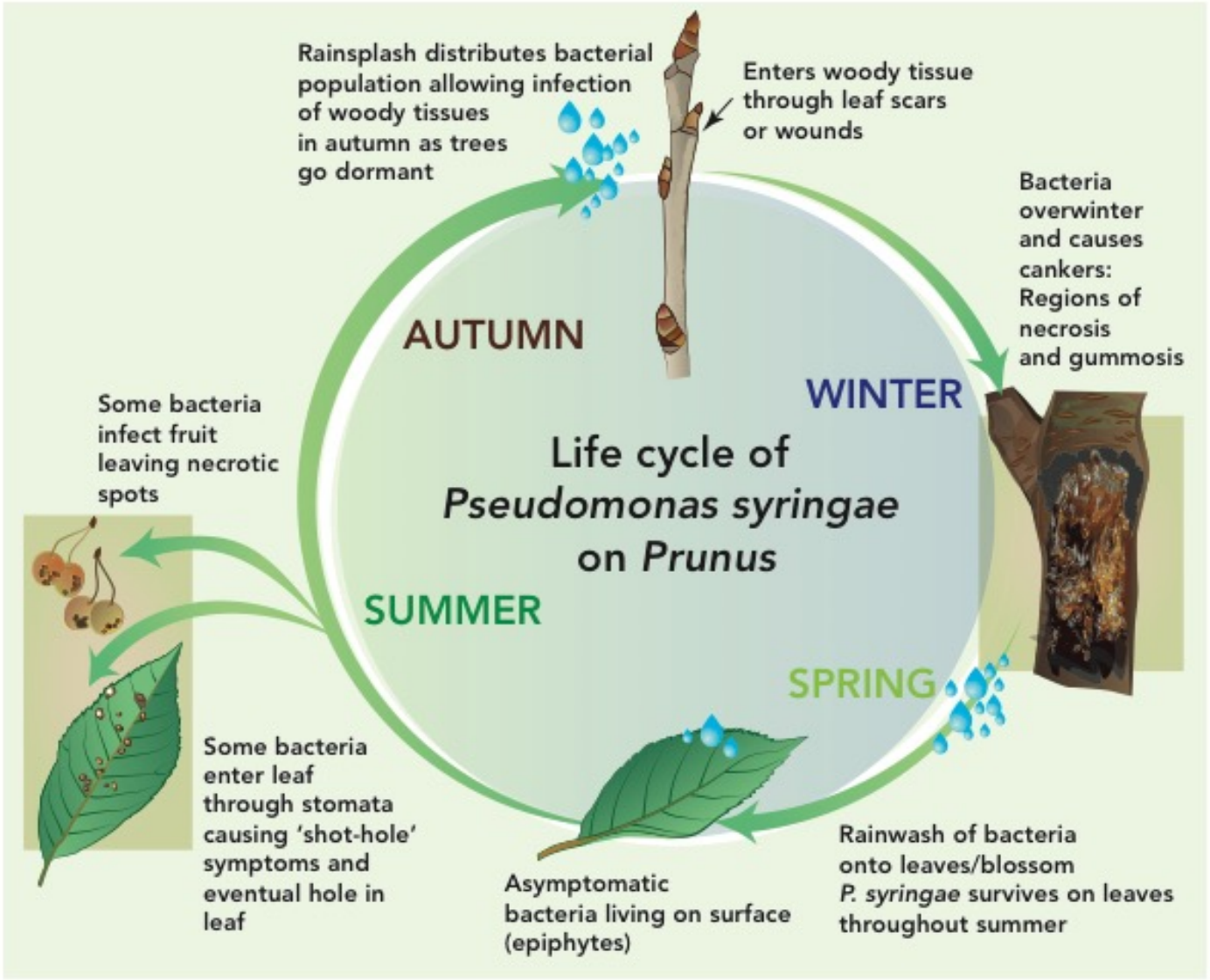
devastating in young orchards or nurseries, where tree losses of up to 70% have been reported in severe cases (Spotts *et al.* 2010).

Research of this disease began in Europe in the early 1900s (Bultreys & Kaluzna 2010). The causal agent of bacterial canker in the UK, *P. syringae* pv. *morsprunorum* (*Psm*), was originally described by Wormald (Wormald 1932). Another bacterium, *P. syringae* pv. *syringae* was later discovered to cause the same disease in the United States (Wilson 1933). After years of studying bacterial canker, phage-typing identified a third organism responsible for the disease, that showed specificity for the cherry cultivar Roundel, so was named *Psm* race 2 (Garrett 1978). *Psm* race 1 and race 2 are now known to belong to distinct clades of *P. syringae*, which have convergently adapted to cause disease on cherry (Marcelletti & Scortichini 2014). These three *P. syringae* clades, although phylogenetically distinct, frequently co-occur in orchards and are thought to be produce similar disease symptoms (Bultreys & Kaluzna 2010). Recent studies have revealed other *P. syringae* clades that can cause disease on cherry, these include *P.s* pv. *avii* which causes bacterial canker on wild cherry and *P.s* pv. *cerasicola* which causes cherry gall disease (Ménard *et al.* 2003; Kamiunten *et al.* 2000).

Thirty years of research on *Psm* Race 1, which was led by Jesse Crosse and Connie Garrett at East Malling Research in the UK, contributed greatly to our understanding of the epidemiology of this and other bacterial diseases. Figure 1.3 presents the lifecycle of this disease. *P. syringae* is spread via rain water and able to infect all aerial parts of the plant, including leaves, blossom, fruit and woody tissues (Crosse 1966). In autumn, when the tree is dormant, bacteria enter the woody tissues of the tree through wounds and leaf scars. They occupy the cambial tissue and produce black necrotic cankers in late winter and spring. During the growing period, there is a large diverse population of epiphytic bacteria that grow on the surface of the leaves. When the bacteria enter the leaf, they may become virulent and induce necrotic lesions that eventually drop out of the leaf, causing shot-hole symptoms. The asymptomatic leaf population, rather than those bacteria causing necrosis, are thought to provide the inoculum for woody tissue infections (Crosse 1959).

Effective control of bacterial canker is a serious issue for cherry growers. Chemical control is limited to copper-based sprays which limit the epiphytic population growth of bacteria on leaves (Wimalajeewa *et al.* 1991). However, these are currently being restricted in European countries and bacterial resistance to copper has been widely reported (Ghorbani &Wilcockson 2007; Sundin & Bender 1993). Measures such as improved sanitation whilst pruning and avoiding pruning during wet conditions can help prevent infection (Roberts 2012). Despite widespread breeding efforts, no fully resistant varieties have been reported although some do show tolerance to the disease (Farhadfar *et al.* 2016; Spotts *et al.* 2010). The different cherry-infecting clades may use subtly different virulence mechanisms and exhibit lifestyle differences. Therefore, to better control and breed for resistance to this disease, an understanding of genetic diversity of *P. syringae* cherry pathogens is required.

***Figure 1.3 (overleaf): The perennial lifecycle of bacterial canker disease on* Prunus.** *Diagram shows the different plant tissues that are inhabited by the bacteria throughout the year and mode of transmission.*


Rainsplash distributes bacterial population allowing infection of woody tissues in autumn as trees go dormant
Enters woody tissue through leaf scars or wounds
Bacteria overwinter and causes cankers: Regions of necrosis and gummosis
AUTUMN
WINTER
Life cycle of Pseudomonas syringae on Prunus
Some bacteria infect fruit leaving necrotic spots
SUMMER
SPRING
Some bacteria enter leaf through stomata causing 'shot-hole' symptoms and eventual hole in leaf
Asymptomatic bacteria living on surface (epiphytes)
Rainwash of bacteria onto leaves/blossom
P. syringae survives on leaves throughout summer

1.7 Thesis Aims

This project explored the genetic basis of host specificity of *P. syringae* on cherry. This thesis is divided into three main chapters that together address this topic. Each results chapter was written in the format of a research article. First the pathogenicity of a range of strains on cherry and plum was assessed using a variety of field and laboratory-based assays. Differences in host cultivar susceptibility to bacterial canker were also determined. A range of *P. syringae* strains were then genome sequenced. Comparative genomics was utilised to determine key similarities and differences in virulence factor repertoires that may explain host specificity. Candidate effector genes involved in the virulence of cherry pathogens and the avirulence of non-pathogens on cherry were then cloned and their role studied *in planta*.

Chapter 2: General materials and methods

As each results chapter was written as a complete research paper, this chapter contains only basic methods used throughout the study. Each results chapter contains a more in-depth methods section.

2.1 Maintenance of bacterial cultures

All *Pseudomonas syringae* strains used in this study were routinely plated onto Kings B agar (KBA) and kept at room temperature. KBA consists of 20 g/L peptone, 1.5 g/L K_2HPO_4, 1.5 g/L $MgSO_4$, 15 g/L agar and 1% v/v glycerol. For all experiments, colonies less than four days old were used. Liquid cultures were generated by transferring single colonies into high salt Luria Broth (LB). LB consists of 10 g/L tryptone, 10 g/L NaCl and 5 g/L yeast extract. Cultures were growth at 25 °C overnight with shaking at 180 rpm. A spectrophotometer (Jenway) was used to measure the concentration of liquid cultures. An OD_{600} of 0.2 corresponded to approximately $2x10^8$ CFU/ml, which was confirmed by dilution plating.

Escherichia coli was plated onto high salt LBA (12 g/L agar added to LB) and grown overnight at 37 °C. Bacteria were grown overnight in LB at 37 °C, 180 rpm for liquid cultures.

For long-term storage, bacteria were kept at -80 °C in 40% glycerol (v/v).

2.2 Maintenance of plant material

Cherry and plum trees were propagated by grafting. Scion wood from multiple varieties was grafted onto standard rootstocks (Gisela 5 for cherry, St Julien A for plum). Potted trees were kept on the sand bed and provided plant tissue for pathogenicity tests. To obtain fresh glasshouse grown leaves, trees that were kept in the cold store to maintain dormancy were moved into a glasshouse to promote new leaf growth. This provided 1-2 week old leaf material not contaminated with fungi or bacteria.

Nicotiana tabacum cv. Samsung was used for hypersensitive response assays. Seeds were germinated in moist tissue paper and after 10 days transferred to compost. They were grown in the glasshouse at 21 °C during the day, 18 °C at night, with 16:8 hour light:dark cycles.

2.3 DNA manipulations

DNA extraction

For most routine PCRs, a colony-based method was used, where a small amount of bacteria was picked from the plate and mixed directly into the PCR reaction. When pure DNA was required, genomic DNA was extracted using a CTAB (cetyl trimethylammonium bromide) method (Feil *et al.* 2012).

Plasmid extractions

To routinely extract plasmids, the Macherey-Nagel NucleoSpin Plasmid Mini-Prep kit was used (Macherey-Nagel). To produce high concentration DNA stocks of frequently used plasmids, the Macherey-Nagel NucleoSpin Plasmid Midi-Prep was used, which allows extraction from a much larger volume of bacterial culture.

General PCR

Primers were designed using the Primer3 Geneious plug-in (Untergasser *et al.* 2012) and purchased from Integrated DNA Technologies (IDT). For all routine PCRs, Taq 5X Master Mix (New England Biolabs) was used. The PCR program was ran using manufacturers instruction, although annealing temperature and extension times were optimised for particular reactions.

Gel electrophoresis

DNA products were run on 0.8x TAE agarose gels which had been pre-stained with 1X GelRed (Biotium). Gels were typically ran at 10 V/cm. Images of gels were captured using a Gel Doc (BioRad).

Chapter 3: Characterisation of the *P. syringae* – *Prunus* interaction with a range of field and laboratory-based assays

3.1 Abstract

Bacterial canker is a major disease of cherry and other stone fruits. It is caused by at least three pathovars of the bacterial phytopathogen *Pseudomonas syringae*. These are *P.s* pv. *morsprunorum* race 1 (*Psm* R1), *P.s* pv. *morsprunorum* race 2 (*Psm* R2) and *P.s* pv. *syringae* (*Pss*). *Psm* R1 and R2 were originally designated as races of the same pathovar, however phylogenetic analysis has revealed them to be distantly related. This study utilised various techniques to characterise the pathogenicity of *P. syringae* strains isolated from *Prunus* and other plant species, on cherry and plum, in the field and the laboratory. A field experiment determined that there were differences in host cultivar susceptibility to the different pathogen clades. The cherry cultivar Merton Glory exhibited a broad resistance to all clades, whilst the cultivars Napoleon and Van showed race-specific resistance towards *Psm* R2. This experiment also revealed that *Psm* R1 may be dividied into a race structure of strains pathogenic to both cherry and plum and other strains that are non-pathogenic on cherry. Next, the results of a range of laboratory-based pathogenicity tests (detached leaves, cherry fruits and cut shoots) were compared to the results obtained on whole-trees. The use of cut shoots was found to best replicate field results. Measuring population growth of bacteria in detached leaves reliably discriminated pathogens and non-pathogens, however was not sensitive enough to detect subtle differences in pathogen virulence between cherry cultivars. When inoculated at higher concentrations, based on the timing of symptom appearance, the leaf assay was able to discriminate between disease lesions caused by *Psm* races and a putative hypersensitive response caused by non-pathogens. *Pss* was found to rapidly induce disease lesions and could not be differentiated from non-pathogens by symptom development. This indicated that *Pss* may exhibit a more necrotrophic pathogenic lifestyle than the hemi-biotrophic *Psm* races. The in-depth study of pathogenic interactions, identification of resistance in

cherry genotypes and optimisation of laboratory-based assays achieved in this study, will provide a reliable framework for future genetic dissection of virulence and host resistance mechanisms.

3.2 Introduction

Bacterial canker caused by *Pseudomonas syringae* is a major disease of cherry, for which there is no effective control (Spotts *et al.* 2010). Several distantly related clades within the three major *Pseudomonas syringae* phylogroups are pathogenic to cherry. Three distinct clades are reported as the main causal agents of this disease. These are *P.s* pv. *syringae* (*Pss*), *P.s* pv. *morsprunorum* (*Psm*) race 1 and *P.s* pv. *morsprunorum* race 2 (Bultreys & Kaluzna 2010). The two *morsprunorum* races have been reported to specifically occur on *Prunus* species, whilst *Pss* strains are more variable and can infect various plant species (Bultreys & Kaluzna 2010). Although distantly related, *Psm* R1 and R2 were originally described as races of *P.s* pv. *morsprunorum*, due to their specific interactions with particular cherry cultivars (Garrett 1978).

An understanding of how the different clades of *P. syringae* cause bacterial canker and how host genotypes vary in resistance is crucial for breeding efforts. Various studies have been conducted in the last century exploring the biology and ecology of the canker pathogens. The perennial lifecycle and epidemiology of this disease was determined through many years of field inoculation studies at East Malling in the UK (Crosse 1966; Crosse & Garrett 1966; Freigoun & Crosse 1975; Garrett 1978). Other studies have investigated pathogen diversity. Molecular techniques such as Repetitive Element Sequence-Based (REP) PCR and Multi-Locus Sequence Typing (MLST), and phenotyping methods, such as the LOPAT test that is used to identify *P. syringae*, have been used to survey the bacterial populations in orchards (Gilbert *et al.* 2008; Vicente & Roberts 2007; Kaluzna *et al.* 2010; Renick *et al.* 2008). These studies revealed the three pathovars often co-exist within orchards with each other and non-pathogenic *Pseudomonas*. To characterise the pathogenicity of isolates, several laboratory- and field-based assays have been developed, including the use of tissue culture plants, dormant shoots and immature fruits (Vicente & Roberts 2003; Gilbert *et al.* 2009; Crosse & Garrett 1966; Farhadfar *et al.* 2016). Improved assays are required, not only

to determine pathogen virulence levels, but also for screening for host resistance. Gilbert *et al.* (2009) used various lab-based assays to determine the pathogenicity of a range of strains isolated from stone-fruits in Belgium. They found that no individual laboratory assay could reliability predict pathogenicity under field conditions. They concluded that field inoculations are required to fully ascertain pathogenicity and differences in host response.

The breeding of resistant cherry cultivars has been hindered due to the complex nature of this disease (Garrett 1979). Early work reported variation in cultivar susceptibility towards the different clades of pathogenic *P. syringae*, with two cultivars Napoleon and Roundel exhibiting differential susceptibility towards the two races of *Psm*. Napoleon was found to be resistant to R2 but susceptible to R1, and vice versa for Roundel (Garrett 1978). It may be challenging to breed resistance to all three of the genetically distinct *P. syringae* clades. Differing environmental factors such as frost exposure and host factors such as tree age, also impact susceptibility (Lamichhane *et al.* 2014; Crosse 1966; Kus *et al.* 2002). Recent studies have identified rootstock selections and cherry varieties exhibiting a degree of resistance (Santi *et al.* 2004; Farhadfar *et al.* 2016; Spotts *et al.* 2010; Li *et al.* 2015). Despite this progress, there is still a lack of totally resistant varieties available and the genetic factors underlying canker resistance remain unknown.

The molecular mechanisms that underlie plant resistance to bacterial pathogens, including PTI and ETI, were discussed in detail in Chapter 1. Very little is known about the resistance mechanisms that are triggered by canker-causing bacteria and other wood-infecting pathogens (Lamichhane *et al.* 2014). In general, resistance can be divided into non-host and host, depending on the plant-pathogen interaction. Non-host resistance occurs when all host genotypes are resistant to all strains of a non-adapted pathogen. This is a broad and durable response, which involves a combination of both PTI and ETI (Lee *et al.* 2016). In comparison, research has mainly focused on identifying sources of host resistance, which is when different host genotypes exhibit variation in resistance towards the pathogen. This can involve specific R protein – effector interactions between particular host genotypes and

pathogen races. Resistance responses based on ETI are sometimes qualitative, whereby the possession of a single R gene provides complete resistance against a particular pathogen effector. Breeding efforts have utilised this knowledge and introduced R genes into susceptible cultivars of various crop species over the last few decades (Pink 2002). However, single R gene-mediated resistance has been found to not be very durable; it can quickly break down in the field as strong selection pressures drive the pathogen to adapt and prevent effector detection. The pyramiding of multiple R genes, particularly those whose products detect core pathogen effectors, could provide a more durable response (Dangl *et al.* 2013).

Single R proteins may therefore provide qualitative and complete resistance through ETI. However, the host resistance of many plant species is quantitative, whereby resistance leads to reduction but not absence of the disease (French *et al.* 2016). This type of resistance is often controlled by more than one gene, with different host genotypes exhibiting a range of susceptibility levels to the pathogen (Poland *et al.* 2009), and even towards single effectors (Iakovidis *et al.* 2016). Quantitative host resistance may encompass a combination of PTI and ETI, and involve allelic variation in a whole range of genes controlling immunity and host physiological factors that dictate susceptibility (Corwin *et al.* 2016). As well as R genes, variation in susceptibility (S) genes may influence host resistance. S gene products are important for compatible interactions with pathogens, allowing pathogens to infect and survive *in planta*. Therefore, mutations or losses of these genes may reduce disease susceptibility (van Schie & Takken 2014). As no fully canker-resistant host genotypes have been identified, the genetic basis of host resistance to bacterial canker is likely to be quantitative. It may involve many different genes that additively contribute to more resistant phenotypes.

To analyse the genetics of pathogenicity and disease resistance to canker, the pathogenicity of a diverse range of *P. syringae* strains, including pathogens of cherry and plum. as well as strains isolated from non-host species, was examined. Using both lab-based and field inoculations, the interaction of these strains with both cherry and

plum was characterised. This will provide a reliable pathogenicity framework for the future genetic dissection of virulence and host resistance mechanisms.

3.3 Methods

3.3.1

Bacterial strains used in this study are presented in Table 3.1.

Table 3.1 (overleaf): Bacterial strains used in this study with host of isolation and reference/source. *Strains sequenced in this study are listed first, followed by the out-group strains* Pph *and* Psav *included in pathogenicity tests and then the rest of the strains used only for phylogenetic analysis. The Genbank accessions of genomes used for the phylogenetic analysis are included.*

Strain	**Pathovar**	**Race**	**Host/Isolation source**	***Prunus* host cv.**	**Host tissue**	**Reference**	**BioProject/accession**
R1-5244	*morsprunorum*	1	*Prunus avium*	unknown	Cankerous wood	This study	PRJNA345357
R1-5300	*morsprunorum*	1	*Prunus domestica*	Victoria	Unknown	This study	PRJNA345357
R1-9326	*morsprunorum*	1	*Prunus domestica*	Victoria	Leaf wash	This study	PRJNA345357
R1-9629	*morsprunorum*	1	*Prunus domestica*	Victoria	Leaf wash	This study	PRJNA345357
R1-9646	*morsprunorum*	1	*Prunus avium*	Stella	Leaf wash	This study	PRJNA345357
R1-9657	*morsprunorum*	1	*Prunus domestica*	Victoria	Leaf wash	This study	PRJNA345357
R2-5255	*morsprunorum*	2	*Prunus avium*	Napoleon	Unknown	This study	PRJNA345357
R2-5260	*morsprunorum*	2	*Prunus avium*	Roundel	Unknown	This study	PRJNA345357
R2-leaf	*morsprunorum*	2	*Prunus avium*	Napoleon	Leaf lesion	This study	PRJNA345357
R2-sc214	*morsprunorum*	2	*Prunus avium*	Wild cherry	Leaf lesion	This study	PRJNA345357
syr9097	*syringae*		*Prunus avium*	unknown	Cankerous wood	This study	PRJNA345357
syr9293	*syringae*		*Prunus domestica*	Victoria	Leaf wash	This study	PRJNA345357
syr9630	*syringae*		*Prunus domestica*	Victoria	Leaf wash	This study	PRJNA345357
syr9644	*syringae*		*Prunus avium*	Stella	Leaf wash	This study	PRJNA345357
syr9654	*syringae*		*Prunus domestica*	Victoria	Leaf wash	This study	PRJNA345357
syr9656	*syringae*		*Prunus avium*	Kiku-Shidare	Leaf wash	This study	PRJNA345357
syr9659	*syringae*		*Prunus avium*	Kiku-Shidare	Leaf wash	This study	PRJNA345357
Ps-9643	-		*Prunus domestica*	Victoria	Leaf wash	This study	PRJNA345357
RMA1	-		*Aquilegia vulgaris*	Winky	Leaf lesion	This study	PRJNA345357
PsavBP631	*avellanae*		*Corylus avellana*			O'Brien *et al.* 2012	AKBS00000000
Pph1448a	*phaseolicola*		*Phaseolus vulgaris*			Joardar *et al.* 2005	CP000058

Strain	Pathovar	Race	Host/Isolate source	*Prunus* Host cv.	Reference	BioProject/accession
acer302273	*aceris*		*Acer* sp.		Baltrus *et al.* 2012	AEAO00000000
act302091	*actinidiae*		*Actinidia deliciosa*		Baltrus *et al.* 2012	AEAL00000000
aes2250	*aesculi*		*Aesculus hippocastanum*		Green *et al.* 2010	ACXT00000000
aes3681	*aesculi*		*Aesculus hippocastanum*		Green *et al.* 2010	ACXS00000000
amy3205	*amygdali*		*Prunus dulcis*		Bartoli *et al.* 2015	JYHB00000000
amyICMP3918	*amygdali*		*Prunus dulcis*		Thakur *et al.* 2016	LJPQ00000000
atroDSM50255	*atrofaciens*		*Triticum aestivum*		Baltrus *et al.* 2014	AWUI00000000
avelVe013	*avellanae*		*Corylus avellana*		O'Brien *et al.* 2012	AKCK00000000
avelVe037	*avellanae*		*Corylus avellana*		O'Brien *et al.* 2012	AKCJ00000000
avii3846	*avii*		*Prunus avium*		Nowell *et al.* 2016	LIIJ00000000
castCFBP4217	*castaneae*		*Castanea crenata*		Nowell *et al.* 2016	LIIH00000000
CC1557	-		Snow		Hockett *et al.* 2014	AVEH00000000
cera6109	*cerasicola*		*Prunus yedoensis*		Nowell *et al.* 2016	LIIG00000000
ceraICMP17524	*cerasicola*		*Prunus yedoensis*		Thakur *et al.* 2016	LJQA00000000
ciccICMP5710	*ciccaronei*		*Ceratonia siliqua*		Thakur *et al.* 2016	LJPY00000000
cit7	-		*Citrus sinensis*		Baltrus et al. 2012	AEAJ00000000
cunnICMP11894	*cunninghamiae*		*Cunninghamia lanceolata*		Thakur *et al.* 2016	LJQE00000000
daphICMP9757	*daphniphylli*		*Daphniphyllum teijsmannii*		Thakur *et al.* 2016	LJQF00000000
glyR4	*glycinea*		*Glycine max*		Qi *et al.* 2011	AEGH00000000
lach301315	*lachrymans*		*Cucumis sativus*		Baltrus *et al.* 2012	AEAF00000000
lach302278	*lachrymans*		*Cucumis sativus*		Baltrus *et al.* 2012	AEAM00000000
lapsaICMP3947	*lapsa*		*Zea* sp.		Thakur *et al.* 2016	LJQQ00000000
mors302280	*morsprunorum*		*Prunus domestica*		Baltrus *et al.* 2012	AEAE00000000
morsU7805	*morsprunorum*		*Prunus mume*		Mott *et al.* 2016	LGLQ00000000

Strain	*Pathovar*	*Race*	*Host/Isolate source*	Prunus Host cv.	Reference	BioProject/accession
myriICMP7118	*myricae*		*Myrica rubra*		Thakur *et al.* 2016	LJQV00000000
neriiICMP16943	*savastanoi*		*Olea europea*		Thakur *et al.* 2016	LJQW00000000
paniLMG2367	*panici*		*Panicum miliaceum*		Liu *et al.* 2012	ALAC00000000
papu1754	*papulans*		*Malus sylvestris*		Nowell *et al.* 2016	JYHI00000000
persNCPPB2254	*persicae*		*Prunus persica*		Zhao *et al.* 2015	LAZV00000000
photICMP7840	*photiniae*		*Photinia glabra*		Thakur *et al.* 2016	LJQO00000000
pisiPP1	*pisi*		*Pisum sativum*		Baltrus *et al.* 2014b	AUZR00000000
R1-2341	*morsprunorum*	1	*Prunus cerasus*	unknown	Nowell *et al.* 2016	LIIB00000000
R1-5269	*morsprunorum*	1	*Prunus cerasus*	unknown	Nowell *et al.* 2016	LIHZ00000000
R2-5261	*morsprunorum*	2	*Prunus avium*	Roundel	Nowell *et al.* 2016	LIIA00000000
sava3335	*savastanoi*		*Olea europea*		Rodriguez-Palenzuela *et al.* 2010	ADMI00000000
sava4352	*savastanoi*		*Olea europea*		Thakur *et al.* 2016	LGKR00000000
soliICMP16925	*solidagae*		*Solidago altissima*		Thakur *et al.* 2016	JYHF00000000
syr1212	*syringae*		*Pisum sativum*		Baltrus *et al.* 2014	AVCR00000000
syr41a	*syringae*		*Prunus armeniaca*		Bartoli *et al.* 2015	JYHJ00000000
syrB301D	*syringae*		*Pyrus communis*		Ravindran *et al.* 2015	CP005969
syrB64	*syringae*		*Triticum aestivum*		Dudnik and Dudler 2013	ANZF00000000
syrB728a	*syringae*		*Phaseolus vulgaris*		Feil *et al.* 2005	CP000075
syrHS191	*syringae*		*Panicum miliaceum*		Ravindran *et al.* 2015	CP006256
syrSM	*syringae*		*Triticum aestivum*		Dudnik and Dudler 2013	APWT00000000
thea3923	*theae*		*Camelia sinensis*		Mazzaglia *et al.* 2012	AGNN00000000
tomDC3000	*tomato*		*Solanum lycopersicum*		Buell *et al.* 2003	AE016853
tomT1	*tomato*		*Solanum lycopersicum*		Almeida *et al.* 2009	ABSM00000000
ulmiICMP3962	*ulmi*		*Ulmus* sp.		Thakur *et al.* 2016	LJRQ00000000

3.3.2 Genome sequencing

Nineteen *P. syringae* strains were genome sequenced. DNA was extracted using the Puregene Yeast/Bact Kit (Qiagen, UK). DNA libraries were then prepared by first fragmenting the DNA using a sonicating water bath for 30 seconds. Fragmented DNA ranging from 400-700bp was extracted from an agarose gel using the Zymogen gel extraction kit (Zymo Research). The NextFlex Rapid-DNA sequencing kit was then used to create each DNA library excluding the size selection step. Barcodes were multiplexed to allow pooling of multiple samples. Libraries were quality checked using the NGS kit on the Fragment Analyzer (Advanced Analytical, USA) and concentration measured on the Qubit (Life Technologies, UK).

Libraries were then sequenced using the Illumina Mi-Seq V3 (Illumina, USA) 300bp paired-end reads. Raw data for each genome was then quality checked and trimmed to remove adapter sequences using fastqc-mcf (https://code.google.com/p/ea-utils/wiki/FastqMcf). The reads were error corrected using Quake prior to assembly(Kelley *et al.* 2010). Each genome was then assembled using SPAdes 3.1.0 (Bankevich *et al.* 2012) and summary statistics generated using Quast(Gurevich *et al.* 2013).

Further details of the genome sequencing and assembly are presented in Chapter 4. However, the basic methods are presented here as phylogenetic analysis was required for the selection of strains for detailed pathology testing.

3.3.3 Phylogenetic analysis

A phylogenetic tree was created for all 19 sequenced strains and 47 other strains which had genome sequences available on NCBI. These other strains included isolates from *Prunus* and out-groups isolated from a range of different plant species. The nucleotide sequences of seven house-keeping genes (*acnB*, *fruK*, *gapA*, *gltA*, *gyrB*, *pgi* and *rpoD*) (Sarkar & Guttman 2004) were extracted from all genomes and individually aligned using Geneious 7.1.9. The alignments were then concatenated and trimmed to produce an overall alignment of 9393bp. This alignment was used to create a Bayesian phylogeny using the Geneious plug-in of MrBayes (Huelsenbeck & Ronquist 2001). The

GTR gamma model of evolution was used with a burn-in length of 100,000 and sub-sampling frequency of 200. The divergent strain CC1557 was used as an out-group to root the tree.

3.3.4 Plant material

All *Prunus* plant material was propagated at NIAB-EMR. For whole-tree inoculations, one-year old grafted trees were used. Cherry cultivars were grafted on the rootstock Gisela 5, whilst plum was grafted on St Julian A. For detached leaf assays, 1-2 week old, fully-expanded leaves was obtained from glasshouse grown trees. Immature cherry fruits were obtained from mature trees.

3.3.5 Pathogenicity assays

Bacterial inoculum was prepared from overnight LB cultures. The bacteria were spun down (3500g, 10 minutes). They were then re-suspended in sterile 10mM $MgCl_2$. A spectrophotometer was used to measure optical density, with an optical density of 0.2 (OD_{600}) being approximately $2x10^8$ CFU/ml (Debener *et al.* 1991).

3.3.5.1 Whole-tree inoculations

Whole-trees were inoculated though either wounds or leaf scars (Crosse & Garrett 1966). Field inoculations were performed during leaf-fall in October 2015, whilst glasshouse wound inoculations were done on dormant trees in February 2015. Bacterial suspensions of approximately 10^7 CFU/ml were used for inoculation or 10mM $MgCl_2$ for the control.

To wound inoculate, a sterile scalpel was used to cut a shallow wound into the trunk of the tree and 200µl of bacterial suspension was pipetted into the wound. To inoculate leaf scars, the leaf was removed and 10µl of bacterial suspension was pipetted on the exposed scar. The infection sites were then covered with parafilm and duct tape to maintain a high humidity. Multiple inoculations were performed on the same tree, with at least four buds between each inoculation. For the field experiment the trees were left for six months before assessment in May 2016, whereas the glasshouse assessment was performed after two months. To assess symptoms, the

tape was removed and bark stripped back using a scalpel. Length of necrosis was measured in mm using a caliper and the disease score was determined as 0: no symptoms, 1: limited browning, 2: black necrosis, 3: necrosis and gumming and 4: necrosis, gumming and spreading from site of infection.

3.3.5.2 Experimental design of whole-tree inoculations

To randomise the glasshouse whole-tree canker assay an incomplete block-design was used. This allowed assessment of 22 different strains across 22 independent trees, with five replicates of each strain, each at different positions on the tree. In the field experiment the virulence of eight different strains was assessed on four cherry cultivars and two plum cultivars, using two different inoculation methods. To limit the number of trees this would require, the eight different strains were divided across two trees, with inoculation sites at least four buds apart. Each tree also had one negative control. This meant that two adjacent trees comprised one experimental unit of all strains and controls inoculated on the same cultivar using one inoculation method. A balanced incomplete design was used to randomise position of each strain and both controls onto the two trees. A balanced complete design was then used to randomise the different cultivars and inoculation methods within ten blocks in the field. Each block contained 24 trees (16 cherry and 8 plum), and the total field contained 240 trees. Further explanation of the experimental design of the field experiment is presented in the Chapter 3 supplementary. The field experiment and glasshouse wound inoculation experiments were only performed once.

3.3.5.3 Cut shoot inoculations

The cut shoot assay was performed as in previous studies (Krzesinska *et al.* 1992; Santi *et al.* 2004). Eight strains were inoculated onto four cherry and two plum cultivars relicated in ten blocks. Bacterial inoculum at a concentration of 10^7 CFU/ml was prepared as above. Dormant one-year old shoots with approximately the same diameter (5 mm) were collected from mature trees in December and cut into 10 cm sections. They were surface sterilised with 0.5% hypochlorite for five minutes and then rinsed with tap water and left overnight to air-dry. Next, 5 mm from the shoot tip was cut with sterile secateurs and the shoot was dipped in bacterial suspension for

five minutes. The wound was covered with parafilm (Fisher Scientific, UK) and the shoot bases were freshly cut (5 mm) and placed in transparent-boxes immersed in water to a depth of 20 mm. The shoots were incubated in the closed boxes at 15 °C with 16-hour light, 8-hour dark cycle for one week. Separate boxes were used for each bacterial isolate to prevent cross-contamination.

Next, the shoots were transferred to −2 °C for one week to simulate frost damage. Finally, the basal 10 mm of each shoot was removed and they were placed in a completely randomised design in water-soaked Oasis Foam (Oasis Floral, UK) in trays containing 30 mm of water. These were incubated for a further 4 weeks at 15 °C. The trays were covered with cling-film to maintain a high humidity.

The shoots were assessed for severity of necrosis by peeling back the uppermost layer of bark from the top 30 mm of the shoot. Images were taken on a digital camera which were then analysed automatically using the machine-learning program developed by Dr Bo Li (Li *et al.* 2015 - paper presented in the Thesis Appendix). This experiment was repeated twice (in 2014 and 2015).

3.3.5.4 Cherry fruit inoculations

To inoculate immature cherry fruits a stab-inoculation method was used (Moragrega & Llorente 2003). First, the fruits were surface sterilised in 0.5% hypochlorite for five minutes before rinsing several times in sterile distilled water. Bacteria were then scraped from 5-day old plates using a 24g needle and stabbed into the plant material. After inoculation, the fruits were placed in transparent boxes lined with moist tissue paper to maintain a high humidity. The fruits were kept at 22 °C and visually assessed over time. Pictures were taken after 10 days. Due to limited material availability these assays were only performed once.

3.3.5.5 Leaf inoculations

Inoculum concentration varied from $2x10^6$ CFU/ml for population count studies to $2x10^8$ CFU/ml for symptom scoring. Freshly picked, 1-2 week old leaves were used for leaf inoculations. The leaves were infiltrated with bacterial suspension from the

abaxial surface using a blunt-ended syringe. The leaves were then placed in plastic trays, which contained a 1 cm layer of water agar (10g/L) covered in damp paper towel to maintain a high humidity. The tray was then sealed inside a transparent bag and incubated at 22 °C (16hr light, 8hr dark). The leaves were left for a maximum of 10 days before assessment. At least three leaves were inoculated for each isolate, with the three replicate leaves coming from different plants.

Bacterial population growth within the leaves was measured over time. Leaf discs were excised using a sterile cork borer (0.5 cm). These were then homogenised in 10mM $MgCl_2$. A dilution series was then plated out to determine bacterial concentration (CFU/ml). Three technical replicates were performed. Leaf assay experiments were repeated at twice, apart from the assay involving different cherry cultivars due to limitations in clean leaf material. A diagram of the experimental design of leaf population assays is presented in supplementary Figure S3.2.

3.3.6 ROS and callose assays

To visualise the host response to infection, several assays that detect molecules involved in immune responses to plant pathogens were trialled. An assay to detect the production of reactive oxygen species (ROS) 6 hours after pathogen infiltration was derived from previous studies of *P. syringae* on *Arabidopsis thaliana* (Torres *et al.* 2005; Daudi & O'Brien 2012). Briefly, freshly detached cherry leaves were inoculated with *P. syringae* (at a concentration of $2x10^7$ CFU/ml) and left for one hour at 22 °C. The leaves were then infiltrated with a 10 mM Na_2HPO_4 3,3'-diaminobenzidine (DAB) staining solution, which is oxidised by hydrogen peroxide to form a brown precipitate. The stained leaves were kept in the dark for 5 hours rotating slowly at room temperature to allow staining to occur. A sterile 10mM $MgCl_2$ control where no bacteria were infiltrated was included for comparison. Several different infiltration methods were trialled, including immersion in the DAB solution and vacuum infiltration either in a vacuum chamber or by hand using a syringe-based method (Matas *et al.* 2014). After five hours of incubation the leaves were cleared using methanol heated to 65 °C and assessed for brown pigmentation. To check the

methodology was appropriate, ROS was also detected in the model plant *Nicotiana tabacum*.

To detect callose deposition within cherry leaves, similar protocols developed in *A. thaliana* were used (Schenk & Schikora 2015; De Torres *et al.* 2006). Detached cherry leaves were infiltrated with *P. syringae* at a concentration of 10^7 CFU/ml and left for 48 hours at 22 °C. The infiltrated section of the leaves were then cleared using hot 65 °C methanol until they became colourless and translucent. After washing in 150mM K_2HPO_4, the sections were incubated in a 150mM K_2HPO_4 0.01% aniline blue staining solution for at least six hours in the dark. The sections were embedded in a 50% glycerol solution and viewed using a Leica AF6000 fluorescence microscope. To visualise callose, a DAPI filter with an excitation wave-length of 370nm was used. To check that the methodology was appropriate, callose in response to *Psm* R2 in the model plant *N. tabacum* was also detected.

3.3.7 Electron microscopy

Electron microscopy was performed by Dr Ian Brown (University of Kent) on infected cherry leaves. Detached leaves were infiltrated as above with bacteria at $2x10^6$ CFU/ml and left for one week at 22 °C.

Inoculated leaves were cut into 2 mm squares using a razor blade in a drop of cold fixative (2.5% glutaraldehyde in 100mM sodium cacodylate buffer pH 7.2 [CAB]) on dental wax. Pieces of tissue were transferred into a glass vial containing 2 ml of cold fixative and stored at 4 ˚C overnight. Multiple washing steps were then performed. Tissue was washed in CAB for 2x10 minutes and then post fixed in 1% Osmium tetroxide in CAB for 1h at room temp (20 ˚C). The tissue was then washed again for 2x10mins in milliQ water. Tissue was then dehydrated in increasing concentrations of ethanol 50, 70, 90% (10mins each) and 3x100% (15mins). Tissue was then washed (2x10mins) in the transition solvent propylene oxide and transferred into 50% propylene oxide 50% Agar low viscosity resin (ALVR) (Agar Scientific) for 30 mins. Tissue was then placed in 100% ALVR 2x1.5h. Tissue pieces were then placed in silicone moulds and polymerised in the oven (60˚C) for 24h. Polymerised blocks were

sectioned (70nm) on a RMC MTXL ultramicrotome and sections were collected on 400 mesh copper grids. Sections were counterstained with 4.5% uranyl acetate in 1% acetic acid for 45mins and Reynolds lead citrate for 7mins. Sections were viewed in a Jeol 1230 TEM with an accelerating voltage of 80kV and images recorded with a Gatan Multiscan 791 digital camera.

3.3.8 Statistical analysis

R software (R Core Team 2012) was used for all statistical analysis. The glasshouse whole-tree experiment was analysed using a REML (Restricted Maximum Likelihood) model due to the imbalanced design, with strain as a fixed effect and position on tree as a random effect. The model: lmer(score ~ strain + (1|tree). The field experiment was complicated by missing values due to tree deaths, meaning a REML was used. Strain and cultivar were fixed effects, and block (with tree number nested within) was treated as a random effect. The model: lmer(log_2(length+0.1) ~ cultivar * strain + (1|block/no.). Wound and scar inoculations were analysed separately due to differences in variance. The cut shoot test was analysed using an ANOVA on the log_2-transformed data set. The model: aov(log_2(area+0.1) ~ strain * cultivar + block). For the cherry fruit experiment comparing all *P. syringae* strains, an ANOVA was used, the model: aov(log_2(length+0.1) ~ strain + replicate). For the cherry fruit experiment comparing cherry cultivars, a REML analysis was used due to missing values. Strain and cultivar were fixed values, whilst box (block) was random. The model: lmer(log2(length+0.1) ~ cultivar * strain + (1|box). To analyse individual leaf experiments, ANOVA was used after log_2 transformation as the data were not normally distributed. The model: aov(log_2(cfu+0.1) ~ strain + leaf/replicate). No bacteria control leaves were not included in the analysis as this would exaggerate the difference between strains in the ANOVA as controls were always 0 CFU/ml. The ANOVA looked for differences between strains and leaves (with technical replicate nested within). If different cherry cultivars were being analysed, cultivar was a factor. To analyse the symptom scores on leaves, the slope of symptom development from 0-24 hours was calculated and an ANOVA used to determine differences between slope. The model: aov(log_2(slope+0.1) ~ strain + leaf). For all analyses, a post-hoc Tukey-HSD (honest significant difference) test (p=0.05), available in the agricolae package of R (de

Mendiburu 2016), was used to group similar strains. The package ggplot2 was used to make graphs of raw data (Warnes *et al.* 2016). To compare the different pathogenicity tests, results were standardised: (data – mean(data) / standard deviation(data)), to allow the means to be plotted on the same scale. Correlation matrices were created using the means and plotted using the R package corrplot. All ANOVA tables are presented in the thesis appendix.

3.4 Results

To determine the diversity of a set of strains isolated from cherry and plum, phylogenetic analysis was peformed. Cherry, plum and non-pathogen strains were then pathogenicity tested on cherry trees in the glasshouse to determine if they were able to cause disease. The virulence of a subset of these strains was then characterised on trees in the field. This field experiment also determined if the different host cultivars varied in susceptibility towards the different pathogens. These whole-tree assays provided a basis to which several rapid laboratory-based pathogenicity assays could be compared.

3.4.1 Phylogenetics

To determine the diversity of strains isolated from cherry and plum, 18 *P. syringae* strains were genome sequenced. The strains from *Prunus* included bacteria representative of all three previously designated clades, *Psm* R1, *Psm* R2 and *Pss*. *Pss* and *Psm* R1 included strains isolated from both cherry and plum, whilst the *Psm* R2 strains all originated from cherry. A strain that did not belong to these clades (Ps 9643), which had been isolated during a plum leaf wash was also included. Finally, an additional strain isolated from the perennial species *Aquilegia vulgaris*, that preliminary analysis had shown to be closely related to *Psm* R2 was sequenced. Seven MLST genes were then extracted from the genomes. Homologous sequences from 59 genome assemblies of other strains within the *P. syringae* complex were then downloaded from NCBI. These included strains that were also isolated from *Prunus* and those isolated from other plant species for comparison. A Bayesian phylogenetic tree was then generated based on a concatenated alignment of the seven genes using MrBayes with a GTR gamma model of evolution.

Figure 3.1 shows the Bayesian phylogenetic tree of *P. syringae* with strains isolated from *Prunus* highlighted and designated as cherry (red) or plum (blue) isolates. *Psm* R1, R2 and Pss were found within phylogroups 3, 1 and 2 respectively. The two *Psm* races fell into discrete monophyletic clade, with individual strains being very closely related. In comparison, *Prunus*-infecting strains of *Pss* exhibited greater diversity and were found across the phylogroup 2. Strains isolated from cherry and plum did not form distinct clusters in any of the pathogenic clades, indicating that they are closely related and may therefore cross-infect the two *Prunus* species. Ps 9643 was found to be closely related to the *Prunus persicae* pathogen (*P.s* pv. *persicae*), whilst RMA1 was an out-group to the clade containing *Psm* R2 and the pathovars *P.s.* pv. *actinidiae* and *P.s.* pv. *theae* (which infect kiwifruit and tea respectively).

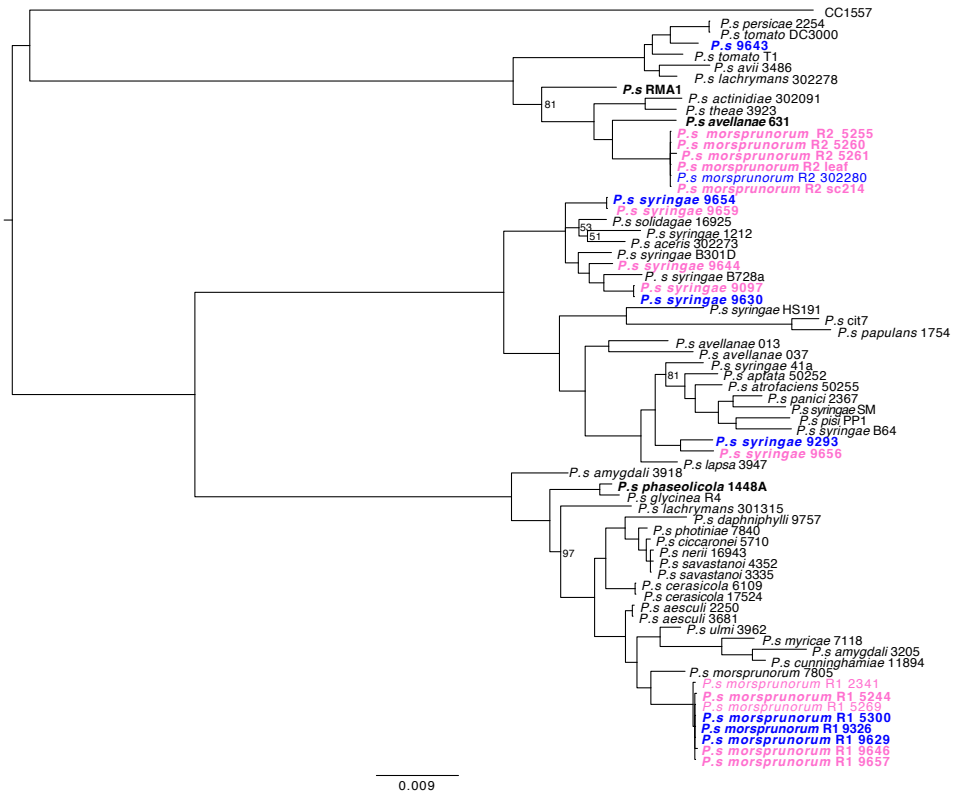
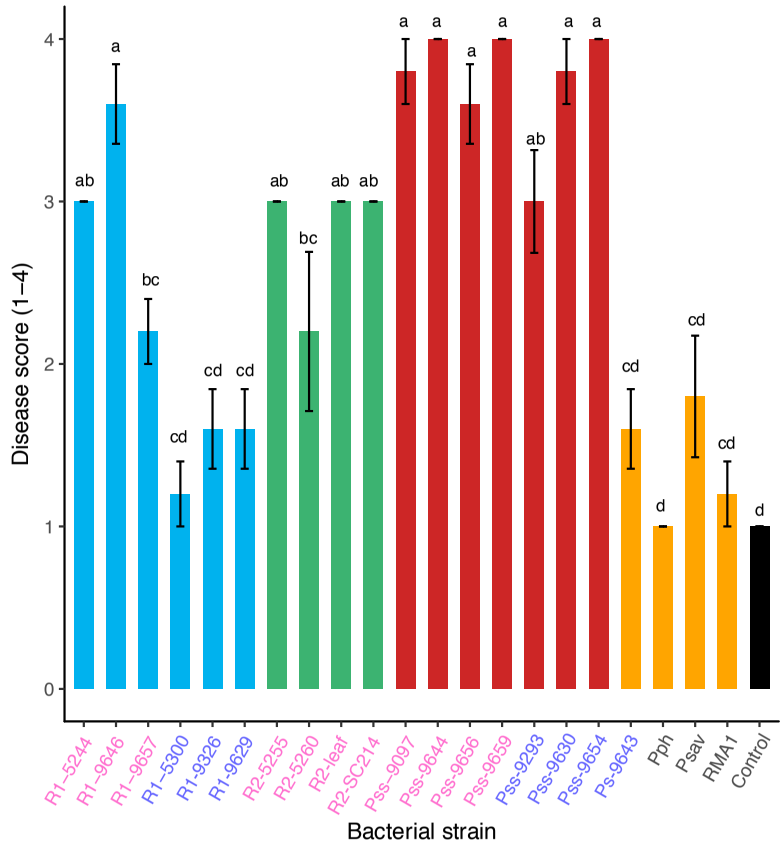


***Figure 3.1: Bayesian phylogenetic tree of* P. syringae.** *The phylogeny was constructed using a concatenated alignment of seven genes (*acnB*,* fruK*,* gapA*,* gltA*,* gyrB*,* pgi *and* rpoD*). Strains from the three major phylogroups were selected for analysis, with the canker-causing clades* Psm *R1,* Psm *R2 and* Pss *highlighted. Strains isolated from cherry are in pink, whilst those from plum are in blue. Strains in bold were tested in this study. Scale shows substitutions per site. Bootstrap support <99% are presented.*

3.4.2 Whole-tree glasshouse experiment

To determine the fundamental ability of each strain to cause bacterial canker on cherry, a whole-tree wound inoculation experiment was performed in the glasshouse during Winter 2015. No heating or lights were used so the conditions closely matched the outside environmental conditions. All strains isolated from *Prunus* were included. Several out-groups from non-host plants*, P.s* pv. *phaseolicola* 1448A (*Pph*) , *P.s* pv. *avellanae* BPIC631 (*Psav*) and RMA1 were included as potential non-pathogens. Comparisons between strains were made based on the level of necrosis produced in the cambial layer underneath the bark at the site of infection after two months of incubation. Figure 3.2 shows that the strains exhibited a wide range of virulence levels on cherry. Both the non-pathogens and negative control gave very limited browning and callusing associated with a wound response (score: 1). Pathogens were then scored for browning of the tissue (2), black necrosis and gumming (3) and black necrosis, gumming and significant spreading from the initial wound site (4). The data were analysed using REML, which confirmed that there were significant differences between strains ($p<0.01$, df=21). A Tukey-HSD post-hoc test grouped the strains by significant difference. There was variation between the different *Prunus*-infecting clades. Strains of *Psm* R1 and R2 showed variation in virulence, but most only scored a maximum of 3. Meanwhile, most strains of *Pss* were able to spread from the site of infection so produced scores of 4. Within *Psm* R1, only two of the strains isolated from cherry (R1-5244 and R1-9646) were able to cause gumming and necrosis, whilst R1-9657 showed reduced virulence. Strains of *Psm* R1 isolated from plum were not significantly different from the non-pathogens and the control. This indicated that there may be host-specificity occurring towards different *Prunus* species within the *Psm* R1 clade. Most strains of *Psm* R2 were pathogenic, however R2-5260 showed reduced virulence. Apart from one strain, *Pss* was highly pathogenic, with strains generally spreading from the site of infection. The strain Ps-9643 isolated from a plum leaf was also not significantly different from non-pathogens and the control, indicating it is unable cause bacterial canker on cherry.



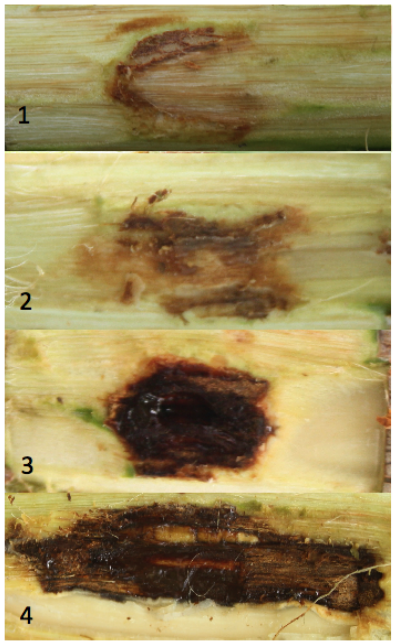


Figure 3.2: Evaluation of symptoms caused by different* P. syringae *strains after wound inoculation of cherry cv. Van trees. *Strains isolated from cherry and plum are highlighted in pink and blue. The bar chart is colour-coded,* Psm *R1: blue,* Psm *R2: green,* Pss*: red, non-pathogens: orange and control: black. The level of disease was measured using a score between 1 and 4, with representative images of each score shown on the right. 1: Limited browning, 2: browning, 3: black necrosis and gumming, 4: necrosis, gumming and spread from site of inoculation. Data presented are the mean values of five replicates, with error bars showing standard error above and below the mean. Tukey-HSD significance groups (p=0.05, confidence level: 0.95) for the different strains are presented above each bar.*

3.4.3 Whole-tree field experiment

A set of strains with contrasting pathogenicity and host of isolation was then chosen for pathogenicity screening under field conditions. The strains included cherry pathogens (R1-5244, R2-leaf, *Pss*-9097 and *Pss*-9293) and non-pathogens (R1-5300, Ps-9643, *Pph* and RMA1). The strain of *Psav* examined above could not be included in the field experiment as being from outside the UK, this strain was considered a licensed pathogen. The experiment was performed on both cherry and plum, across a variety of cultivars, to allow the comparison of host susceptibility to the different pathogens. Concerning cherry, the cultivar Merton Glory is reported to be tolerant to canker (APS 1966), Napoleon and Roundel show race-specific differences, with Napoleon being susceptible to *Psm* R1 and tolerant to R2 (and vice versa in Roundel) (Garrett, 1978). The cultivar Van is reported to be highly susceptible (Long & Olsen 2013). Whilst in plum, Victoria is reported as highly susceptible and Marjorie's Seedling is more resistant (RHS, n.d.). Trees were inoculated through both leaf scars or artificial wounds, which are both known to be natural entry routes for pathogens in the field (Crosse 1966). The trees were left over winter, representing the normal disease cycle of bacterial canker, and assessed in spring. Each disease lesion was given a disease score and the length of necrosis measured. The variance of length data from the wound inoculations was much greater than that for leaf-scar inoculations. In addition, the variance on cherry was much greater than that on plum. This meant that the wound and scar data sets were analysed separately for each *Prunus* species.

In cherry, inoculation of the previously characterised pathogenic strains led to clear disease symptoms using both leaf scars and wounds (Figure 3.3). In both inoculations the pathogens generally caused a large degree of necrosis, whilst non-pathogens were similar to the control. The resulting phenotypes reflected those recorded in the glasshouse experiment, with pathogens causing cankerous black necrosis and gumming. However, under field conditions, strains from all three pathogenic clades were able to spread from the site of infection (in the glasshouse experiment only *Pss* appeared to spread). In the field, contamination by wild *Pseudomonas* strains may have occurred, and this could explain why some control and non-pathogen

inoculations appeared to generate disease symptoms (6% of cherry controls scored 3 or higher). The REML analyses indicated there were significant differences between strains and cultivars for both the leaf scar and wound experiments ($p<0.01$, df=8 and $p<0.01$, df=3 respectively). There was no significant interaction between strain and cultivar, as pathogen and non-pathogen responses were generally consistent across cultivars. However, there did appear to be variation in *Psm* R2 virulence between the cultivars, with reduced virulence compared to *Psm* R1 on Van and Napoleon, but a similar level of virulence on Roundel. The plum strain *Psm* R1 5300 was not significantly different to the non-pathogens, indicating that it lacks pathogenicity for cherry. The two *Pss* strains varied considerably in virulence, with the cherry isolate *Pss* 9097 always causing significantly greater necrosis than the plum isolate *Pss* 9293. This is consistent with the results of the glasshouse inoculation, where *Pss* 9293 showed reduced ability to cause canker. The cultivar Merton Glory appeared to be more tolerant to canker, with reduced necrosis levels for all pathogens.

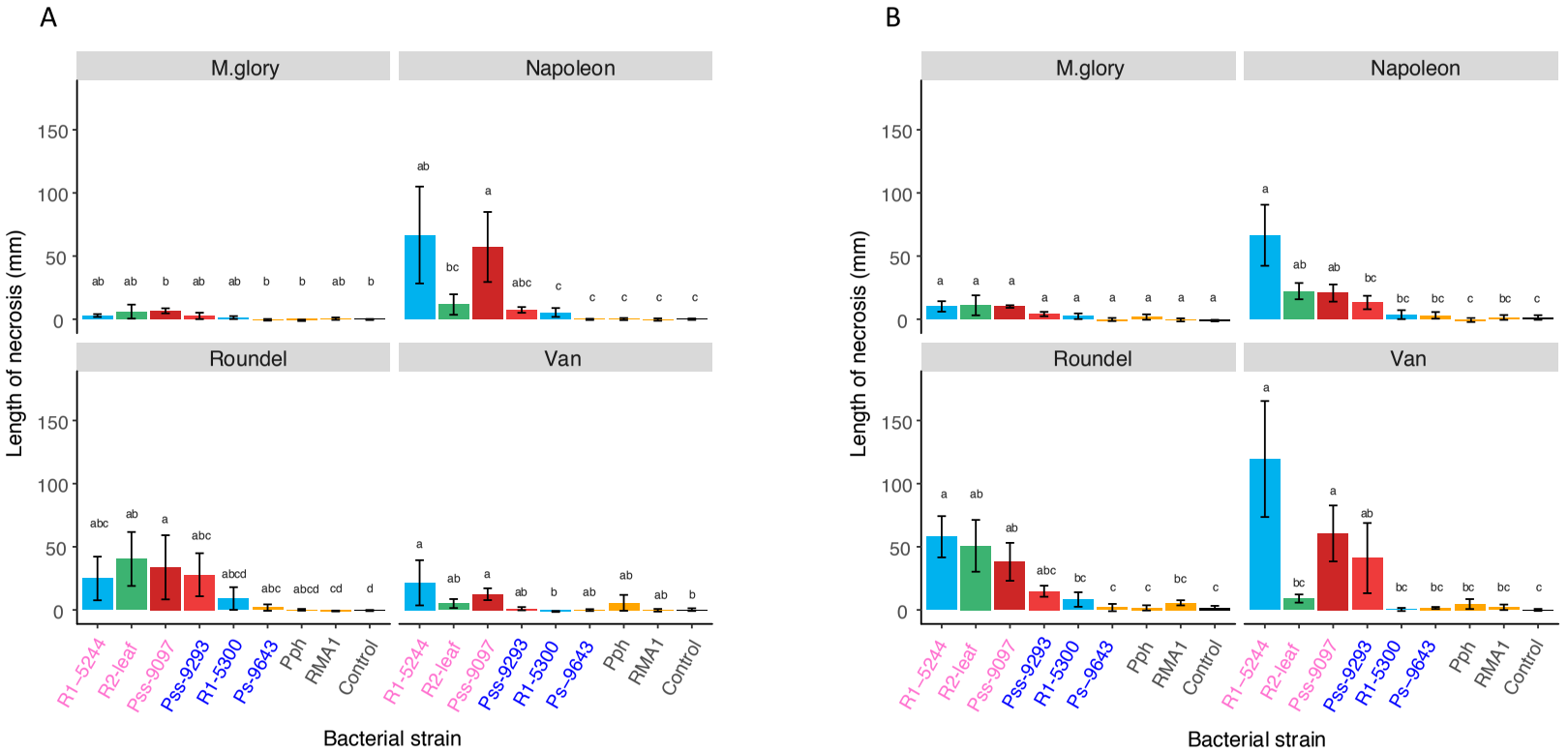


Figure 3.3: Evaluation of symptoms caused by* P. syringae after inoculation of *different cherry cultivars in the field. *A: Leaf-scar inoculations. B: Wound inoculations. The bar plot shows length of necrosis (mm) caused by each* P. syringae *strain. Strains isolated from cherry and plum are highlighted in pink and blue. The bar chart is colour-coded,* Psm *R1: blue,* Psm *R2: green,* Pss*: red, non-pathogens: orange and control: black. Data presented are the mean of ten replicates standardised to the negative control of the most susceptible cultivar, with error bars showing standard error above and below the mean. Tukey-HSD (p=0.05, confidence level: 0.95) significance groups for the different strains for each separate cultivar are presented above each bar.*

In plum (Figure 3.4), with both types of inoculation method there were significant differences between strains ($p<0.05$, df=8 for both), but no significant differences between the two cultivars ($p=0.20$, df=1 for scar and $p=0.35$, df=1 for wound). Symptoms produced were similar to those on cherry, with necrosis and gumming being indicative of disease. A post-hoc Tukey-HSD multiple comparison test ($p=0.05$) revealed that with both methods of inoculation, the only strains that were significantly different from the control were *Pss* 9097 and *Psm* R1 5300. The plum cultivar Marjorie's Seedling did not appear to be susceptible to leaf-scar infection, as no strain caused a necrosis length significantly different from the $MgCl_2$ control. The plum cultivar Victoria was more susceptible to leaf scar infections, with all strains of *Pss* and *Psm* R1 causing some necrosis. *Psm* R2 appeared unable to invade through plum leaf scars and cause disease.

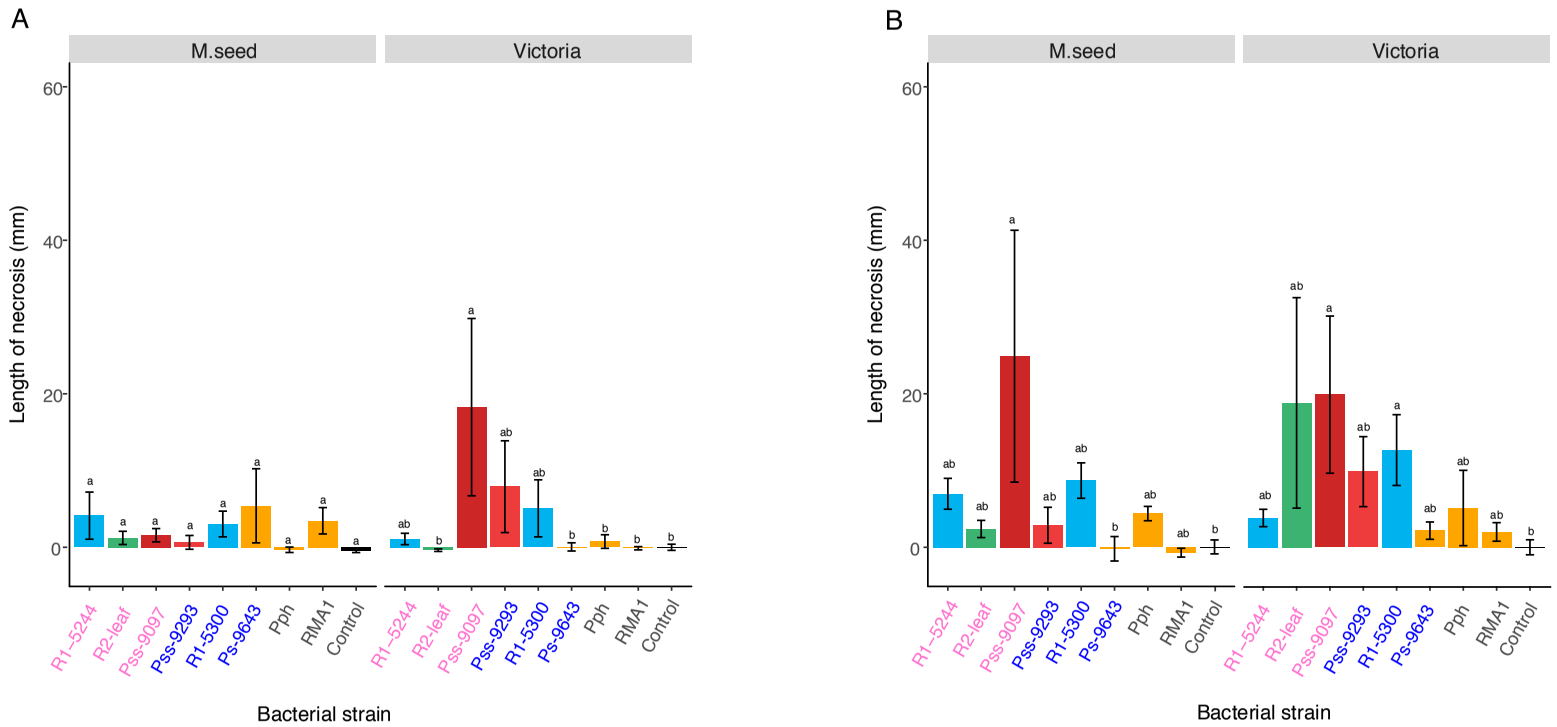


Figure 3.4: Evaluation of symptoms caused by* P. syringae *on different plum cultivars in the field. *A: Leaf-scar inoculations. B: Wound inoculations. The bar plot shows length of necrosis (mm) caused by each P. syringae strain. Strains isolated from cherry and plum are highlighted in pink and blue. The bar chart is colour-coded,* Psm *R1: blue,* Psm *R2: green,* Pss*: red, non-pathogens: orange and control: black. Data presented are the mean values of ten replicates standardised to the negative control of the most susceptible cultivar, with error bars showing standard error above and below the mean. Tukey-HSD (p=0.05, confidence level: 0.95) significance groups for the different strains for each separate cultivar are presented above each bar.*

3.4.4 Laboratory-based pathogenicity assays

The whole-tree inoculations in the glasshouse and field allowed the pathogenicity and virulence levels of different *P. syringae* strains on *Prunus* to be determined. However, to screen for differences in virulence these methods are very slow and involve the destruction of whole trees. To undertake large-scale screens of *Prunus* mapping populations or perform molecular studies of pathogenicity these methods are not very suitable. Therefore, several laboratory-based assays were trialled to both compare the pathogenicity of different *P. syringae* strains and screen different cherry cultivars for variation in susceptibility.

3.4.4.1 Cut shoot inoculations

Several studies have documented the use of detached cut shoots for rapid screening for resistance to bacterial canker (Santi *et al.* 2004; Krzesinska *et al.* 1992). Using most of the strains included in the field assay, the same cultivars of cherry and plum were screened with the cut shoot method. This involved using one-year old dormant shoots and inoculating at a cut end by dipping in a bacterial suspension. The extent to which initial necrosis spread down the shoot from this point could then be used to measure differences in bacterial virulence/host resistance. As cutting back the bark from shoots was a long and difficult process, only the first 3cm from the point of infection was assessed. Shoots were imaged and Dr Bo Li developed a rapid image analysis algorithm for determining percentage area diseased from these images (see Appendix paper documenting development of this method).

Figure 3.5 shows the results on both cherry and plum. An analysis of variance (ANOVA) was performed on the log transformed percentage data. It revealed that there were significant differences between *Pseudomonas* strains ($p<0.01$, df=8), no significant difference between the susceptibility of the two *Prunus* species ($p=0.65$, df=1) and there was a significant interaction between *Prunus* species and *P. syringae* strain ($p<0.01$, df=8), as well as interactions between strain and individual cultivars ($p<0.01$, df=36). Focusing on cherry, pathogenic strains *Psm* R1-5244, *Psm* R2-5255 and *Pss* 9097 were able to cause necrosis >5% of the shoot area. The two *Psm* races varied in virulence on the different cultivars. As in the field experiment, *Psm* R2 was more

virulent on Roundel, but less virulent on Van compared to *Psm* R1. However, in comparison with the field experiment *Psm* R2 caused similar necrosis to *Psm* R1 on Napoleon. Whereas, in the field experiment *Psm* R2 had caused minimal disease on this cultivar. The cut shoot test also confirmed that Merton Glory shows some tolerance compared to the other cultivars.

On plum, the level of necrosis on Victoria was generally higher than that on Marjorie's Seedling. As observed in the field experiment, the plum strain of *Psm* R1 (R1-5300), was able to cause necrosis where it had failed on cherry, indicating again that *Psm* R1 strains tested show host-specificity for the two *Prunus* species. The non-pathogens were able to match some of the pathogenic strains in virulence on plum. On cv. Victoria the *Aquilegia* pathogen RMA1 caused necrosis similar to *Pss* 9097, although the standard error bars were large indicating there was substantial variation in ability to cause necrosis.

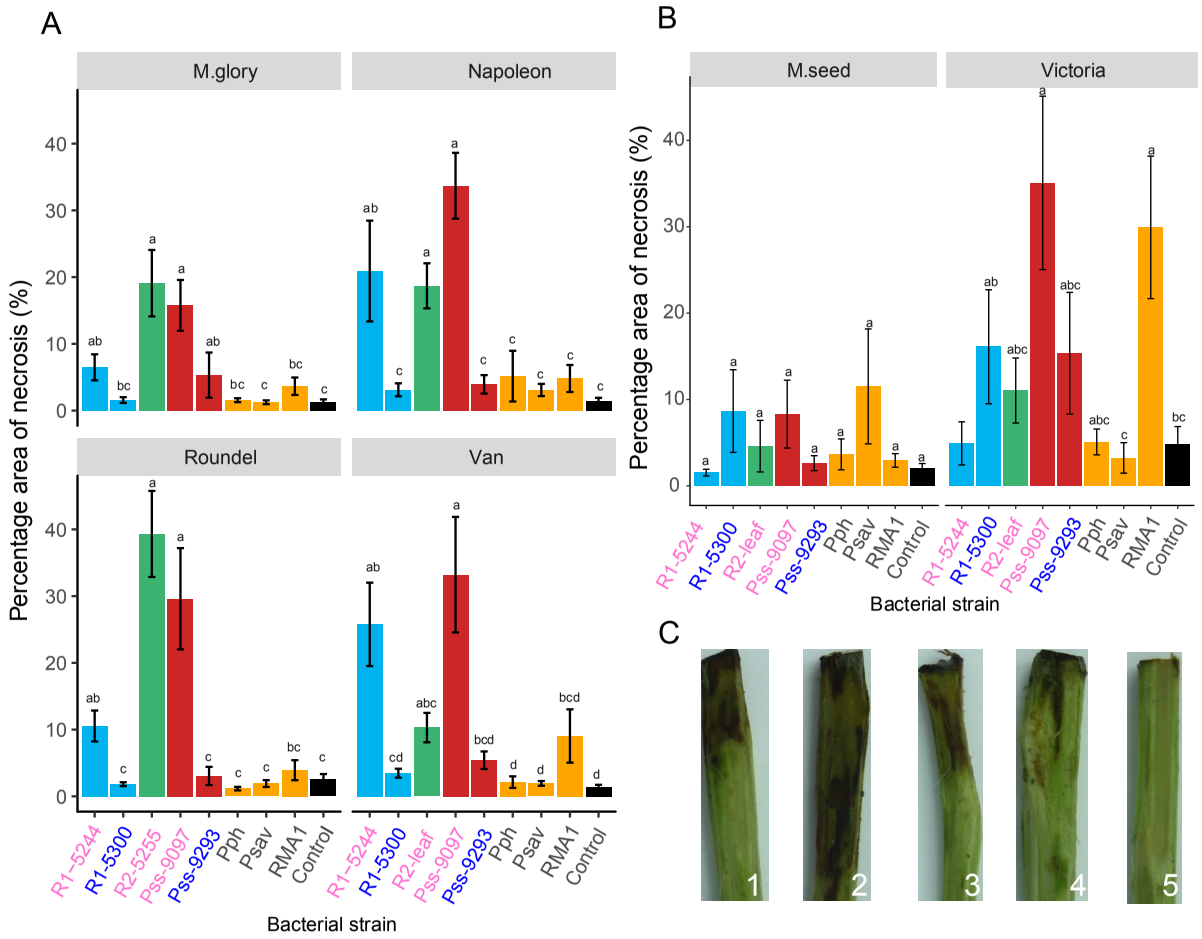


Figure 3.5: A cutshoot assay using different inoculated strains of* P. syringae *on cherry and plum cultivars revealed quantitative symptom differences. *A: Barplot of percentage area of necrosis of different* P. syringae *strains on cherry cultivars. Strains isolated from cherry and plum are highlighted in pink and blue. The bar chart is colour-coded,* Psm *R1: blue,* Psm *R2: green,* Pss*: red, non-pathogens: orange and control: black. Data presented are the mean values of ten replicates, with error bars showing standard error above and below the mean. Tukey-HSD (p=0.05, confidence level: 0.95) significance groups for the different strains for each separate cultivar are presented above each bar. B: The same parameters for two different plum cultivars. C: Images of the cutshoot infection caused by* Pss *9097 on cv. Napoleon (1-4) and the no bacteria control (5).*

3.4.4.2 Inoculation of detached immature cherry fruits

Next, the suitability of immature cherry fruits was assessed for screening for bacterial canker resistance. Green immature fruits were found to be very easy to inoculate by stabbing the fruit with bacteria using a 24g needle. The symptoms developed rapidly, often within a few days. The different clades that infect *Prunus* produce remarkably different symptoms on this tissue. Strains of *Pss* rapidly produced large necrotic lesions on cherry fruits, whereas the *Psm* races both produced grey, water-soaked lesions (Bultreys & Kaluzna 2010). An initial screen of all *P. syringae* strains used in this study inoculated onto cherry fruits (Figure 3.6), confirmed significant differences between strains ($p<0.01$, $df=21$). Strains of the *Pss* clade caused a large amount of necrosis, often spreading across the whole fruit within the first 48 hours. Strains within the *Psm* races that were confirmed to be pathogenic on whole trees in the glasshouse assay, caused small water-soaked lesions. However, the ANOVA revealed that diameters of the water-soaked lesions were mostly not significantly different from non-pathogens (see Tukey-HSD groups, $p=0.05$). Most of the non-pathogens caused very limited browning, however the non-pathogenic strain Ps-9643 caused a similar level of water-soaking to the pathogenic Psm races.

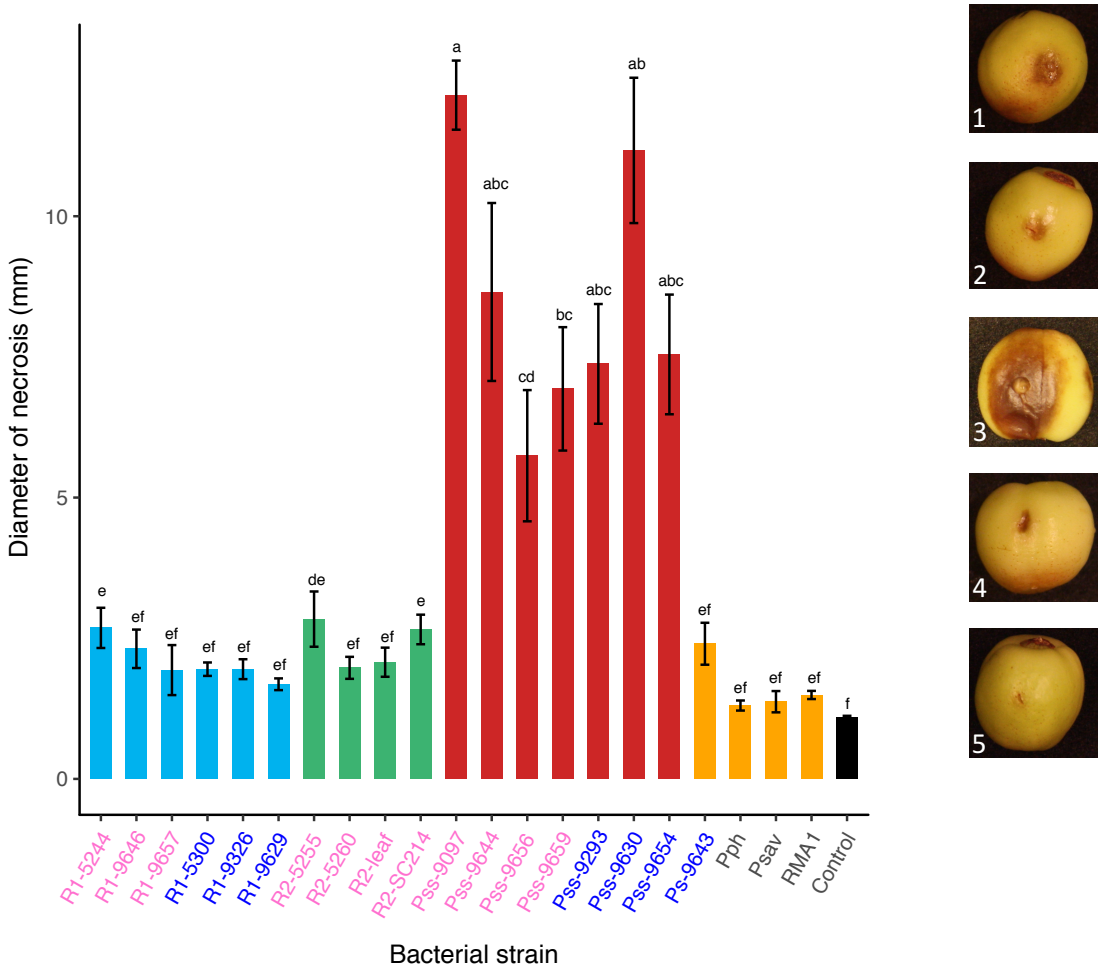


Figure 3.6: Diameter of necrosis caused by different* P. syringae *strains on immature cherry fruits. *Strains isolated from cherry and plum are highlighted in pink and blue. The bar chart is colour-coded,* Psm *R1: blue,* Psm *R2: green,* Pss*: red, non-pathogens: orange and control: black. Data presented are the mean values of five replicates. Error bars show standard error above and below the mean. Representative images are presented. 1:* Psm *R1, 2:* Psm *R2, 3:* Pss*, 4: non-pathogens, 5: control. Tukey-HSD (p=0.05, confidence level: 0.95) significance groups are presented above each bar.*

One member of each pathogenic clade was then used to screen for differences between cherry cultivars. As the fruits were sourced from trees in the field there was a substantial amount of fungal contamination, meaning that 19% of the results were missing values. Therefore, a REML analysis was performed to assess for differences between strains and cultivars. The results can be seen in Figure 3.7. There were significant differences between strains ($p<0.01$, $df=3$), cultivars ($p<0.01$, $df=3$) and an

interaction between strain and cultivar (p=0.012, df=9). However, previously reported differences in virulence between the two *Psm* races, which were observed in the field experiment, were not picked up using this assay due to the limited symptom development.

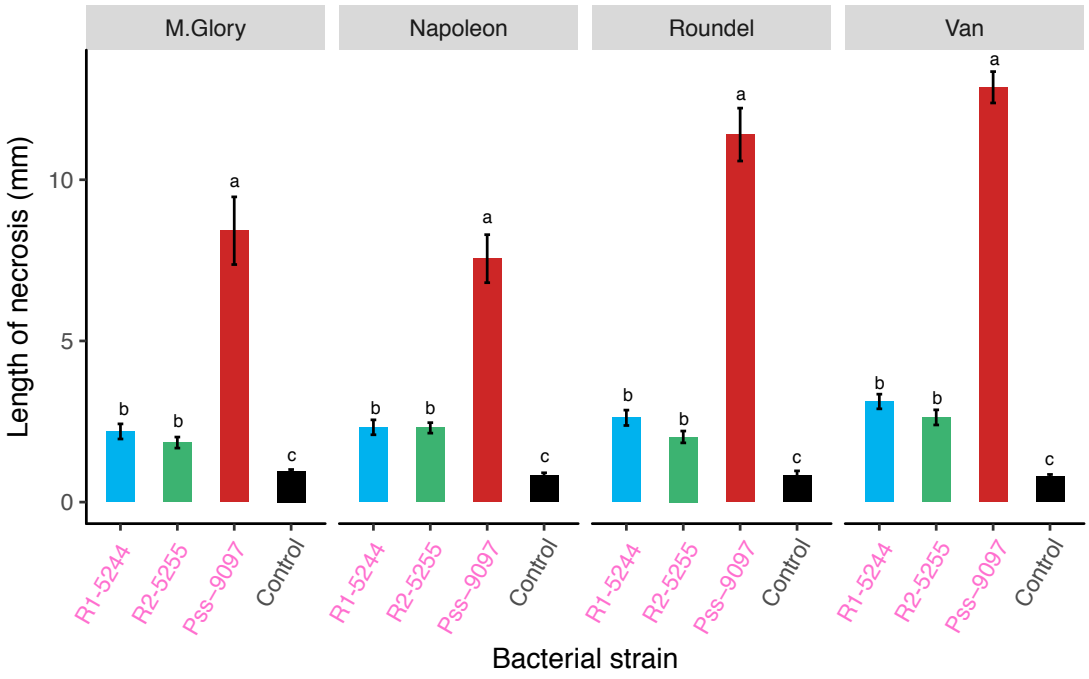


Figure 3.7: Diameter of necrosis caused by cherry pathogens on four cherry cultivars using immature green cherry fruits. *Strains are colour-coded with those isolated from cherry in pink and the no bacterial control in black. The bars are colour-coded:* Psm *R1: blue,* Psm *R2: green,* Pss*: red and no bacteria control: black. Data presented are the mean values of ten replicates, with error bars showing standard error above and below the mean. Tukey-HSD (p=0.05, confidence level: 0.95) significance groups for the different strains for each separate cultivar are presented above each bar.*

3.4.4.3 Inoculation of detached leaves

For comparison, bacterial population growth in detached cherry leaves was also measured. As many of the strains were originally isolated from leaves, those that showed reduced virulence in cherry woody tissues, may show tissue-specificity for leaves. A pilot experiment determined the best method of inoculation was syringe-infiltration (Figure 3.8). The other tested methods (adding a droplet of bacteria, wounding then adding a droplet and stabbing a needle spiked with bacteria) produced

little or no symptoms, likely because there were too few bacteria entering the leaf. Following previous studies on *Arabidopsis* and bean (Cunnac *et al.* 2011; Baltrus *et al.* 2012) a low concentration of inoculum ($2x10^6$ CFU/ml) was used for bacterial population counts.

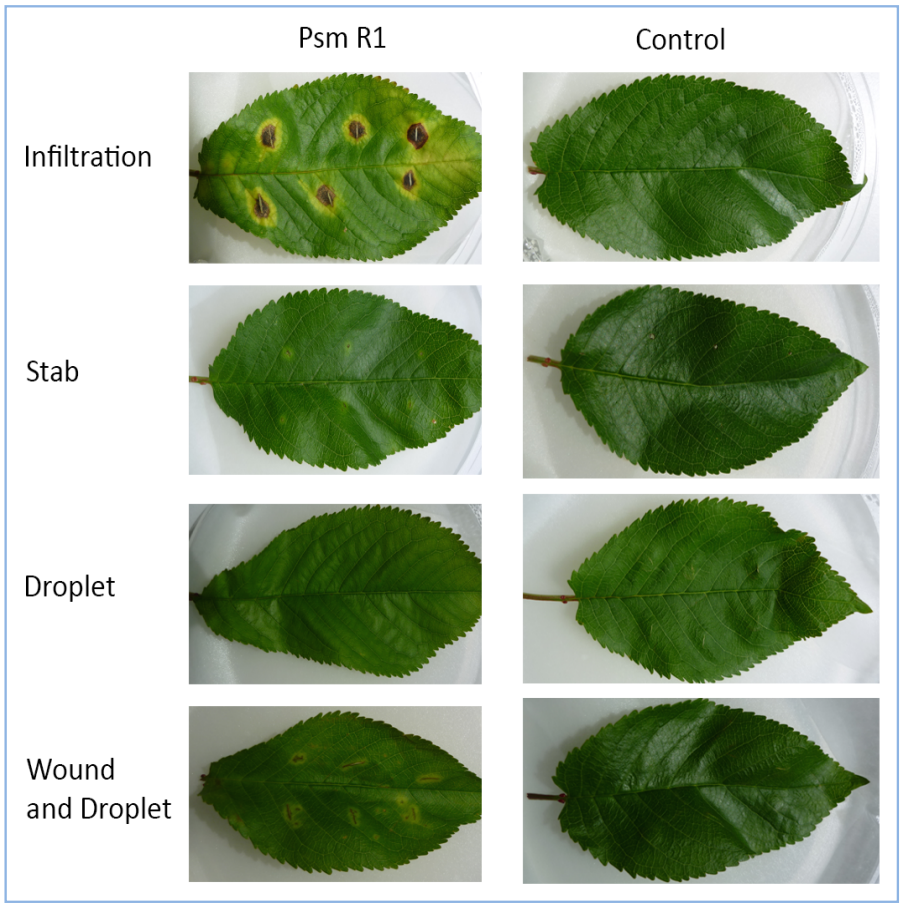


Figure 3.8: Symptoms observed in detached cherry leaves using different leaf inoculation methods*. Representative images of the four methods – infiltration, stab, droplet and wound + droplet. Leaves show inoculation with* Psm *R1-5244 or a 10mM* $MgCl_2$ *control.*

To first assess how the different bacteria behaved *in planta*, a population count of several strains over time was conducted. Figure 3.9 shows the population growth of these strains over time on cherry and plum. On cherry (Figure 3.9A), the pathogens (R1-5244, R2-leaf and *Pss*-9097) exceeded levels of 10^7 CFU/ml within four days and caused black necrosis at the site of infection. The non-pathogen strains, including the plum R1-5300 strain all failed to reach 10^6 CFU/ml even after 10 days *in planta* and failed to produce symptoms. These results support those found on whole-trees, with

only those strains capable of causing bacterial canker being able to reach high levels within leaves. An ANOVA performed on the final day 10 population counts showed a significant difference between strains ($p<0.01$, $df=6$) and the post-hoc Tukey-HSD test indicated that pathogenic strains were significantly different to the non-pathogens. On plum (Figure 3.9B), the pathogens exceeded 10^6 CFU/ml after 4 days. However, some of the strains that were non-pathogenic on cherry were able to grow to similar levels as the pathogens. Again, there were significant differences between strains ($p<0.01$, $df=6$). The *Psm* R1 plum strain R1-5300 and RMA1 isolated from *A. vulgaris* were found to be reach high levels *in planta* and were not significantly different from the pathogenic strains R1-5244, R2-leaf and *Pss*-9097. In the case of R1-5300, this result follows on from its clear pathogenicity on plum in the field experiment, indicating that it is a pathogen of plum but not of cherry. The ability of RMA1 however to grow within plum leaves did not support the field experiment where it caused similar symptoms to the negative control. Plum leaves may therefore lack the non-host resistance occurring towards this pathogen in woody-tissues.

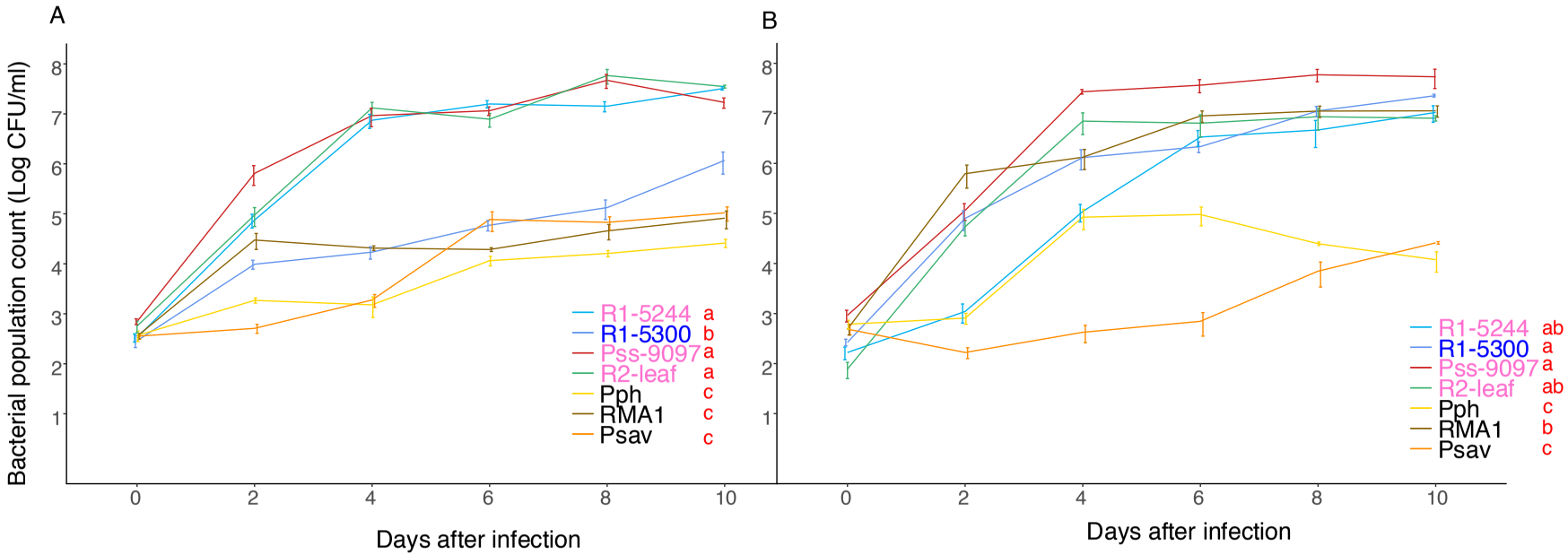


Figure 3.9: Population counts of different strains over time on cherry cv. Van (A) and plum cv. Victoria (B) leaves. *The strains isolated from cherry and plum are highlighted in pink and blue. Population counts are Log CFU/ml. Data presented are the mean values of three replicates, with error bars showing standard error above and below the mean. Tukey-HSD (p=0.05, confidence level: 0.95) significance groups for the different strains (based on day 10 populations) are presented.*

Next, the population growth of all strains used in this study on leaves was tested. An end-point measure of bacterial growth was taken after 10 days (day 0 population counts for all leaf population assays are presented in Figures S3.3-S3.6). All pathogens produced clear necrosis at the site of infection, whilst non-pathogens and the four low virulence *Psm* R1 strains failed to produce symptoms. An ANOVA revealed significant differences between strains ($p<0.01$, df=20). Post-hoc Tukey-HSD analysis grouped the pathogens and non-pathogens into separate groups. The results broadly match the whole-tree wound inoculations on cherry (Figure 3.10). There was clear differentiation of *Psm* R1 strains into pathogens and non-pathogens. The non-pathogenic isolates from other plant species again failed to exceed 10^7 CFU/ml so were clearly differentiated. Non-pathogens such as *Psav* grew to slightly higher levels than in the previous population count (Figure 3.9), showing that there is some variation between experiments, perhaps due to differences in leaf material or inoculum. All pathogenic strains reached approximately 10^8 CFU/ml: Populations may stabilise at this concentration because the available intercellular spaces at the infection site become saturated with bacteria.

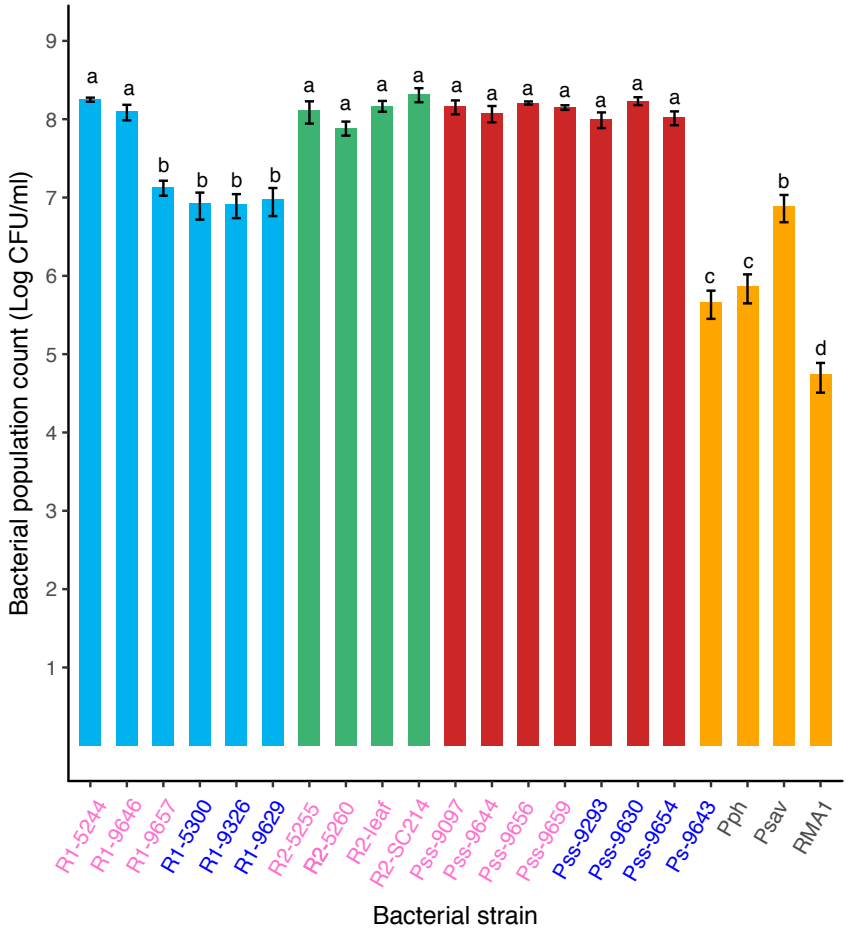


Figure 3.10: Day 10 population count of all strains used in this study on cherry cv. Van leaves. *Strains are colour-coded with those isolated from cherry in pink, plum in blue and non-pathogens in black. The bars are coloured:* Psm *R1: blue,* Psm *R2: green,* Pss*: red. The 10mM* $MgCl_2$ *control is not included as no bacteria were found. The data presented are the mean values of three replicates, with error bars showing standard error above and below the mean. Tukey-HSD (p=0.05, confidence level: 0.95) significance groups for the different strains are presented.*

The bacterial population counts were performed using a low concentration of inoculum, meaning that non-pathogens on cherry never reached concentrations high enough to induce symptoms. One hypothesis was that cherry may be recognising the non-pathogen effector complement and triggering a hypersensitive response, leading to reduced bacterial growth. However, at these low concentrations the HR is likely microscopic, occurring in only a few cells. At higher concentrations the HR becomes macroscopic and can be visualised (Mur *et al.* 2008). Therefore, to determine the best

concentration to detect a HR, symptom devlopment was assessed at different concentrations, for several strains in cherry cv. Van over time. The different inoculum concentrations were obtained using dilution series starting at $2.0x10^8$ and performing 2-fold dilutions until $2.5x10^7$. Figure 3.11 shows the results for all strains at the different concentrations. The scores ranged from 0 (no symptoms), 1 (limited browning), 2 (<50% area brown), 3 (>50% area brown), 4 (completely brown area) to 5 (completely brown with spreading from site of infection). The strains varied in their ability to cause lesions at the different concentrations, particularly the non-pathogen RMA1 which failed to reach a score higher than 2 except at the highest concentration. The *Pss* strain induced rapid lesion formation within 24 hours and sometimes was found to be able to spread slightly from the site of inoculation. At the higher concentrations the final lesions of all strains were similar in appearance. However, the timing of lesion onset clearly differentiated the pathogenic *Psm* strains (R1-5244 and R2-leaf) with all other strains. The non-pathogens (including R1-5300) and *Pss* 9097 all induced rapid lesion formation at the highest concentration, reaching a score of 4 within the first 48 hours, which was suggestive of a HR. The pathogenic *Psm* strains however induced slower symptom development, likely indicative of a hemibiotrophic interaction with the host.

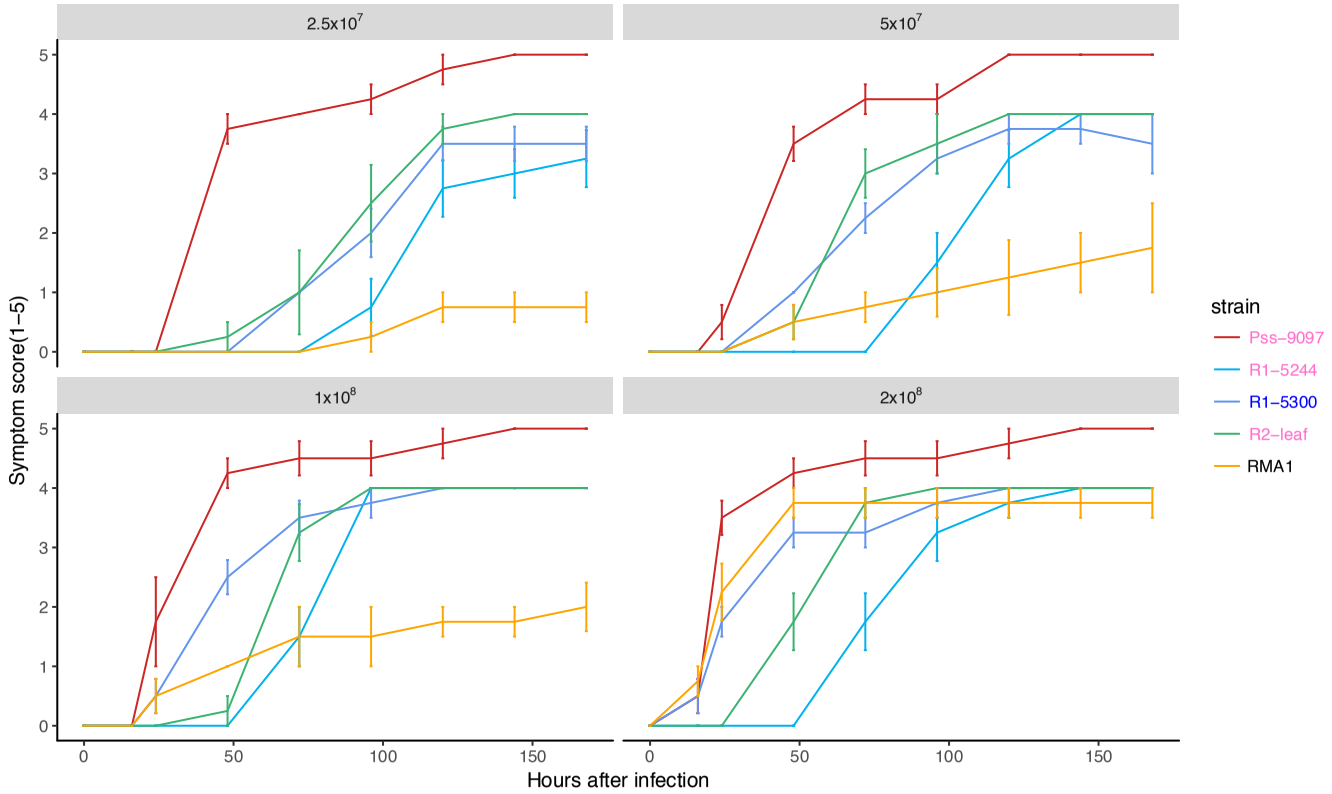


Figure 3.11: Symptom development over time after inoculation of various* P. syringae *strains in cherry cv. Van at different concentrations. *Strains are colour-coded, with those isolated from cherry in pink and plum in blue. Symptoms were scored from 0-5. 0: no symptoms, 1: limited browning, 2: browning <50% of inoculated site, 3: browning >50% of inoculated site, 4: Complete browning, 5: Spread from site of inoculation. Values represent the mean of four replicates and error bars show the standard error above and below the mean. The lines for each strains are colour-coded with* Psm *R1: blue,* Psm *R2: green,* Pss*: red, non-pathogen RMA1: orange.*

To compare the host reactions on cherry and plum, the population count and symptom scoring experiments were repeated with a group of representative strains on both cherry and plum leaves. As before, population counts clearly differentiated pathogens and non-pathogens in cherry (Figure 3.12A). Pathogens exceeded 10^7 CFU/ml and produced necrotic lesions. In comparison, non-pathogens failed to induce symptoms and did not exceed 10^6 CFU/ml. In plum (Figure 3.12B), all strains were able to exceed 10^6 CFU/ml. These differences in host specificity, indicated that immune responses may differ between the two *Prunus* species.

In the symptom scoring experiment all strains gave symptoms in the leaves, however the timing of symptoms was used to differentiate pathogenicity and hypersensitivity. Figure 3.13 shows representative images of symptoms on cherry and plum over time. To statistically analyse the difference in symptom development, the slope of symptom score over time from 0-24 hours was determined. In cherry (Figure 3.12B), there were significant differences ($p<0.05$, df=8) in slope of symptom development from 0-24 hours, between pathogenic *Psm* strains and some of the non-pathogens. In particular, the non-pathogens *Psm* R1 5300, Ps-9643 and RMA1 developed symptoms within 24 hours. This indicated that they trigger a hypersensitive response in cherry leaves, which leads to their reduced population growth. The other non-pathogens (*Pph* and *Psav*) exhibited slower symptom development and were not significantly faster than the pathogen *Psm* R2-leaf. As before, despite being pathogenic and achieving large populations *in planta*, the *Pss* strains 9097 and 9293 triggered rapid symptoms. As they grew to reach high levels *in planta*, they were unlikely to be triggering a HR and these symptoms may instead be due to rapid onset of disease lesions.

In plum, there were significant differences between strains in symptom development from 0-24 hours ($p<0.01$, df=8). Again the two *Pss* strains rapidly induced symptoms. A post-hoc Tukey-HSD test showed that the non-pathogens from other plant species did not produce symptoms significantly faster than the pathogen R2-leaf. Also symptom development of *Psm* R1 5300 was not significantly different from the other *Psm* R1 strain R1-5244, both inducing symptoms after 48 hours post-inoculation, indicating that in plum the two pathogens behave similarly.

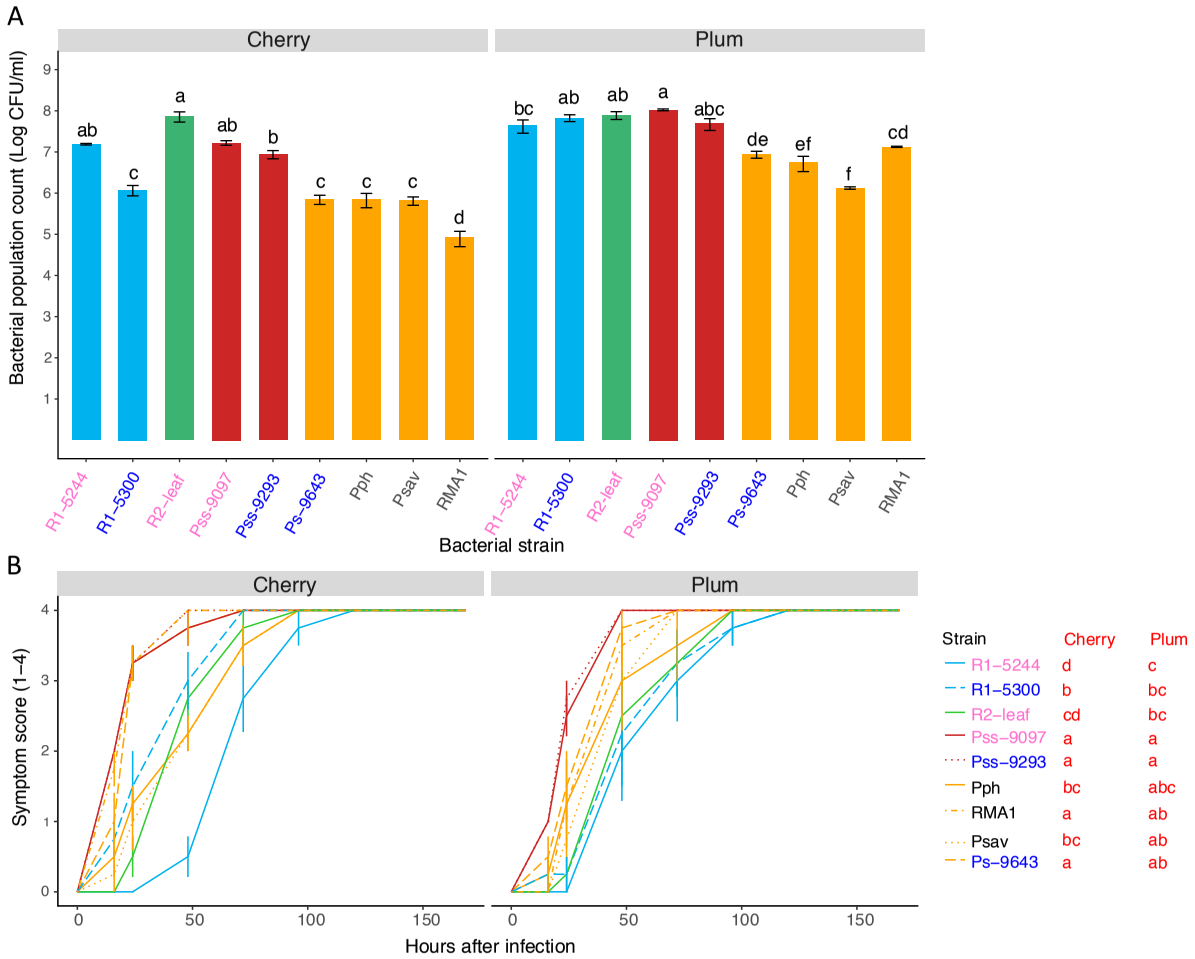


Figure 3.12: Pathogenicity of different strains, assessed by population counts and symptom scores, on cherry and plum leaves*. A: Day 10 population counts on cherry cv. Van and plum cv. Victoria.The names of strains isolated from cherry are pink whilst plum are blue. Strains are colour-coded with* Psm *R1: blue,* Psm *R2: green,* Pss*: red, non-pathogen: orange. Data presented are the mean values of three replicates, with error bars showing standard error above and below the mean. Tukey-HSD (p=0.05, confidence level: 0.95) significance groups for the different strains are presented. B: Symptom development over time. Strains are colour-coded as in A and the data are mean values of four replicates with error bars showing standard error above and below the mean. Tukey-HSD (p=0.05, confidence level: 0.95) significance groups for the appearance of symptoms after 24 hours for the different strains are presented.*

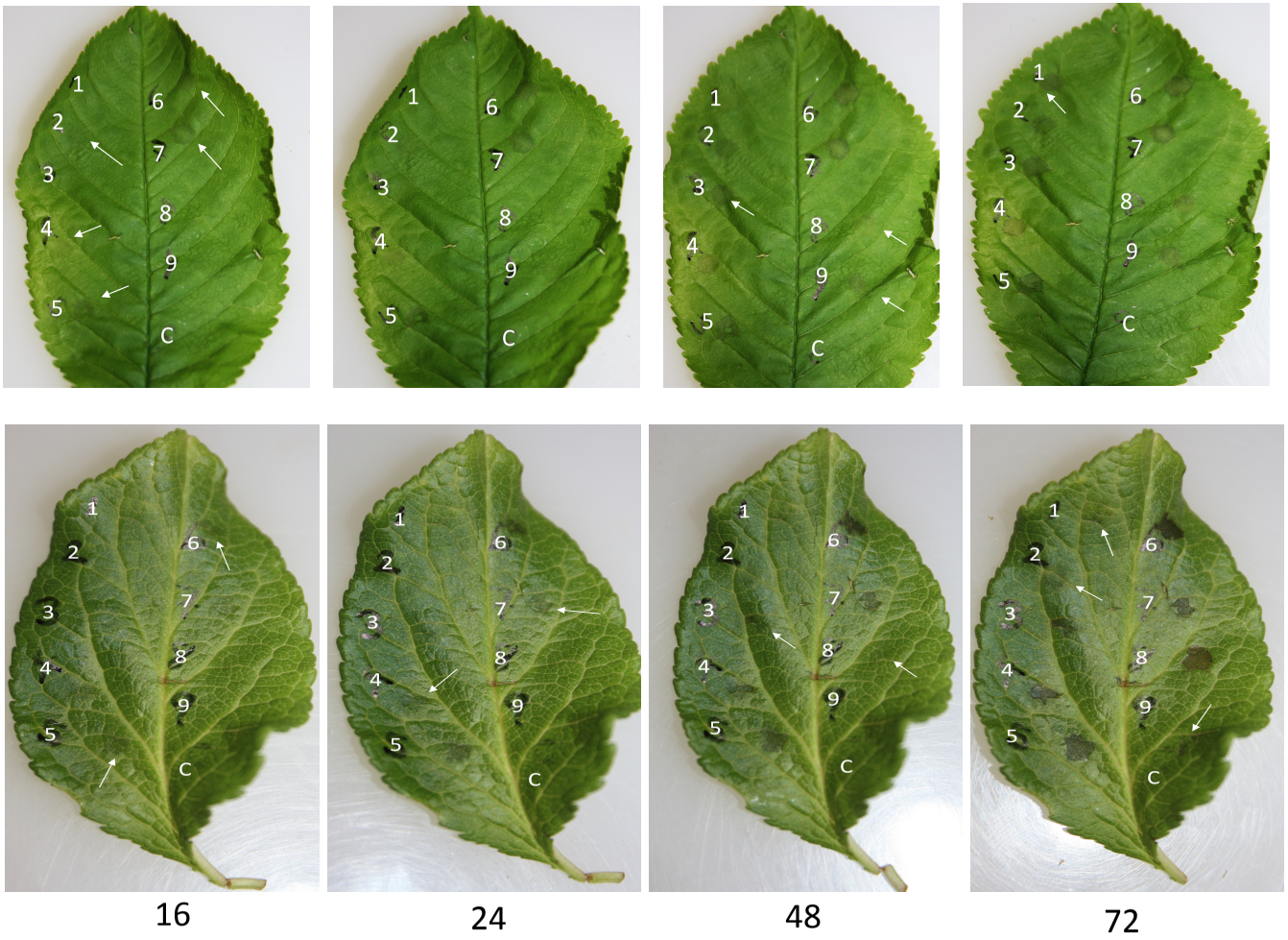


Figure 3.13: Images of symptom development overtime on cherry and plum. A: Cherry cv. Van, B: Plum cv. Victoria. *The same leaf was imaged 16, 24, 48 and 72hpi. Arrows indicate the first appearance of symptoms for that particular strain. Strains are labelled: 1:* Psm *R1-5244, 2:* Psm *R1-5300, 3:* Psm *R2-leaf, 4: Ps-9643, 5:* Pss*-9097, 6:* Pss*-9293, 7: RMA1, 8:* Psav*, 9:* Pph*, C: No bacteria control*

3.4.4.5 Suitability of leaves for resistance screening

The leaf population assay differentiated pathogens from non-pathogens. However, a screen for canker resistance would involve discriminating subtle differences in pathogen growth on different cherry genotypes. To see if detached leaves were able to discriminate cultivar differences, the assay was tested on four cultivars with differences in susceptibility recorded in the field. Strains representing the three cherry-infecting pathovars were tested. Figure 3.14 shows that the three strains were able to grow to exceed 10^6 CFU/ml and cause symptoms in all cultivars. This suggested that on leaves, host-resistance to the pathogens is limited. The leaf system, although useful for comparing strains with very different virulence levels may not be sensitive enough to detect subtle differences between the pathogens found in the field. An ANOVA revealed significant differences between strains, cultivars and interaction between strain and cultivar ($p<0.01$, df=2, $p<0.01$, df-3 and $p<0.01$, df=6). Some of the differences support observations in the field. For example, *Psm* R1 and R2 on the cultivar Roundel show some specificity. However, on Napoleon and Van, unlike in the field experiment, *Psm* R2 grew equally as well as *Psm* R1.

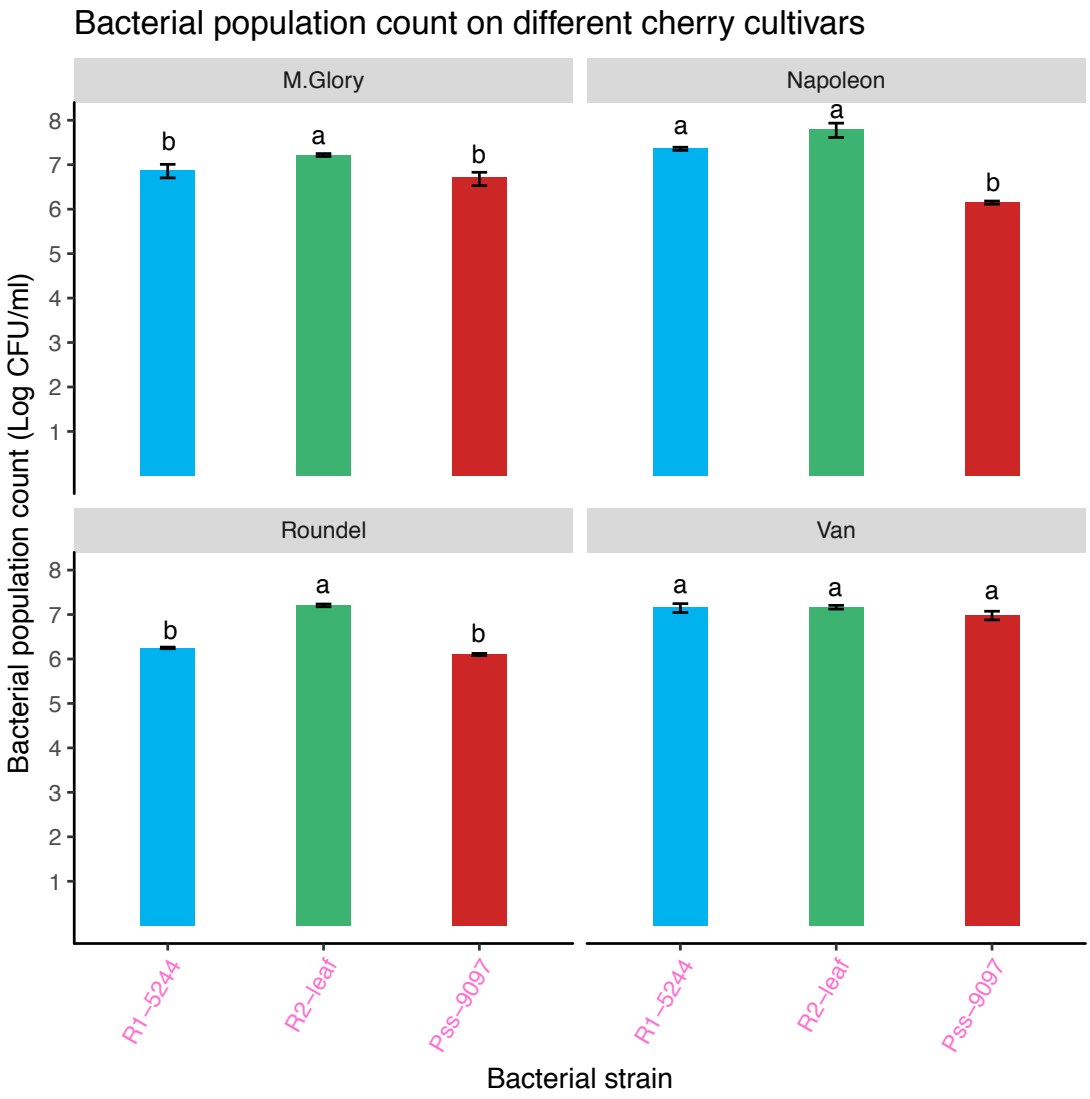


Figure 3.14: Day 10 population counts of three pathogenic* P. syringae *strains on different cherry cultivars. *Strains are colour-coded with* Psm *R1: blue,* Psm *R2: green,* Pss*: red. All three are cherry isolates so the names are coloured pink. The control is not included as no bacteria were found. The data presented are the mean values of three replicates, with error bars showing standard error above and below the mean. Tukey-HSD (p=0.05, confidence level: 0.95) significance groups for the different strains on each separate cultivar are presented.*

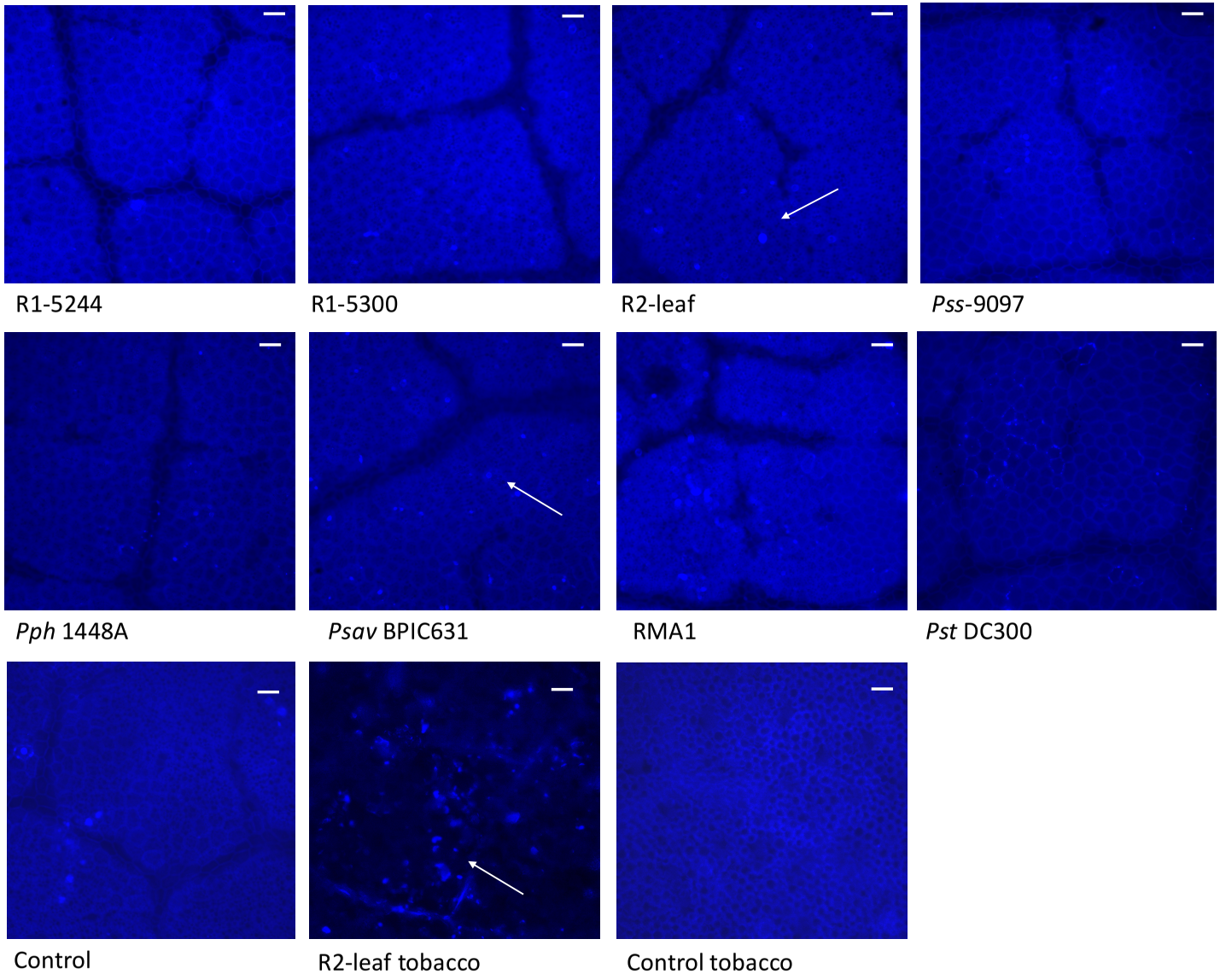
3.4.5 Callose and ROS detection

To further determine how the cherry immune system reacts to both cherry pathogens and non-pathogens, staining can be utilised to detect phenotypes commonly involved in defence reactions. Previous studies in model systems such as *Arabidopsis* and tobacco have developed staining assays that allow rapid visualisation and

quantification of defence responses induced by either *P. syringae* bacteria or molecular elicitors such as flg22 isolated from the flagellum (Torres *et al.* 2005; De Torres *et al.* 2006; Matas *et al.* 2014).

During the PTI response, plants produce a stress-induced defensive *β1-3*-glucan polymer, callose, which is deposited on the symplastic face of the cell wall at the sites of pathogen attack (Ellinger & Voigt 2014). To see if cherry produces callose in response to pathogens and non-pathogens and if the response varies between them, cherry leaves were stained for callose deposition using aniline blue. Cherry leaves were inoculated with *P. syringae* and left to incubate for 48 hours as callose deposition is a late immune response (Schenk & Schikora 2015). The model strain *P.s* pv. *tomato* DC3000 was included for comparison, as this strain has been used in previous callose-deposition assays. Leaf sections were cleared and stained by immersion in aniline blue staining solution before viewing on a fluoresence microscope. Figure 3.15 shows some representative images. Limited callose deposition occurred in all treatment. There were also no immediately obvious differences in callose between the different bacterial strains and the no bacteria control. To determine that the protocol was appropriate, the staining was repeated on *N. tabacum*. Figure 3.15 shows that *Psm* R2 induced callose compared to the no bacteria control in the leaves of this plant species. This suggested that the stain was suitable. The experiment was repeated several times in cherry leaves with little success.

Figure 3.15 (overleaf): Representative fluorescent microscopy images from callose staining experiments of cherry and tobacco leaves. *Images were taken of leaf inoculation sites 48 hours after inoculation with* P. syringae *or a no bacteria 10mM* $MgCl_2$ *control. Leaves were stained for callose using 0.01% aniline blue, cleared in 100% methanol (65°C) and embedded in 50% glycerol solution. These sections were viewed using a fluorescent microscope with a DAPI filter. Scale bar on each image = 50μm. Arrows indicate potential callose deposits.*


R1-5244
R1-5300
R2-leaf
Pss-9097
Pph 1448A
Psav BPIC631
RMA1
Pst DC300
Control
R2-leaf tobacco
Control tobacco

Another well-characterised immune response involves the production of reactive oxygen species in the first few hours after infection. The production of ROS occurs during both PTI and ETI. It is strongly associated with the hypersensitive cell-death induced by avirulence proteins. ROS is detected within plant tissue using a stain called 3,3'-diaminobenzidine (DAB) which precipitates in the presence of hydrogen peroxide. The system was tested first with the model strain *P.s* pv. *tomato* DC3000 which has been used in previous studies involving ROS assays. Tobacco was included as a control to test the methodology was working correctly. Leaves were inoculated with DC3000, stained with DAB and viewed 6 hours after infection to visualise ROS. Figure 3.16 shows some representative leaf samples. The ROS staining worked for DC3000 on tobacco as there is browning of the area of infection (Figure 3.16). However, no cherry leaf samples produced brown pigment at all (Figure 3.16). This lack of staining could be due to the cherry leaves being very difficult to infiltrate compared to tobacco. Several different methods to try and infiltrate the cherry leaves, including using a vacuum pump and a syringe-based infiltration method were trialled with no success.

The results of both callose and ROS staining of cherry leaves were therefore inconclusive. Cherry leaves may be more difficult to infiltrate with staining solutions, and therefore further optimisation of the staining methodology is required.

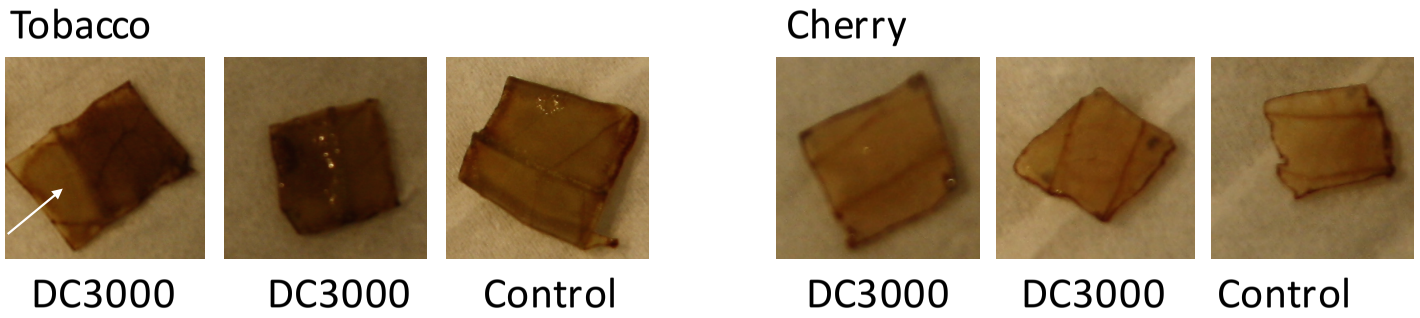


Figure 3.16: ROS staining of leaf inoculation sites *6hpi with the model strain* P.s *pv.* tomato *DC3000 or a no bacteria 10mM* $MgCl_2$ *control. Leaves were stained with DAB solution and cleared with 100% methanol (65°C) before visual assessment. Brown precipitate was evident for inoculations of tobacco leaves (arrow), however absent in cherry leaves.*

3.4.6 Electron microscopy

Finally, electron microscopy was used to visualise the interactions occurring at the cellular level between *P. syringae* and cherry. Detached leaves were infiltrated with a pathogen (*Psm* R2-leaf), non-pathogen (RMA1) (both were at $2x10^6$ CFU/ml which is the concentration used in population counts) and a negative 10mM $MgCl_2$ control. The bacteria were left to grow for one week and then, with the help of Dr Ian Brown (University of Kent), the leaves were sectioned and viewed using transmission electron microscopy. Figure 3.17 shows some representative images of the pathogen *Psm* R2-leaf. There are clear examples of the pathogen residing next to plant cells and not inducing any response, indicating it is suppressing the host immune system. Some regions were found where the plant cells appeared to have responded to bacterial colonies. There was restructuring of the cell wall with putative papilla formation (red arrows in Figure 3.17). Next to these formations there were dead bacterial cells. It is unclear if the bacteria triggered the cell wall restructuring when dead or alive. Unfortunately, the non-pathogen RMA1 did not grow to sufficient levels to detect any bacteria.

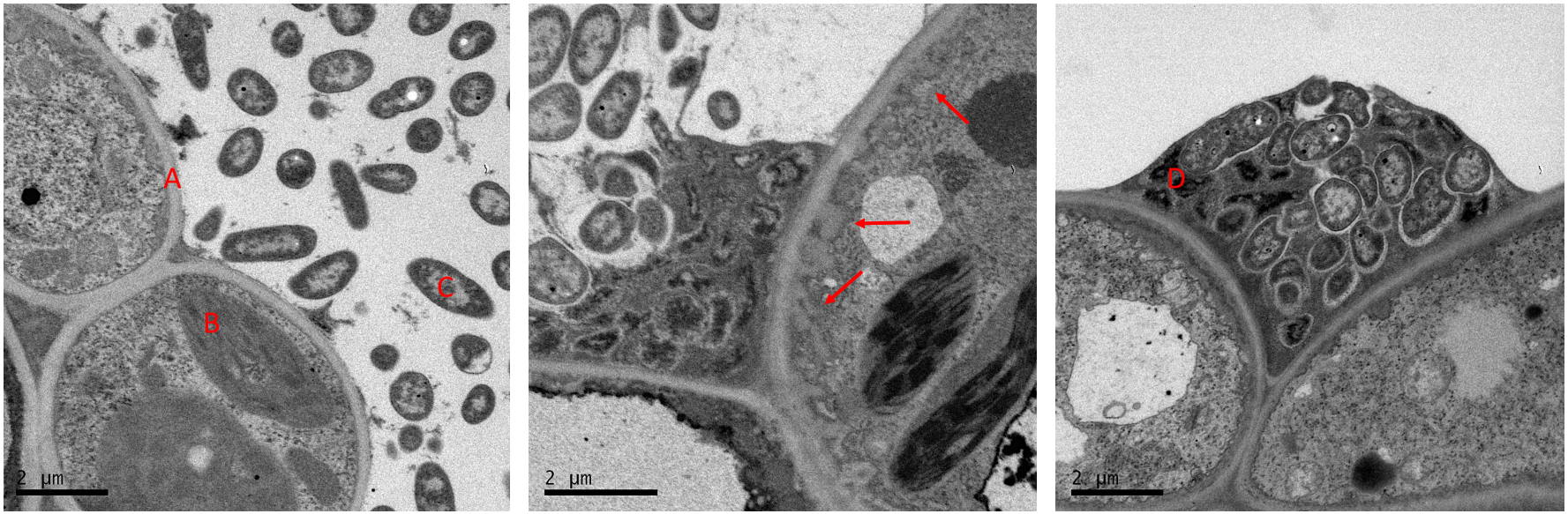


Figure 3.17: TEM images of* Psm *R2-leaf in a detached cherry leaf one week after inoculation. *Arrows point to putative papilla formation in the plant cell wall next to a bacterial colony containing dead bacterial cells. A: Plant cell wall, B: Plant cell chloroplast, C: Bacterial cell, D: Bacterial colony. Scale bar = 2µm.*

3.4.7 Comparison of pathogenicity tests

To compare the results of all laboratory-based assays to the field experiments, a series of correlation matrices comparing the means of each inoculation type were produced. (Figure 3.18). Figure 3.18A presents a line plot which visually highlights how each of the three pathogenic strains (Psm R1, Psm R2 and Pss) performed across the four cherry cultivars. For example, *Psm* R1 caused the greatest level of disease on cv. Van in the wound, cutshoot and fruit assays. Whilst on leaves and in the leaf scar test this strain caused the greatest disease on Napoleon. The heatmap shows that generally the cut shoot assay results agreed with those of the field experiment. For example, *Psm* R2 caused the greatest disease on Roundel in both leaf-scar, wound and cut shoot experiments. The fruit and leaf assays however did not represent the results seen in the field or on shoots. *Psm* R2 grew to higher levels in Napoleon leaves and caused greater necrosis on Van fruits, despite its higher level of virulence on Roundel in the tests on woody tissues. These differences may be due to tissue-specific differences in pathogen virulence or host resistance. As the assays performed on fruit and leaf tissue take only a few days, they are unlikely to be sensitive enough to detect the subtle differences seen in the field, in a disease that takes many months to develop.

***Figure 3.18 (overleaf): Comparison of different inoculation methods**. A: Line plots of mean standardised disease score for each bacterial strain on each cherry cultivar for each different inoculation type. Error bars show standard error above and below the mean. Lines were drawn to connect the means for each inoculation type for visualisation purposes only. B:* Psm *R1 correlation matrix of mean disease scores between inoculation types. Statistically significant correlations ($p<0.05$) are in boxes. The different inoculation types are colour coded based on the colours in the corresponding graph. Correlation coefficients are colour coded based on values between -1 and 1 shown on the scale. C:* Psm *R2 correlation matrix of mean disease scores between inoculation types. D:* Pss *correlation matrix of mean disease scores between inoculation types. E: Correlation matrix of the combined data set (including all bacterial strains).*

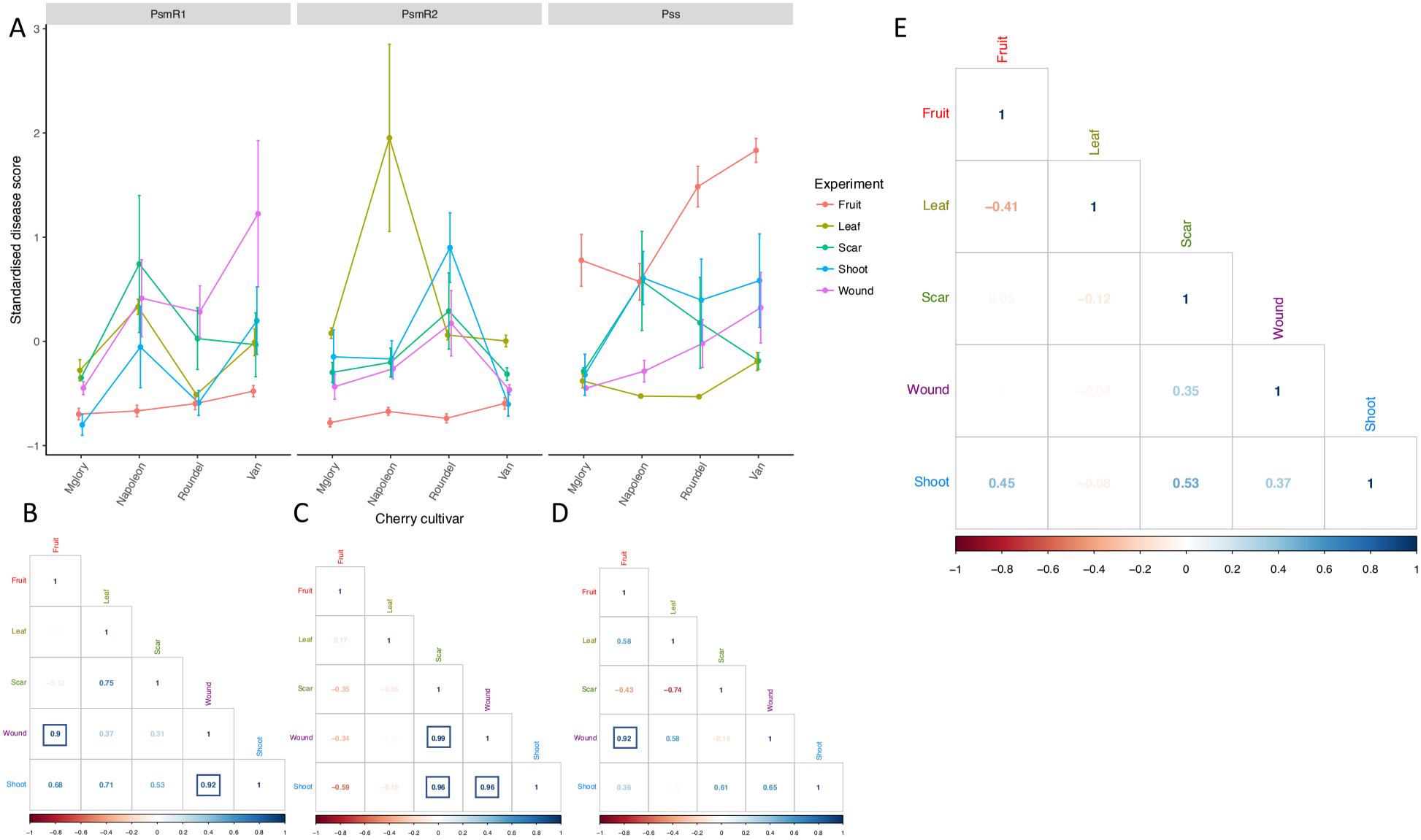

A
PsmR1
PsmR2
Pss
Standardised disease score
Mglory
Napoleon
Roundel
Van
Cherry cultivar
Experiment
Fruit
Leaf
Scar
Shoot
Wound
B
C
D
E

Figure 3.18 visually presents how the virulence of each bacterial strain varied on different cherry cultivars depending on inoculation method. For example, *Psm* R1 scored highly in most experiments on Napoleon and Van, whilst *Psm* R2 caused the greatest disease scores on Roundel in all tests apart from the fruit inoculations and the leaf population assay. To determine if the results using different inoculation methods correlated with each other, correlation matrices were produced for the mean disease scores of each bacterial strain. For *Psm* R2 (Figure 3.18C) there was clear correlation between the inoculation of woody tissues (shoots, leaf scar and wound inoculations), whilst leaf populations and fruit inoculations did not correlate with the other experiments. Similarly, in *Psm* R1 (Figure 3.18B) the cutshoot assay correlated with cherry wound inoculations indicating that this lab-based assay is reflecting the virulence of this strain during wound inoculations in the field. By contrast, for *Pss* (Figure 3.18D), the woody tissue inoculations were less well correlated with each other, whilst the fruit assay correlated well with whole-tree wound inoculations. An overall correlation matrix combining the results of all bacterial strains (Figure 3.18E) indicated that no tests correlated strongly, however there was generally stronger correlations between the woody tissue inoculation experiments.

3.5 Discussion

In this study, the virulence of a range of *P. syringae* strains on cherry and plum was characterised and a subset of these strains was used to screen cultivars for canker resistance. Breeding for resistance towards this complex disease is particularly challenging due to the large number of divergent strains that appear to be pathogenic. The John Innes cherry breeding programme initially screened for resistance to *Psm* R1 in the 1960s-70s, only to later discover some of their new selections were susceptible to the newly discovered *Psm* R2 (Matthews 1979; Garrett 1986). Host resistance to cherry canker is likely to be multi-factorial and potential mechanisms of resistance towards the different cherry-infecting clades may operate at different stages of the disease cycle. Pathogenicity towards cherry has arisen multiple times in the *P. syringae* species complex, with the three main clades falling in separate phylogroups.

Therefore, as the different clades have convergently evolved, it is likely that host resistance mechanisms targeted towards them may differ significantly.

3.5.1 Whole-tree tests revealed differences in pathogenicity and host susceptibility to *P. syringae*

First, the ability of individual bacterial strains to cause cherry canker was assessed using a glasshouse whole-tree inoculation. This provided a baseline to compare with the results of laboratory-based assays, that may or may not correlate with ability to cause canker. Although strains of *Psm* R1 were phylogenetically indistinct, they could be divided into pathogenic and non-pathogenic isolates, with non-pathogenic isolates failing to cause gumming and black necrosis. Non-pathogenic strains isolated from distantly related plant species were unable to cause disease, supporting the theory that individual pathovars are mostly specialised to their particular host plant (Sarkar *et al.* 2006). All *Psm* R1 isolates from plum were non-pathogenic on cherry. Their lack of cherry pathogenicity may be due to host-specific factors. By contrast, all isolates of *Pss* (from plum and cherry) were able to causing necrosis and gumming, indicating that these strains exhibit a greater host range.

A subset of strains with variable virulence levels were then pathogenicity tested under field conditions, in assays which should be most representative test of natural disease. The different host-specificities of *Psm* R1 strains on cherry and plum was confirmed. The cherry isolate *Psm* R1 5244 was pathogenic to both cherry and plum, whereas *Psm* R1 5300 was only pathogenic on plum trees. This is an interesting result as phylogenetics revealed this clade to be highly homogeneous (Figure 3.1). As the phylogenetic analysis was based only on core house-keeping genes in the core genome, it may be missing divergence in the flexible genome that are responsible for differences in pathogenicity. Genomic analysis of these strains could reveal important differences in virulence factor repertoires that dictate host specificity. Interestingly the results support studies done at East Malling Research looking at *Psm* R1 host specificity (Crosse, & Garrett 1970). *Psm* R1 was originally designated as a race based on differences with *Psm* R2, however it is now known that these are two divergent clades, so should not really be designated races of the same pathovar. However, the

differences in pathogenicity of members of *Psm* R1 may indicate that, at least within the bacterial populations occupying orchards in UK, there may be a race structure within this clade, with the different groups varying in ability to infect different *Prunus* species. Members of the group containing *Psm* R1 5300 may be restricted in growth on cherry due to the expression of avirulence factors. Further sampling of a diverse range of strains from different *Prunus* species and cultivars would be needed to confirm this.

The field inoculations of cherry revealed significant differences between host cultivars. In both wound and leaf-scar inoculations the cultivar Merton Glory exhibited a broad level of tolerance to all three pathogenic clades. This cultivar is therefore a candidate for further study of the mechanisms underlying resistance. Although the REML analysis did not identify an overall strain by cultivar interaction, there was variation in resistance to *Psm* R2. This strain caused only limited disease on Napoleon and Van, whilst the cultivar Roundel was highly susceptible. Napoleon and Van are therefore candidate cultivars exhibiting race-specific resistance. Overall on cherry, all pathogenic strains seemed able to invade both through both wounds and leaf scars. This does not support previous reports that *Pss* and *Psm* R2 are less able to invade leaf-scars than *Psm* R1 (Freigoun & Crosse 1975; Crosse & Garrett 1966). Differences in experimental procedure could have led to this result, as the original studies used fully mature trees, which may have contrasting resistance mechanisms to the young trees used in this study. The rootstocks used in this study were also different to past experiments. It is widely known that rootstocks are known to communicate with the scion via the transfer of metabolites and phytohormones, which may influence scion traits including dwarfing, fruit quality and disease resistance (Albacete *et al.* 2015). As various studies have identified variability in the susceptibility of different rootstocks to canker (Garrett 1979; Krzesinska *et al.* 1992), it would be interesting to determine if rootstocks can influence the progression of the disease, by altering resistance responses in scion wood.

The field experiment on plum produced significant differences between strains, but not between cultivars. The cultivar Marjorie's Seedling was not susceptible to any

strains inoculated through leaf-scars indicating this is unlikely to be a natural entry point for pathogens. Indeed, previous reports suggest that plum pathogens do not naturally enter through the leaf scars (Crosse 1966). However, some pathogenic strains were capable of causing disease on the cultivar Victoria when inoculated through leaf scars.

3.5.2 Comparison of laboratory-based assays

With a set of strains whose virulence had been confirmed on whole trees, several more rapid laboratory-based assays could be tested for their suitability for resistance screening. Assuming that field inoculations (leaf scar and wound) are closest to the natural disease, the results of lab-based tests were correlated (Figure 3.18) with those of the field study to see if they are truly representative of the of the disease. Several studies have reported the use of woody shoots as a rapid screening method (Santi *et al.* 2004; Farhadfar *et al.* 2016). This assay was tested using several strains of known virulence. In collaboration with Dr Bo Li, a rapid image analysis program was developed to automatically determine the percentage area of necrosis of shoot images (Li *et al.* 2015 paper in Appendix). The assay differentiated pathogens and non-pathogens on cherry shoots and was sensitive enough to detect variation in pathogen virulence on the different cultivars. The results of this test were correlated with the leaf-scar and wound field experiments, although the correlations varied depending which bacterial strain was used (Figure 3.18). For example, *Psm* R2 caused the most necrosis on cv. Roundel in all three woody tissue inoculation tests, indicating there is strong support that this cultivar is highly susceptible to *Psm* R2. In addition, differential virulence of *Psm* R1 and *Psm* R2 on cv. Van was supported by both the cut shoot test and field experiments, suggesting this cultivar may possess some resistance to *Psm* R2. However, on cv. Napoleon the field experiment showed *Psm* R1 to be more virulent than *Psm* R2, a result that supports previous studies (Freigoun & Crosse 1975; Garrett 1978), but this difference was not seen on the cut shoots. Therefore, the use cut shoot tests to assay pathogen virulence may not always represent virulence levels that are seen in the field. Other studies that utilised cut shoot inoculations of *P. syringae* and fungal canker-causing pathogens have found differences in virulence of the same isolate between field and laboratory results (Farhadfar *et al.* 2016; Gomez-

Cortecero *et al.* 2016). Overall, the cut shoot assay provided a means to rapid high-throughput screening, with speed aided through automated image analysis of shoots. Such a test is highly suitable for screening for resistance to bacterial canker and other tree diseases. Therefore, the results of such tests may narrow down a list of putatively resistant lines, but resistance would still need to be confirmed using whole-trees.

Next, an immature fruit assay was assessed. A screen of all strains used in this study revealed that the fruit test was clearly able to differentiate pathogenic *Pss* strains from other isolates. However, it was more difficult to differentiate pathogenic *Psm* from non-pathogens based on diameter of necrosis as the lesions caused by *Psm* were small. This assay was then tested on different cultivars of cherry. Due to the size of *Psm* lesions, it was not sufficiently sensitive to pick up any of the *Psm* race-specific cultivar differences found in the cut shoot and field experiments. The differences observed in race-specific resistance of the different cherry cultivars in the field and shoot experiments may be tissue-specific, and therefore resistance phenotypes in fruit may differ from those found in dormant woody tissues. The limited availability of immature fruits (1-2 months per year) also made them less suitable for use in screening assays.

Finally, the use of detached leaves for screening in the laboratory was explored. Pathogenicity was assessed both through symptom development and population counts of bacteria extracted from leaf material. Cherry leaf population counts clearly discriminated pathogenic and non-pathogenic strains. When inoculated at a low concentration only pathogenic strains (R1-5244, R1-leaf and Pss-9097) were able to cause disease lesions on cherry, which appeared 7-10 dpi. On Plum, the non-pathogen RMA1 and plum isolate R1-5300 were able to grow to similar levels to the cherry pathogens (Figure 3.10) and cause disease symptoms. The non-pathogens *Pph* and *Psav* failed to produce disease symptoms. The fact that RMA1 was able to grow to high levels in plum leaves does not correspond to its pathogenicity in the field assay (Figure 3.4). Interestingly, in the cutshoot assay (Figure 3.5) RMA1 caused necrosis on plum similar to the *Pss* pathogen. The field experiment showed that RMA1 is not a

true pathogen of plum, however its virulence in the lab-based assays may indicate it has adaptive potential to cause disease when inoculated in unnaturally high concentrations directly onto plant tissue. Its inability to cause any disease on cherry in all lab-based assays, indicated that cherry may exhibit a robust non-host immune response towards this non-pathogen, which is different to that expressed in plum.

Despite their utility in differentiating pathogenicity, when a leaf population count on different cherry cultivars was performed, all three cherry pathogenic clades were able to exceed 10^6 CFU/ml *in planta* and cause symptom development in the area of inoculation. The presence of similar symptoms meant that this method would not be very applicable for large-scale screening. However, the ANOVA did produce strain and cultivar differences and an interaction. Figure 3.18 shows that the results of this assay did not correlate with the results obtained for each bacterial strain on the different cultivars in the field. This suggested that this test alone cannot be used to screen different cultivars for resistance to mimic that seen in woody tissues in the field.

Overall, correlation analysis (Figure 3.18) indicates that the virulence of different bacterial strains varied between inoculation methods. This is interesting as it shows that they behave differently on different plant tissues. This reflects the complexity of this disease as each strain may be more/less adapted to survive and cause disease on each plant tissue type. In general, woody tissues (whole tree inoculations and cut shoots) correlated well indicating that resistance screening must involve one or more of these tests.

3.5.3 Timing of symptom development on cherry leaves is indicative of pathogenicity or the HR

Leaves provided a good model for studying the cherry immune system within the laboratory as the hypersensitive and pathogenic responses could be differentiated based on symptom timing. When inoculated at high concentrations all strains produced necrotic lesions, however non-pathogens were found to induce symptoms earlier than pathogens of *Psm* R1 and R2. The activation of the HR may mean that effector-triggered immunity is operating against non-pathogens in cherry leaves, and

differences in effector repertoires between cherry-infecting strains and non-pathogens could reveal those effectors that are detected. In particular, there were clear differences in pathogenicity of the two *Psm* R1 strains on cherry, which agreed with the whole-tree assay. The HR was clear for non-pathogens *Psm* R1 5300, Ps-9643 and RMA1, whereas symptom development associated with *Pph* and *Psav* was slower. This slower onset of symptoms may mean that any hypersensitive response induced by these strains is weaker or that more basal resistance mechanisms such as PAMP-triggered immune responses are responsible for preventing their population growth in leaves. Interestingly, although *Pss* strains reached high population levels in the leaves, they triggered symptom development at a similar rate to the HR caused by non-pathogens. *P. syringae* is traditionally described as a hemi-biotrophic pathogen (Lindeberg *et al.* 2012), with delayed symptom onset during the biotrophic phase followed by symptoms during a more necrotrophic phase. The results indicated that on leaves *Pss* may be more necrotrophic as it triggers symptoms very quickly. Further study could reveal the factors causing these rapid symptoms. The production of non-ribosomal peptide toxins is common in strains of phylogroup 2, which includes *Pss* (Dudnik & Dudler 2014), and if expressed early could cause the necrotic symptoms seen. This is supported by a study of *Pss* toxins (Yin-Yuan & Gross 1991) that showed that syringomycin is expressed within the first 24 hours of inoculation of immature cherry fruits. *Pss* could also be deliberately triggering the HR like other necrotrophic pathogens to aid disease development (Govrin & Levine 2000).

To further understand the immune responses occurring, leaves were stained for both ROS and callose, which are produced during the different stages of plant immunity in many plants. Unfortunately, both assays were inconclusive. Both assays were based on studies of model systems such as *Arabidopsis* and tobacco. Successful staining of ROS and callose was achieved in *N. tabacum* leaves, indicating that the staining solutions and imaging technique were appropriate. The differences in the leaf composition of cherry may have prevented adequate uptake of the stain solution. This protocol therefore would require further optimisation if used in future pathology experiments.

Transmission electron microscopy can also be used to study *P. syringae* – plant cell interactions. Leaves were inoculated with *Psm* R2. Numerous sites were observed at which bacterial cells were adjacent to plant cells but had not induced any cell wall or cytoplasmic alterations, indicating that the pathogen is either suppressing or failing to activate the plant immune system. There was some evidence of cell wall papilla formation, particularly near dead bacterial cells.

3.5.4 Conclusions

This study has focused on the detailed analysis of pathogenicity in strains used for genome sequencing. Results show that representatives of the three clades of *P. syringae* that cause bacterial canker may utilise distinct mechanisms of virulence and trigger differing host resistance mechanisms in cherry. A hypersensitive response is putatively triggered by non-pathogenic strains in leaves, indicating that effector-triggered immunity may be operating in cherry. Cherry leaves and fruit failed to sufficiently reveal varietal differences to the same extent as experiments on woody tissues. This suggests that some resistance mechanisms are tissue-specific. A whole range of complex variable traits could be involved in these varietal differences in susceptibility. These include timing of leaf drop, rootstock, phellogen activity and differences in leaf-surface bacterial populations which act as inocula for wood infections, as discussed by Crosse (1966). Breeding resistance to at least three rather distinct groups of a pathogen remains a challenging prospect. Cultivars such as Merton Glory that exhibit resistance to all three clades may be useful for determining the genetic basis of broad-spectrum resistance mechanisms.

Chapter 3: Supplementary methods

Cherry and plum field inoculations

Experimental design:

- 10 replicated blocks of the experiment.
- 4 varieties of cherry (Mgl, Nap, Rou, Van) and 2 varieties of plum (Mse, Vic)
- 8 bacterial strains and a 10mM $MgCl_2$ control (R1-5244, R1-5300, R2-leaf, Pss9097, Pss-9293, Ps9643, *Pph*, RMA1)
- 2 inoculation methods (leaf scar and scalpel wound)

To minimise the number of trees required, different strains were inoculated onto the same tree. As a minimum amount of space of 4 buds between sites of inoculation was required (to prevent merging of the disease lesions), two trees were required for all strains. Therefore, two trees were effectively one experimental unit (or tree). The 8 strains and 2 controls (1 per tree) were randomised across the two trees (Figure S3.1) to control for position effects. This involved an R script that randomly assigned the position of the control on tree number 1 and then randomly placed 4 strains on this tree. The script then randomly positioned the other control on tree number 2, and the four remaining strains were assigned to the other positions. The position of these two-tree units was then completely randomised in ten blocks in the field. Each block contained all different combinations of plant variety and inoculation method, with 24 trees per block.

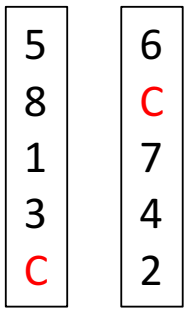


Figure S3.1: Diagram of randomisation of* P. syringae *strains across two trees in field experiment. *A control (in red) was randomly positioned on both trees and then the 8 strains were randomly positioned.*

To analyse the experiment

The variances of the two inoculation methods (leaf-scar and wound) was very different meaning that the two methods were analysed separately. The host (cherry and plum) variances were also different.

Variances:

- Cherry scar: 1258.317
- Plum scar: 132.8194
- Cherry wound: 1844.474
- Plum wound: 331.2786

Therefore, the analysis was done for each plant species and inoculation method combination e.g. cherry leaf-scar, cherry wound, plum leaf-scar and plum wound. There were a lot of missing values in the final data set due to completely dead trees. A REML analysis was used:

model <- lmer(log2(length+0.1) ~ cultivar * strain + (1|block/no.)

The model treated cultivar and strain as fixed effects and block (1-10), with tree number (position 1-24 within the block) nested within as random effects.

Leaf population counts

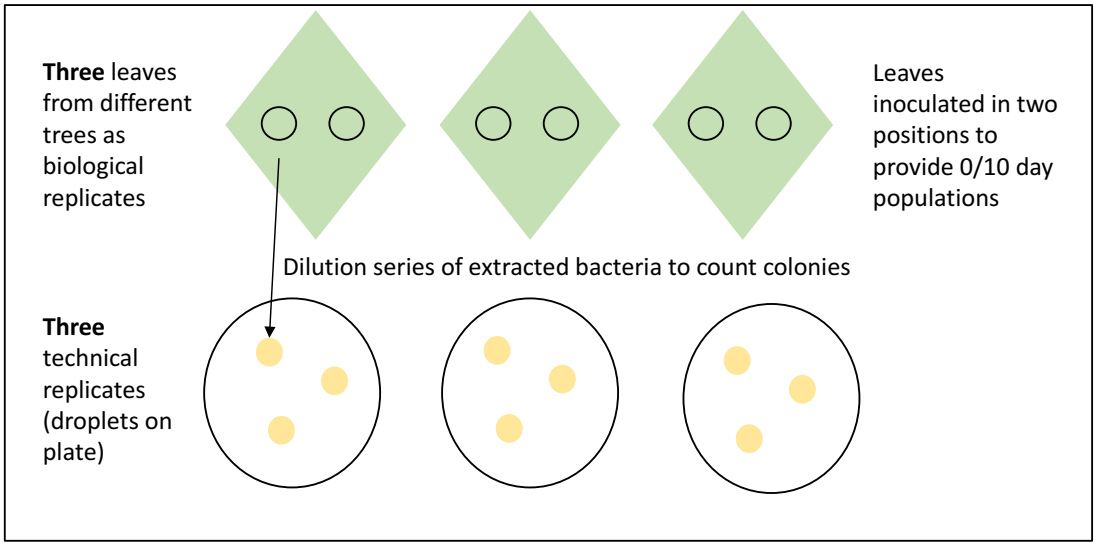


Figure S3.2: Diagram depicting experimental design of population counts. *Three replicate leaves were taken from three cherry trees. These were inoculated at least twice. Bacteria were extracted and serially diluted. For each concentration, there were three technical replicates (10 μl droplets).*

Chapter 3: Supplementary results

Day 0 population counts were measured to ensure that bacteria were present in the leaves and that inoculum were of a similar concentration.

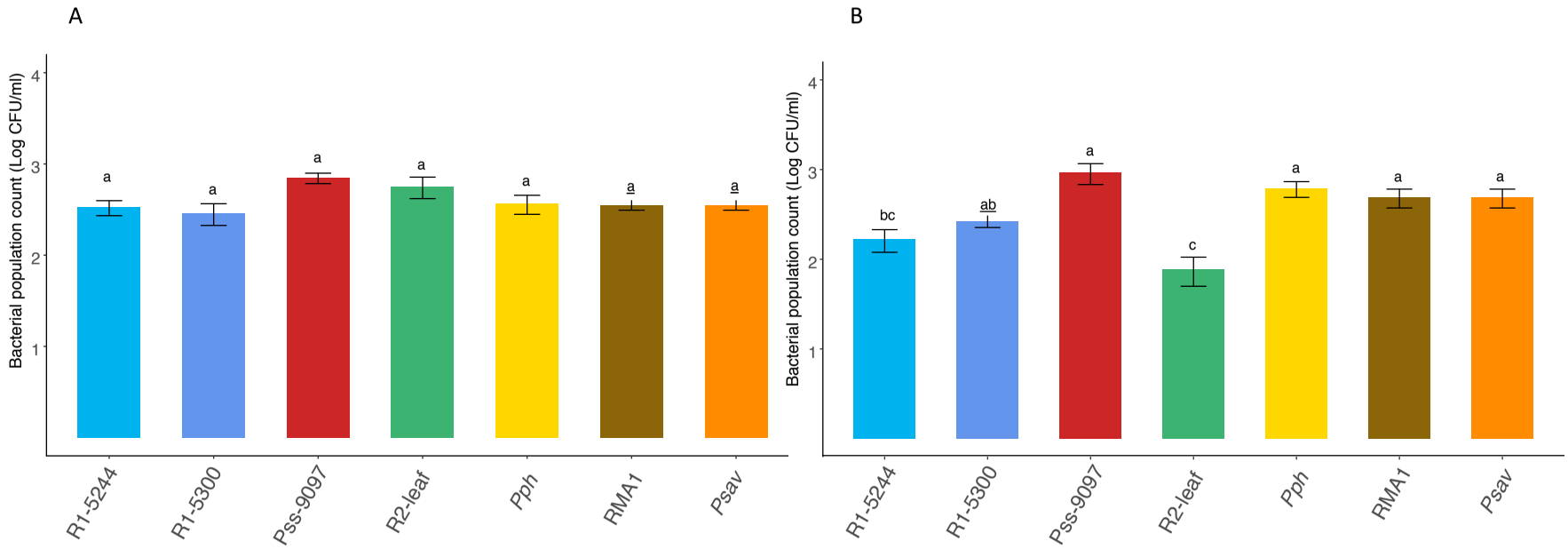


Figure S3.3: Population counts of different strains 0 dpi on cherry cv. Van (A) and plum cv. Victoria (B) leaves (for Figure 3.9). *Population counts are Log CFU/ml. Data presented are the mean values of three replicates, with error bars showing standard error above and below the mean. Tukey-HSD (p=0.05, confidence level: 0.95) significance groups for the different strains are presented above each bar.*

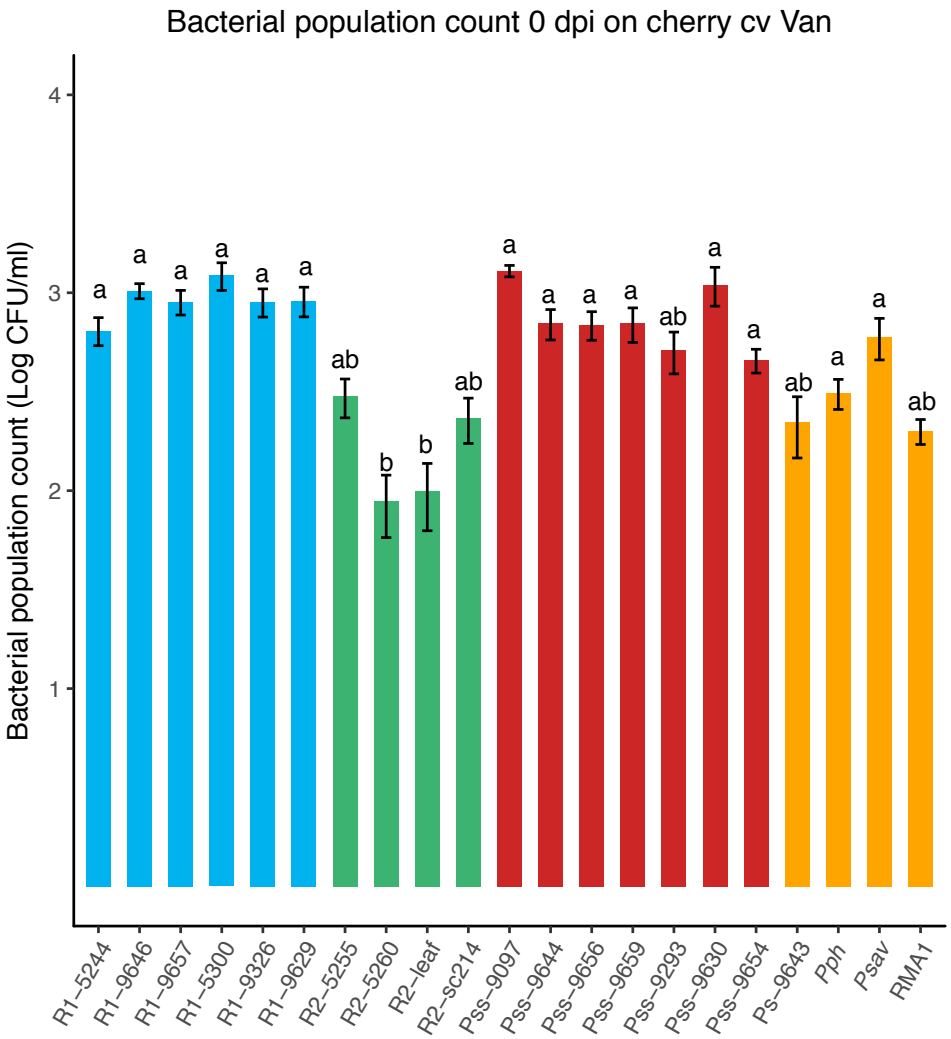


Figure S3.4: Day 0 population count of all strains used in this study on cherry cv. Van leaves (for Figure 3.10). *Strains are colour-coded with* Psm *R1: blue,* Psm *R2: green,* Pss*: red. The control is not included as no bacteria were found. The data presented are the mean values of three replicates, with error bars showing standard error above and below the mean. Tukey-HSD (p=0.05, confidence level: 0.95) significance groups for the different strains are presented.*

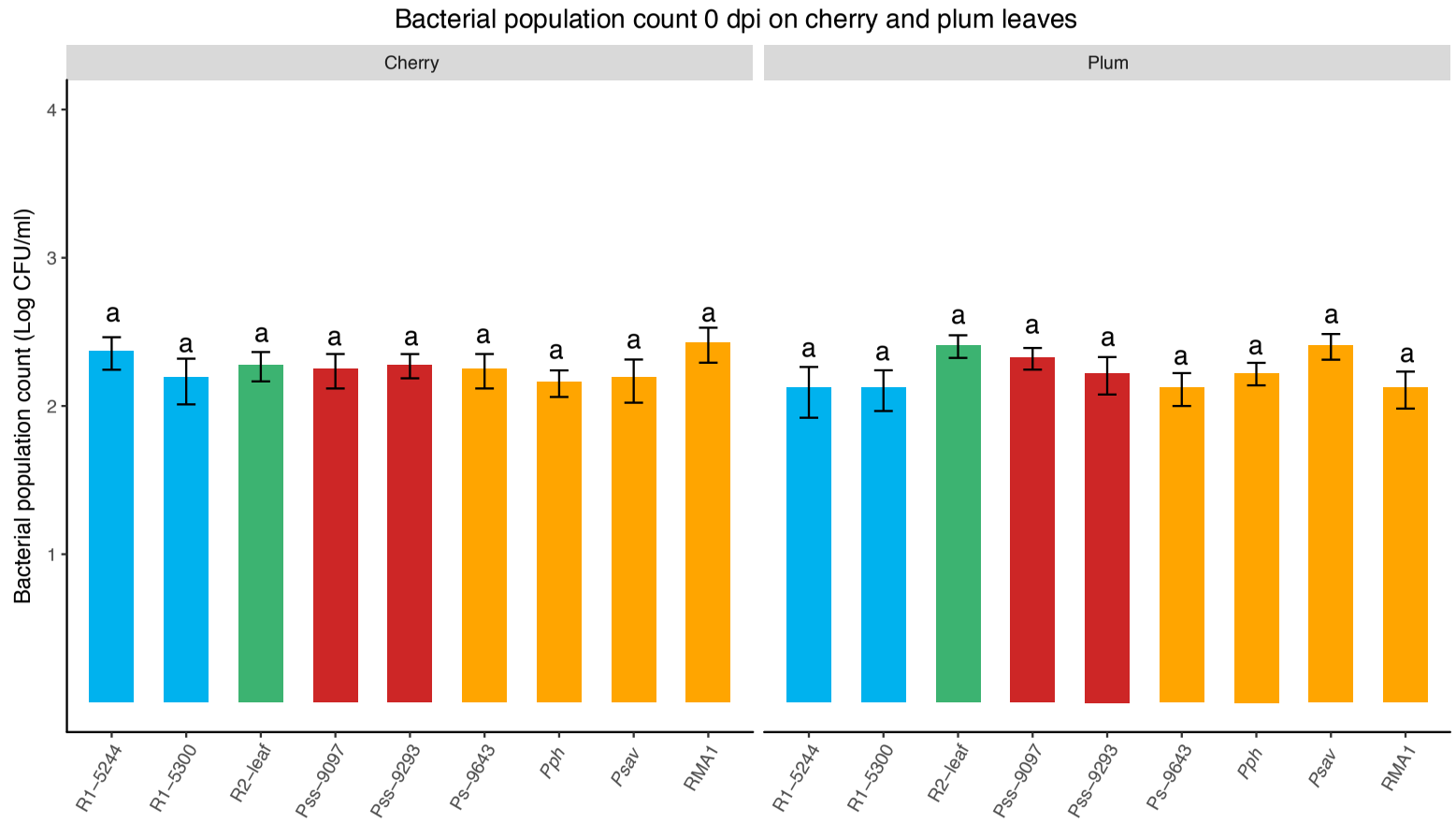


Figure S3.5: Day 0 population counts on cherry cv. Van and plum cv. Victoria (for Figure 3.12). *Strains are colour-coded with Psm R1: blue, Psm R2: green, Pss: red, non-pathogen RMA1: orange. Data presented are the mean values of three replicates, with error bars showing standard error above and below the mean. Tukey-HSD (p=0.05, confidence level: 0.95) significance groups for the different strains are presented.*

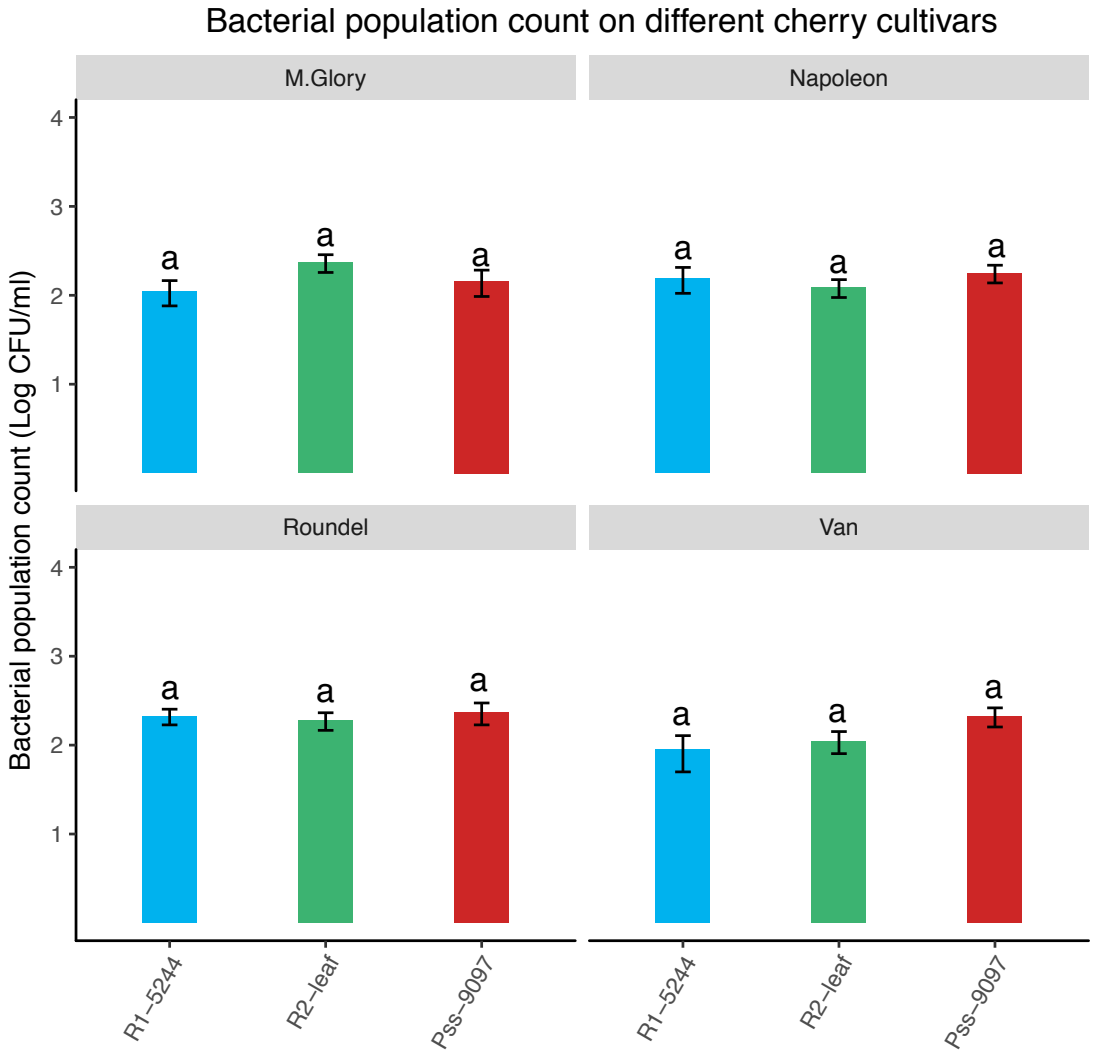


Figure S3.6: Day 0 population counts of three pathogenic* P. syringae *strains on different cherry cultivars (for Figure 3.14). *Strains are colour-coded with* Psm *R1: blue,* Psm *R2: green,* Pss*: red. The control is not included as no bacteria were found. The data presented are the mean values of three replicates, with error bars showing standard error above and below the mean. Tukey-HSD (p=0.05, confidence level: 0.95) significance groups for the different strains on each separate cultivar are presented.*

Chapter 4: Comparative genomics of *P. syringae* to explore the evolution of host specificity for cherry (*Prunus avium*)

4.1. Abstract

Pseudomonas syringae is a globally important plant pathogen that infects over 180 different plant species. Individual strains are highly specialised and can only infect one or a few hosts. This study utilised comparative genomics to explore host specificity of strains that cause disease on cherry (*Prunus avium*). At least five pathovars of *P. syringae* have been described to infect cherry and cause a range of symptoms such as canker and gall formation. These are *P.s* pv. *morsprunorum* race 1 (*Psm* R1), *P.s.* pv. *morsprunorum* race 2 (*Psm* R2, which is now known to be distantly related to race 1), *P.s.* pv. *syringae* (*Pss*), *P.s.* pv. *avii* and *P.s.* pv. *cerasicola*. A range of strains were sequenced from *Psm* R1, *Psm* R2 and *Pss* using both short and long-read sequencing technology (Illumina MiSeq and PacBio/MinIon). Genome-based phylogenetics revealed that individual clades were generally monophyletic, apart from *Pss*, where cherry-infecting strains were interspersed with strains isolated from other plant species. The clades that infect cherry were found to possess very different repertoires of type III effectors. *Pss* possessed a small repertoire of core effectors and several non-ribosomal toxin biosynthesis clusters, whilst the other clades had much larger effector repertoires and generally lacked toxins. There was convergent acquisition of several effector genes, such as *hopAR1* and *hopBB1* in multiple clades, which suggests they are important in host specificity for cherry. Convergent loss and pseudogenisation of members of the effector family *hopAB* suggested the product of this gene could be an avirulence factor, detected by the cherry immune system. The program BayesTraits identified further putative effectors, the presence/absence of which is significantly correlated with cherry pathogenicity. There was also convergent gain of gene clusters whose products are involved in the degradation of aromatic compounds across the cherry-infecting clades, suggesting that these genes are important for survival in this niche. Finally, analysis of virulence gene phylogenies showed that horizontal gene

transfer of effectors has occurred frequently between cherry-infecting clades, with many key effectors being present on plasmids. All strains of *Pss* (apart from 9644) lacked plasmids, indicating that plasmids may not be a major mechanism of genetic exchange for this clade.

4.2 Introduction

Over the last decade, high-throughput sequencing has become a major tool in the study of bacteria (Edwards & Holt 2013). With the reductions in the cost and improvements in technology and bioinformatics, genome sequencing is now common-place in research. This has led to huge leaps in our understanding of bacterial pathogenicity, ecology and evolution (Binnewies *et al.* 2006). With the increasing number of bacterial genomes available and the inclusion of non-model species, population-level studies can now be conducted, to ask complex evolutionary questions about bacterial populations, such as how disease epidemics emerge and what ecological processes drive the evolution of pathogenicity and host specialisation (Guttman *et al.* 2014; Monteil *et al.* 2016).

Before the genomics era, functional characterisation studies of *P. syringae* revealed the importance of type III secretion system effector (T3SE) repertoires, toxins and other virulence factors in pathogenesis (Lindgren 1997; Bender *et al.* 1999; Alfano *et al.* 2000). Type III effectors were found to not only be involved in disease, but also act as avirulence factors when detected by the plant immune system (Jones & Dangl 2006). The identification of effector genes through genomics has led to the discovery that *P. syringae* has evolved a large functionally redundant repertoire of effectors, which allows pathogen populations to easily lose/modify effectors that are detected, with minimal impact on overall virulence (Arnold & Jackson 2011). The selection pressures resulting from the detrimental effects of disease on host plants has triggered a similar expansion of plant receptor genes involved in pathogen effector detection (Karasov *et al.* 2014). It is believed that as pathogen lineages specialise towards a particular host, they will fine-tune their effector repertoires to maximize virulence and avoid detection. Host range becomes restricted during specialisation

because the pathogen may lose effectors important for disease on other hosts or gain effectors detected by other plant species (Schulze-Lefert & Panstruga 2011). Therefore, many comparative genomics studies to date have focused on identifying patterns that link effector complements with a particular diseases (O'Brien *et al.* 2012; Baltrus *et al.* 2012).

Early genomics studies focused on the three pathovars infecting the model systems of bean, tomato and *Arabidopsis* (Joardar *et al.* 2005; Buell *et al.* 2003; Feil *et al.* 2005). These studies revealed both substantial overlap in effector gene content and pathovar-specific effectors in the three *P. syringae* pathovars. They also indicated the important roles mobile elements play in genome plasticity and how distinct gene sets may facilitate adaptation towards particular pathogen lifestyles. For example, the expansion of gene sets involved in UV resistance, such as the *rulAB* operon, may be important for the epiphytic lifestyle of *P.s* pv. *syringae* on bean (Feil *et al.* 2005). The availability of these high quality reference genomes led to a wider study of host specificity using microarray-based comparative genomic hybridization, a technique that allows the identification of a strains gene content based on genomic DNA hybridization to a microarray of reference strain loci. Sarkar *et al.* (2006) used this technique with great effect to elucidate putative host specificity and virulence factors in a variety of *P. syringae* pathovars. They were able to statistically associate various genes with particular plant hosts such as tomato and *Brassica* species.

Advances in sequencing technology in the last decade have led to >100 additional *P. syringae* strains being genome sequenced (Martínez-García *et al.* 2015). This has enabled a range of comparative genomics studies of non-model plant pathogens, leading to great advances in our understanding of *P. syringae* pathogenicity, host specificity and population genetics. Baltrus *et al.* (2011), sequenced 19 diverse strains to explore broad genomic trends across the *P. syringae* complex. They found that the size of effector repertoires varied greatly between pathovars and that in strains with fewer effectors, phytotoxin genes may instead be providing important virulence functions. This indicated that convergent mechanisms of pathogenicity have evolved in *P. syringae*. This study showed that particular effector families show rapid

evolution, with genes frequently being gained, lost and mutated across the phylogeny. This is likely due to the selective pressures exerted by host immune detection of these effector proteins. By examining multiple homologues of the same effector in distantly related strains they were able to identify host-specific mutations that potentially have evolved to prevent recognition on new host plant species. Other studies have focused on characterising strains that cause particularly important diseases. Population genomics of *P.s* pv. *actinidiae*, the causal agent of kiwifruit canker in Asia and Australasia, revealed that three distinct clades of the pathogen have arisen during the domestication of kiwifruit. These clades have similar core genomes but vary in effector genes, indicating they may utilise different effectors to cause the same disease (McCann *et al.* 2013). By sampling from diverse populations in wild and cultivated kiwi fruit species in South-East Asia, they traced the origins of the lineage responsible for the recent global pandemic of kiwifruit canker to China (McCann *et al.* 2016).

Much of the understanding of the *P. syringae* – plant molecular interaction has been achieved using herbaceous plant models such as *Arabidopsis*, bean and tomato. Woody, perennial patho-systems provide a greater challenge due to the difficulty in conducting pathogenicity tests and seasonality of these interactions. Pathogens often occupy different host tissues at different times of the year and may use distinct tissue-specific strategies (Leben 1981; Lu *et al.* 2008; Lamichhane *et al.* 2014). Perennial hosts may also differ in susceptibility at different developmental stages. For example, temporal changes in hormones levels may influence immune responses as a plant ages (Crosse 1966; Garrett 1979; Kus *et al.* 2002).

Bioinformatics and functional analysis has revealed insights into *P. syringae* diseases on perennial crops. A study of the olive pathogen *P.s* pv. *savastanoi* revealed two effector genes (*hopBL1* and *hopBL2*) are over-represented in wood-infecting pathovars so may be important for virulence in this niche (Matas *et al.* 2014). Apart from effectors, genes involved in the metabolism of aromatic compounds, phytohormone production and tolerance to reactive oxygen species have been implicated in virulence on woody olive plants (Rodríguez-Palenzuela *et al.* 2010; Buonaurio *et al.* 2015). Following on from this, Bartoli *et al.* (2015), showed the

importance of one pathway involved in the metabolism of aromatic compounds in survival on woody hosts. They found that the degradation of catechol via the β-ketoadipate pathway was important for endophytic survival and symptom development of *P.s* pv. *actinidiae* on kiwifruit.

Green *et al.* (2010), completed a comparative study of the horse chestnut pathogen *P.s* pv. *aesculi*. They found that differences in the tissue specificity of strains could be related to the ability to metabolise sucrose. Wood-infecting strains possessed a sucrose metabolism gene cluster, which is potentially important in the colonisation of the phloem as it was missing from the leaf-infecting strain. With a greater number of genomes for comparison, Nowell *et al.* (2016) were able to identify genes significantly associated with the woody niche, and plot the gain and loss of genes onto the phylogeny leading to *P.s* pv. *aesculi*. They found candidate effectors, xylose degradation and the β-ketoadipate pathway were associated with horse chestnut pathogenicity and woody pathovars generally.

In this study, genomics was used to examine the evolution of strains that cause bacterial canker on sweet cherry (*Prunus avium*). Clades of *P. syringae* that constitute the main causal agents of bacterial canker include *P.s* pv. *morsprunorum* race 1, *P.s* pv. *morsprunorum* race 2 and a *P.s* pv. *syringae* (Bultreys & Kaluzna 2010). Two other pathovars have also recently been characterised as cherry pathogens, these are the wild cherry pathogen *P.s* pv. *avii* and gall-forming pathogen *P.s* pv. *cerasicola* (Ménard *et al.* 2003; Kamiunten *et al.* 2000). The cherry-infecting clades of *P. syringae* are reported to exhibit differences in virulence, host-range, geographical distribution and lifestyle (Scortichini 2010; Crosse & Garrett 1966). Due to this complexity, the *P. syringae*-cherry pathosystem provided an intriguing opportunity to study multiple occurrences of convergent specialisation at the genome level.

4.3 Materials and methods

4.3.1 Bacterial strains

The bacterial strains used in this study are presented in Table 4.1.

Strain	Pathovar	Race	Host/Isolation source	*Prunus* host cv.	Reference	BioProject/accession
R1-5244	*morsprunorum*	1	*Prunus avium*	unknown	This study	PRJNA345357
R1-5300	*morsprunorum*	1	*Prunus domestica*	Victoria	This study	PRJNA345357
R1-9326	*morsprunorum*	1	*Prunus domestica*	Victoria	This study	PRJNA345357
R1-9629	*morsprunorum*	1	*Prunus domestica*	Victoria	This study	PRJNA345357
R1-9646	*morsprunorum*	1	*Prunus avium*	Stella	This study	PRJNA345357
R1-9657	*morsprunorum*	1	*Prunus domestica*	Victoria	This study	PRJNA345357
R2-5255	*morsprunorum*	2	*Prunus avium*	Napoleon	This study	PRJNA345357
R2-5260	*morsprunorum*	2	*Prunus avium*	Roundel	This study	PRJNA345357
R2-leaf	*morsprunorum*	2	*Prunus avium*	Napoleon	This study	PRJNA345357
R2-sc214	*morsprunorum*	2	*Prunus avium*	Wild cherry	This study	PRJNA345357
R2-9095	*morsprunorum*	2	*Prunus avium*	Wild cherry	This study	PRJNA345357
syr9097	*syringae*		*Prunus avium*	unknown	This study	PRJNA345357
syr9293	*syringae*		*Prunus domestica*	Victoria	This study	PRJNA345357
syr9630	*syringae*		*Prunus domestica*	Victoria	This study	PRJNA345357
syr9644	*syringae*		*Prunus avium*	Stella	This study	PRJNA345357
syr9654	*syringae*		*Prunus domestica*	Victoria	This study	PRJNA345357
syr9656	*syringae*		*Prunus avium*	Kiku-Shidare	This study	PRJNA345357
syr9659	*syringae*		*Prunus avium*	Kiku-Shidare	This study	PRJNA345357
Ps-9643	unknown		*Prunus domestica*	Victoria	This study	PRJNA345357
RMA1	unknown		*Aquilegia vulgaris*	Winky	This study	PRJNA345357
syr2675	*syringae*		*Phaseolus vulgaris*		This study	PRJNA345357
syr2676	*syringae*		*Phaseolus vulgaris*		This study	PRJNA345357
syr2682	*syringae*		*Phaseolus vulgaris*		This study	PRJNA345357
syr3023	*syringae*		*Syringa vulgaris*		This study	PRJNA345357
syr100	*syringae*		*Phaseolus lunatus*		This study	PRJNA345357
acer302273	*aceris*		*Acer* sp.		Baltrus *et al.* 2012	AEAO00000000
act302091	*actinidiae*		*Actinidia deliciosa*		Baltrus *et al.* 2012	AEAL00000000
actCRAFRU	*actinidiae*		*Actinidia deliciosa*		Butler *et al.* 2013	ANGD00000000

Strain	Pathovar	Race	Host/Isolation source	*Prunus* host cv.	Reference	BioProject/accession
actKW41	*actinidiae*		*Actinidia deliciosa*		McCann *et al.* 2013	AGNP00000000
actNCPPB3871	*actinidiae*		*Actinidia deliciosa*		Marceletti *et al.* 2011	ANGD00000000
aes089323	*aesculi*		*Aesculus hippocastanum*		Baltrus *et al.* 2012	AEAD00000000
aes2250	*aesculi*		*Aesculus hippocastanum*		Green *et al.* 2010	ACXT00000000
aes3681	*aesculi*		*Aesculus hippocastanum*		Green *et al.* 2010	ACXS00000000
amy3205	*amygdali*		*Prunus dulcis*		Bartoli *et al.* 2015	JYHB00000000
amyICMP3918	*amygdali*		*Prunus dulcis*		Thakur *et al.* 2016	LJPQ00000000
atroDSM50255	*atrofaciens*		*Triticum aestivum*		Baltrus *et al.* 2014	AWUI00000000
avelBP631	*avellanae*		*Corylus avellana*		O'Brien *et al.* 2012	AKBS00000000
avelVe013	*avellanae*		*Corylus avellana*		O'Brien *et al.* 2012	AKCK00000000
avelVe037	*avellanae*		*Corylus avellana*		O'Brien *et al.* 2012	AKCJ00000000
avii3846	*avii*		*Prunus avium*		Nowell *et al.* 2016	LIIJ00000000
BRIP34876	unknown		*Hordeum vulgar*		Gardiner *et al.* 2013	AMXK00000000
BRIP34881	unknown		*Hordeum vulgar*		Gardiner *et al.* 2013	AMXL00000000
BRIP39023	unknown		*Hordeum vulgar*		Gardiner *et al.* 2013	AMZX00000000
castCFBP4217	*castaneae*		*Castanea crenata*		Nowell *et al.* 2016	LIIH00000000
CC1416	unknown		Epilithon		Baltrus *et al.* 2014	AVEP00000000
CC1458	unknown		*Dodecatheon pulchellum*		Baltrus *et al.* 2014	AVEN00000000
CC1543	unknown		Lake water		Baltrus *et al.* 2014	AVEJ00000000
CC1544	unknown		Lake water		Baltrus *et al.* 2014	AVEI00000000
CC1557			Snow		Hockett *et al.* 2014	AVEH00000000
CC1559	unknown		Snow		Baltrus *et al.* 2014	AVEG00000000
CC440	unknown		*Cantaloupe*		Baltrus *et al.* 2014	AVEC00000000
CC457	unknown		*Cantaloupe*		Baltrus *et al.* 2014	AVEB00000000
CC94	unknown		*Cantaloupe*		Baltrus *et al.* 2014	AVEA00000000
cera6109	*cerasicola*		*Prunus yedoensis*		Nowell *et al.* 2016	LIIG00000000
ceraICMP17524	*cerasicola*		*Prunus yedoensis*		Thakur *et al.* 2016	LJQA00000000
ciccICMP5710	*ciccaronei*		*Ceratonia siliqua*		Thakur *et al.* 2016	LJPY00000000

Strain	Pathovar	Race	Host/Isolation source	*Prunus* host cv.	Reference	BioProject/accession
cunnICMP11894	*cunninghamiae*		*Cunninghamia lanceolata*		Thakur *et al.* 2016	LJQE00000000
daphICMP9757	*daphniphylli*		*Daphniphyllum teijsmannii*		Thakur *et al.* 2016	LJQF00000000
glyR4	*glycinea*		*Glycine max*		Qi *et al.* 2011	AEGH00000000
ICMP11293	unknown		*Actinidia deliciosa*		Visnovsky *et al.* 2016	LKEP00000000
ICMP19498	unknown		*Actinidia deliciosa*		Visnovsky *et al.* 2016	LKCH00000000
lach301315	*lachrymans*		*Cucumis sativus*		Baltrus *et al.* 2012	AEAF00000000
lachrymans302278	*lachrymans*		*Cucumis sativus*		Baltrus *et al.* 2012	AEAM00000000
lapsaICMP3947	*lapsa*		*Zea* sp.		Thakur *et al.* 2016	LJQQ00000000
macuES4326	maculicola		*Raphanus sativus*		Baltrus *et al.* 2012	AEAK00000000
mors302280	*morsprunorum*		*Prunus domestica*		Baltrus *et al.* 2012	AEAE00000000
morsU7805	*morsprunorum*		*Prunus mume*		Mott *et al.* 2016	LGLQ00000000
myriICMP7118	*myricae*		*Myrica rubra*		Thakur *et al.* 2016	LJQV00000000
neriiICMP16943	*savastanoi*		*Olea europea*		Thakur *et al.* 2016	LJQW00000000
paniLMG2367	*panici*		*Panicum miliaceum*		Liu *et al.* 2012	ALAC00000000
papu1754	*papulans*		*Malus sylvestris*		Nowell *et al.* 2016	JYHI00000000
persNCPPB2254	*persicae*		*Prunus persica*		Zhao *et al.* 2015	LAZV00000000
photICMP7840	*photiniae*		*Photinia glabra*		Thakur *et al.* 2016	LJQO00000000
pisiPP1	*pisi*		*Pisum sativum*		Baltrus *et al.* 2014b	AUZR00000000
pph1448a	*phaseolicola*		*Phaseolus vulgaris*		Joardar *et al.* 2005	CP000058
R1-2341	*morsprunorum*	1	*Prunus cerasus*	unknown	Nowell *et al.* 2016	LIIB00000000
R1-5269	*morsprunorum*	1	*Prunus cerasus*	unknown	Nowell *et al.* 2016	LIHZ00000000
R2-5261	*morsprunorum*	2	*Prunus avium*	Roundel	Nowell *et al.* 2016	LIIA00000000
rhapCFBP4220	*rhaphiolepidis*		*Rhaphiolepis umbellata*		Nowell *et al.* 2016	LIHV00000000
sava3335	*savastanoi*		*Olea europea*		Rodriguez-Palenzuela *et al.* 2010	ADMI00000000
sava4352	*savastanoi*		*Olea europea*		Thakur *et al.* 2016	LGKR00000000
savaDAPP-PG722	*savastanoi*		*Olea europea*		Moretti *et al.* 2014	JOJV00000000
savaPseNe107	*savastanoi*		*Olea europea*		Bartoli *et al.* 2015	JYHF00000000

Strain	Pathovar	Race	Host/Isolation source	*Prunus* host cv.	Reference	BioProject/accession
soliICMP16925	*solidagae*		*Solidago altissima*		Thakur *et al.* 2016	JYHF00000000
syr1212	*syringae*		*Pisum sativum*		Baltrus *et al.* 2014	AVCR00000000
syr2339	*syringae*		*Prunus avium*		Nowell *et al.* 2016	LIHU00000000
syr2340	*syringae*		*Pyrus* sp.		Nowell *et al.* 2016	LIHT00000000
syr41a	*syringae*		*Prunus armeniaca*		Bartoli *et al.* 2015	JYHJ00000000
syr7872	*syringae*		*Prunus domestica*	Opal	Nowell *et al.* 2016	LIHS00000000
syr7924	*syringae*		*Prunus cerasus*		Nowell *et al.* 2016	LIHR00000000
syrB301D	*syringae*		*Pyrus communis*		Ravindran *et al.* 2015	CP005969
syrB64	*syringae*		*Triticum aestivum*		Dudnik and Dudler 2013	ANZF00000000
syrB728a	*syringae*		*Phaseolus vulgaris*		Feil *et al.* 2005	CP000075
syrCRAFRU12	*syringae*		*Corylus avellana*		Scortichini *et al.* 2013	ATSV00000000
syrHS191	*syringae*		*Panicum miliaceum*		Ravindran *et al.* 2015	CP006256
syrSM	*syringae*		*Triticum aestivum*		Dudnik and Dudler 2013	APWT00000000
syrUMAF0158	*syringae*		*Mangifera indica*		Martínez-García *et al.* 2015	CP005970
thea3923	*theae*		*Camelia sinensis*		Mazzaglia *et al.* 2012	AGNN00000000
tomDC3000	*tomato*		*Solanum lycopersicum*		Buell *et al.* 2003	AE016853
tomT1	*tomato*		*Solanum lycopersicum*		Almeida *et al.* 2009	ABSM00000000
UB303	unknown		Lake water		Baltrus *et al.* 2014	AVDZ00000000
ulmiICMP3962	*ulmi*		*Ulmus* sp.		Thakur *et al.* 2016	LJRQ00000000
USA007	unknown		Stream water		Baltrus *et al.* 2014	AVDY00000000
USA011	unknown		Stream water		Baltrus *et al.* 2014	AVDX00000000

Table 4.1: List of bacterial strains used in this study, *with pathovar designation, host of isolation and reference if sequenced elsewhere.*

4.3.2 Plasmid profiling

To determine the number and sizes of native plasmids present in *Prunus*-infecting strains, plasmids were extracted and viewed by gel electrophoresis (Moulton *et al.* 1993; Neale *et al.* 2013). Briefly, 400 µl of an overnight *P. syringae* culture or 200 µl of an *E. coli* culture was spun down at 12,000g for 1 min and the supernatant removed. 40µl of TE was added without disturbing the pellet, followed by 30 µl of a 0.2% SDS, 0.2M NaOH solution and mixed gently with the pipette tip. The cells were left for 5 min to lyse. Plasmids were then extracted by 70 µl of phenol:chloroform:isoamyl in a ratio of 25:24:1 (Sigma-Aldrich), the tube was flicked to mix and spun in the centrifuge for 10 min at 12,000g. The top phase was then run on a 0.6% gel at 5V/cm for at least five hours. The gel was post-stained with GelRed (Biotium) using manufacturer's instructions.

4.3.3 Genome sequencing, assembly and annotation

4.3.3.1 Illumina

To prepare genomic DNA sequencing libraries, genomic DNA was extracted from 25 *P. syringae* strains using the Puregene Yeast/Bact Kit (Qiagen, UK). The DNA was fragmented using a sonicating water bath for 30 seconds. Fragmented DNA ranging from 400-700bp was then extracted from an agarose gel using the Zymogen gel extraction kit (Zymo Research). To make DNA libraries, the NextFlex Rapid-DNA sequencing kit was then used, excluding the size selection step. The barcodes were multiplexed to allow pooling of multiple samples in one run. Libraries were quality checked using the NGS kit on the Fragment Analyzer (Advanced Analytical, USA) and concentration measured using the HS-DNA Qubit kit (Life Technologies, UK).

Libraries were sequenced using the Illumina Mi-Seq V3 (Illumina, USA) 300bp paired-end reads. A preliminary run was performed on the pooled DNA of four strains. One of these strains was then re-sequenced along with 15 others (16 in the final run).

Raw data for each genome was quality checked and trimmed to remove adapter sequences using fastqc-mcf (https://code.google.com/p/ea-utils/wiki/FastqMcf). The

reads were then error corrected using Quake prior to assembly (Kelley *et al.* 2010). Each genome was assembled using SPAdes 3.1.0 (Bankevich *et al.* 2012) and summary statistics generated using Quast (Gurevich *et al.* 2013).

4.3.3.2 Pacbio and MinIon sequencing

To produce more complete bacterial genomes, four representative strains from different clades were sequenced using the long-read technologies PacBio (Pacific Biosystems) and MinIon (Oxford Nanopore). High molecular weight DNA was extracted from four *P. syringae* strains (R1-5244, R2-leaf and syr9097 for PacBio and R1-5300 for MinIon) using a CTAB method (Feil *et al.* 2012). For the PacBio sequencing, this DNA was sent TGAC (Norwich) for PacBio RSII (C4-P6 chemistry) sequencing with one SMRTcell per strain. The MinIon sequencing was performed in-house. The DNA library was prepared using Oxford Nanopore instructions (RAD001 rapid-prep kit) with the optional fragmentation step using the covaris G-tube (Covaris Inc) to generate an approximate library size of 8kb. The libraries were then run on the Oxford Nanopore, and the raw data sent to Metrichor for 1D basecalling.

For the MinIon data, the reads were extracted from Fast5 files using the program Poretools (Loman & Quinlan 2014). The sequencing reads for both technologies were then trimmed and assembled using the program Canu with default options (Berlin *et al.* 2015). The program Circlator was used to circularise the contigs, removing the duplicated regions that occur at the end of contigs during assembly of long-read sequence data (Hunt *et al.* 2015). Circlator successfully circularised nearly all contigs. One contig each from the R1-5244, R1-5300 and R2-leaf assemblies failed to circularise so these were manually trimmed to remove duplicated regions. The final circularised contigs were then polished using the error-corrected Illumina reads with bowtie2, samtools and Pilon 1.17 (Langmead & Salzberg 2013; Li *et al.* 2009; Walker *et al.* 2014).

All assemblies were annotated using the online annotation server RAST (Aziz *et al.* 2008) which annotates protein-encoding, tRNA and rRNA genes as well as assigning

genes into metabolic subsystems. The genomes were then submitted to Genbank under the BioProject ID: PRJNA345357.

4.3.4 Phylogenetic analysis of *P. syringae*

Three different methods were used to produce phylogenies of *P. syringae*. The first involved using seven Multi-Locus Sequence Typing genes (*acnB, fruK, gapA, gltA, gyrB, pgi and rpoD*) which are commonly used in phylogenetic studies of bacteria (Baltrus *et al.* 2011). These gene sequences were extracted from the genome assemblies of all strains aligned using clustalW (Larkin *et al.* 2007) and concatenated. This alignment was then trimmed of poorly aligned positions using Gblocks 0.19 (Castresana 2000). A phylogeny was built using the MrBayes plugin (Huelsenbeck & Ronquist 2001) on Geneious 7.1.9 (Kearse *et al.* 2012) with a GTR gamma model of substitution, a chain-length of 1,100,000, burn-in length of 100,000 and sub-sampling frequency of 200. The strain ES4326 (initially identified as *P.s*. pv *maculicola*) was used as an outgroup as it has been shown to actually be a closer relative of *Pseudomonas cannabina* (Bull *et al.* 2010).

To create a core genome phylogeny, individual single-copy genes identified as being present in all genomes through orthology analysis (see below) were extracted from each genome and separately aligned using ClustalW. The gene alignments were concatenated using Geneious and Gblocks was used again to remove poorly aligned positions. A Maximum-likelihood approach was taken to build the phylogeny as with the increased number of genes a Bayesian approach would be too computationally intensive (Moriarty Lemmon & Lemmon 2013). The program jmodeltest 2.1.10 (Posada 2008) was used to determine the correct evolutionary model. RaxML (AVX) 8.1.15 (Stamatakis 2014) was then used to build the phylogeny with a GTR gamma model of evolution and 100 non-parametric bootstrap replicates.

The final method of phylogeny construction used the program RealPhy 112 (Bertels *et al.* 2014). RealPhy 112 maps sequencing reads (or assembly contigs cut into 50bp fragments) to a set of reference genomes to build whole genome alignments for the construction of phylogenetic trees. Three different reference genomes were used of strains that had complete genome assemblies (*P.s* pv. *tomato* DC3000, *P.s* pv.

phaseolicola 1448A and *P.s* pv. *syringae* B728a) to build alignments which were then merged into a final alignment. A RaxML phylogeny was then built using this merged alignment with the GTR gamma model of evolution and 100 non-parametric bootstrap replicates.

4.3.5 Whole-genome alignment

To align whole genomes the program Mauve was used (Darling *et al.* 2010). FASTA sequences were given as input and the progressiveMauve algorithm with default settings was run to create an alignment.

4.3.6 Effector identification

To search for homologs of known *P. syringae* effectors, all available effector protein sequences were downloaded from *pseudomonas-syringae*.org. This set included all defined effectors from all sequenced *P. syringae* isolates that were available. Although this is relatively up-to-date, a recent study (Lo *et al.* 2016) reclassified members of the *hopF* family, so this classification system was used in the analysis.

The program tBLASTn, within the Basic Local Alignment Search Tool (BLAST) suite (Altschul *et al.* 1990), was used to search genome sequences for the presence of effector gene homologues. Effector homologues were concluded to be present with a BLAST score of ≥70% identity and ≥40% query length to at least one sequence in each effector family. Python (Python Software Foundation) was then used to extract the nucleotide sequences from the genome sequence. These sequences were manually examined for frameshifts, truncations and the presence of a *hrp*-box motif within 500bp of the start of the gene. A heatmap showing presence/absence of effector homologs was then generated using the function heatmap.2 in the R library gplots 3.0.1. (R Core Team 2012; Warnes *et al.* 2016)

A similar analysis was performed for other potential virulence genes, encoding proteins involved in phytotoxin biosynthesis and wood-degradation. Protein sequences were obtained from NCBI and then blasted against the genome sequence

using tblastn. A homologue was said to be putatively present for BLAST scores of ≥70% identity and ≥40% query length.

4.3.7 Gain and loss analysis

To plot the gain and loss of genes on the phylogenetic tree, the online webserver GLOOME (Gain Loss Mapping Engine) was used (Cohen *et al.* 2010). First a presence and absence matrix of effector genes for all strains was converted into FASTA format by using Python. GLOOME was then run using the core genome or RealPhy phylogeny with a variable gain/loss ratio evolutionary model, an optimisation level of "very high", and four rate categories, to improve accuracy.

4.3.8 BayesTraits

To correlate the association of particular genes with cherry pathogenicity (Pagel 2004; Barker & Pagel 2005), the program BayesTraits was used in discrete mode. BayesTraits took a binary matrix of presence and absence of two traits and a phylogenetic tree, and modelled if the two traits evolved independently or in a mutually dependent manner. A binary matrix was created for each effector (presence/absence in different strains) and pathogenic ability of each strain on cherry. BayesTraits was then run in discrete mode for both the independent model of evolution (where the two traits are not correlated) and dependent mode (where the evolution of the two traits is correlated). The likelihood score for each model was then generated, from which a ratio (2*(Lh(D)-Lh(I))) is calculated for each gene. A Chi-squared test was then used to generate a p value of significance.

The analysis was ran 100 times using both the core genome and RealPhy phylogenies. Only those genes that were statistically associated with pathogenicity ($p \leq 0.05$) in >90% of the runs, as in Press *et al.* (2013) were labelled as significant.

4.3.9 Phylogenetic analysis of individual effector genes

For each effector family, the nucleotide sequences of best hits were extracted from the genome and aligned using clustalW (Larkin *et al.* 2007). The program FastTree was used to build phylogenies from each alignment (Price *et al.* 2010). To visualise the

phylogenies, the program ColorTree (Chen *et al.* 2009) was used to automatically change the font colour of strains of interest and the phylogenies were viewed using Dendroscope (Huson *et al.* 2007).

The phylogenetic trees were used to infer if effectors gained in cherry-infecting strains had been horizontally transferred, as described in O'Brien *et al.* (2012). Each tree was manually assessed to see if effector gene sequences for cherry pathogens clustered with those sequences of their close relatives within the core genome phylogeny. Incongruence between the core genome phylogeny and effector phylogeny suggested that a horizontal transfer had occurred.

4.3.10 Effector genome region alignments

The regions containing individual effectors of interest were extracted from genome assemblies and aligned using MAFFT (Katoh *et al.* 2002). They were visualised and figures created using Geneious 7.1.9 (Kearse *et al.* 2012).

4.3.11 Effector domain identification

Candidate effector proteins that were thought to be important for host specificity, were analysed for particular domains. The online version of Interproscan (Quevillon *et al.* 2005) was used to identify domains and the online program Illustrator for Biological Sequences (IBS) was used to create figures (Liu *et al.* 2015).

4.3.12 Search for novel effector genes

To determine if sequenced strains contained any novel previously unidentified effectors, the program Effective (Jehl *et al.* 2011) was used. Given an input protein FASTA file, Effective predicted candidate effectors based on the presence of eukaryotic-domains and Sec pathogen signal peptides in amino acid sequences. The nucleotide regions (+/- 500bp) containing each candidate gene that scored $\geq$0.95 were then extracted from the genome. A python regular-expression script to search for the *hrp*-box motif (allowing for sequence variation) was then used to narrow down the candidate list to only those genes with a nearby *hrp*-box motif.

4.3.13 Identification of genomic islands

To identify putative genomic islands in the PacBio sequenced pathogens, IslandViewer3 was used (Dhillon *et al.* 2015). IslandViewer3 incorporates three different methods to predict genomic islands based on identifying unique sequences in comparison to a closely related reference, changes in dinucleotide and codon-usage bias and the proximity of mobile element genes. As the identification of islands in fragmented genome assemblies may not produce accurate results, this analysis was only performed on the PacBio genome assemblies. The predicted islands were then manually delimitated using Geneious, based on the presence of key elements such as type three effectors, insertion sequences (Siguier *et al.* 2006), phages (Zhou *et al.* 2011), plasmid-encoding genes, lineage-specific genes identified through OrthoMCL, hypothetical proteins, the presence of nearby tRNA genes and changes in GC content (McCann *et al.* 2013).

A BLAST approach was utilised to determine if these genomic islands were present in other strains within the *P. syringae* complex. As the genomic islands were quite large (most > 10kb), they were split into 5000bp chunks. This was because most of the genomes being queried were in several hundred contigs, meaning that a BLAST of the original sequences could lead to false negatives if the island was spread across multiple contigs. BLASTn was used and an island was taken to be fully present if all 5000bp chunks produced hits with a query length >30%. As a control to check that this approach was detecting the genomic islands correctly, the Illumina-sequenced SPAdes genome assemblies of the PacBio-sequenced strains were searched for their own islands. All islands were predicted to be present, however, some were not complete which is likely to be due to contig breaks in these fragmented genomes. An island was therefore taken to be partially present if not all of the 5000bp regions were found in a genome.

4.3.14 Orthology analysis

The program OrthoMCL (Li *et al.* 2003) was used for orthology analysis. All sequenced genomes and 79 other *P. syringae* genomes were submitted to the RAST annotation server to obtain amino-acid FASTA files. To ensure the annotation was similar for all

sequenced strains, the Illumina Miseq SPAdes assemblies of the four PacBio/MinIon reference strains were used instead of the PacBio/MinIon assemblies. OrthoMCL took the gene set of each strain and uses a BLAST approach followed by Markov clustering (MCL) to cluster genes into potential orthogroups of similar function. A cut-off length of 50 residues was used to prevent the identification of orthogroups made up entirely of spurious protein fragments produced when contig breaks occur within a gene where two putative coding sequences are predicted. To further reduce the impact of protein fragments, all but one genome assembly had less than 1000 contigs (*P.s* pv. *aceris* 302273 has 1179 contigs, but was included as it is a woody-host-infecting pathogen in phylogroup 2, closely related to *Pss*).

Single-copy core genes were obtained from the orthology analysis for use in phylogenetic analysis of *P. syringae*. This was achieved by first removing multi-copy orthogroups (where an orthogroup is present more than once in a strains genome) from the total list of orthogroups. The remaining orthogroups were examined to determine how many strains possessed each orthogroup. Those orthogroups that were present in all strains were retained and the corresponding gene sequence in each strain was extracted from the gene nucleotide FASTA files generated by RAST.

4.3.15 Visualisation of genomes using Circos

To visualise the location of candidate pathogenicity regions in the PacBio strain genomes, the Circular genome plotter Circos was used (Krzywinski 2009). The locations of particular effector, toxin, plasmid-encoding and lineage-specific genes in the PacBio strains were confirmed using tBLASTn. To identify insertion elements and prophages the genomes were uploaded to the online webservers ISFinder (Siguier *et al.* 2006) and PHAST (Zhou *et al.* 2011).

4.4 Results

In this study, the genomes of a diverse set of *P. syringae* strains isolated from cherry and plum were sequenced. To aid comparisons, a closely related out-group isolated from *Aquilegia vulgaris* (RMA1) which is related to *Psm* R2 was sequenced. Strains of

P.s pv. *syringae* isolated from bean which belong to the same phylogroup as *Prunus*-infecting *Pss* were also genome sequenced.

Different phylogenetic methods were used to accurately infer the relationships between *Prunus*-infecting strains. Virulence factors were then identified and compared to the rest of the *P. syringae* species complex. An accurate phylogeny allowed the prediction of how these virulence factors have evolved during the specialisation of different *P. syringae* clades towards *Prunus*. The importance of horizontal transfer in pathogenicity for *Prunus* was also assessed.

This chapter describes a large comparative analysis of *P. syringae* strains. To simplify and avoid confusion over the naming of individual strains in figures, the first few letters of the pathovar name were used, followed by the strain identity number. For example, "Pss" becomes "syr", as otherwise *Pss* could refer to multiple pathovars beginning with "s", e.g. *syringae* and *savastanoi*. Full pathovar names and shortened versions are listed in Table 4.1.

4.4.1 Genome statistics and plasmid profiles

The genomes of 25 *P. syringae* strains were sequenced using the Illumina Miseq (strain details can be found in Table 4.1). Eighteen of these were isolated from either cherry or plum, and a further seven out-group strains were sequenced to aid comparisons. Three strains with well characterised virulence (R1-5244, R2-leaf and syr9097) representative of the three major cherry-infecting clades, were also sequenced using PacBio long-read sequencing to obtain complete genomes. Within the *Psm* R1 clade, which contained both pathogenic and non-pathogenic isolates, R1-5300 (non-pathogenic on cherry) was sequenced using the Oxford Nanopore Minion to produce a more complete genome and aid comparisons. Table 4.2 summarises the genome sequencing statistics for all sequenced strains. To determine plasmid content, plasmids were extracted and profiled using gel electrophoresis (Figure 4.1). The numbers of plasmids identified are also presented in Table 4.2. Strains of *Psm* R1 and R2 possessed between 1-5 plasmids, whereas only one *Prunus*-infecting *Pss* strain (syr9644) possessed a plasmid.

Assembly	Sequencing	no. contigs	predicted plasmids	Total length	GC%	N50	Average coverage	Features
Ps-9643	Illumina	58	1	5937102	58.78	243355	212	5386
R1-5244	Illumina	198	4	6302385	58.08	227422	265	5810
R1-5300	Illumina	201	4	6342586	57.88	142021	180	5844
R1-9326	Illumina	268	4	6353636	57.91	142021	81	5874
R1-9629	Illumina	216	3	6341664	57.94	142021	172	5856
R1-9646	Illumina	171	3	6302776	58.03	235429	180	5801
R1-9657	Illumina	191	4	6317852	57.91	145272	158	5848
R2-5255	Illumina	206	2	6448834	58.38	102760	112	5966
R2-5260	Illumina	223	3	6495620	58.41	101794	458	5995
R2-9095	Illumina	304	1	6418849	58.48	92453	100	5887
R2-leaf	Illumina	203	5	6366714	58.48	100658	180	5846
R2-sc214	Illumina	203	3	6253818	58.56	108341	180	5747
syr9097	Illumina	66	0	5892389	59.35	316078	158	5117
syr9293	Illumina	73	0	6135031	58.84	557853	196	5302
syr9630	Illumina	57	0	5940819	59.33	347701	206	5175
syr9644	Illumina	75	1	6173193	59.13	251053	208	5334
syr9654	Illumina	49	0	5941610	59.37	245023	147	5148
syr9656	Illumina	39	0	5980728	59.1	1007808	205	5184
syr9659	Illumina	51	0	5943090	59.37	235830	116	5148
RMA1	Illumina	95	1	6306889	58.73	187448	320	5825
syr100	Illumina	23	0	5872916	59.33	893822	83	5140
syr2675	Illumina	65	0	5994384	59.34	227612	83	5177
syr2676	Illumina	90	1	6158476	59.3	259660	78	5387
syr2682	Illumina	185	2	6259099	59.21	242212	84	5405
syr3023	Illumina	228	0	6203212	58.9	456738	88	5365
R1-5244	PacBio	5	4	6445963	58.05	6109228	82	6024

Assembly	Sequencing	no. contigs	predicted plasmids	Total length	GC%	N50	Average coverage	Features
R2-leaf	PacBio	6	5	6576340	58.41	6242845	141	6093
syr9097	PacBio	1	0	5929959	59.3	5929959	100	5147
R1-5300	MinION	6	4	6645601	57.87	5688034	100	6449

Table 4.2: Assembly statistics for all sequenced strains. *N50: The weighted median contig size in the assembly, it is the minimum contig length required to cover 50% of the genome. Features: The number of protein encoding and RNA sequences in the annotated genome.*

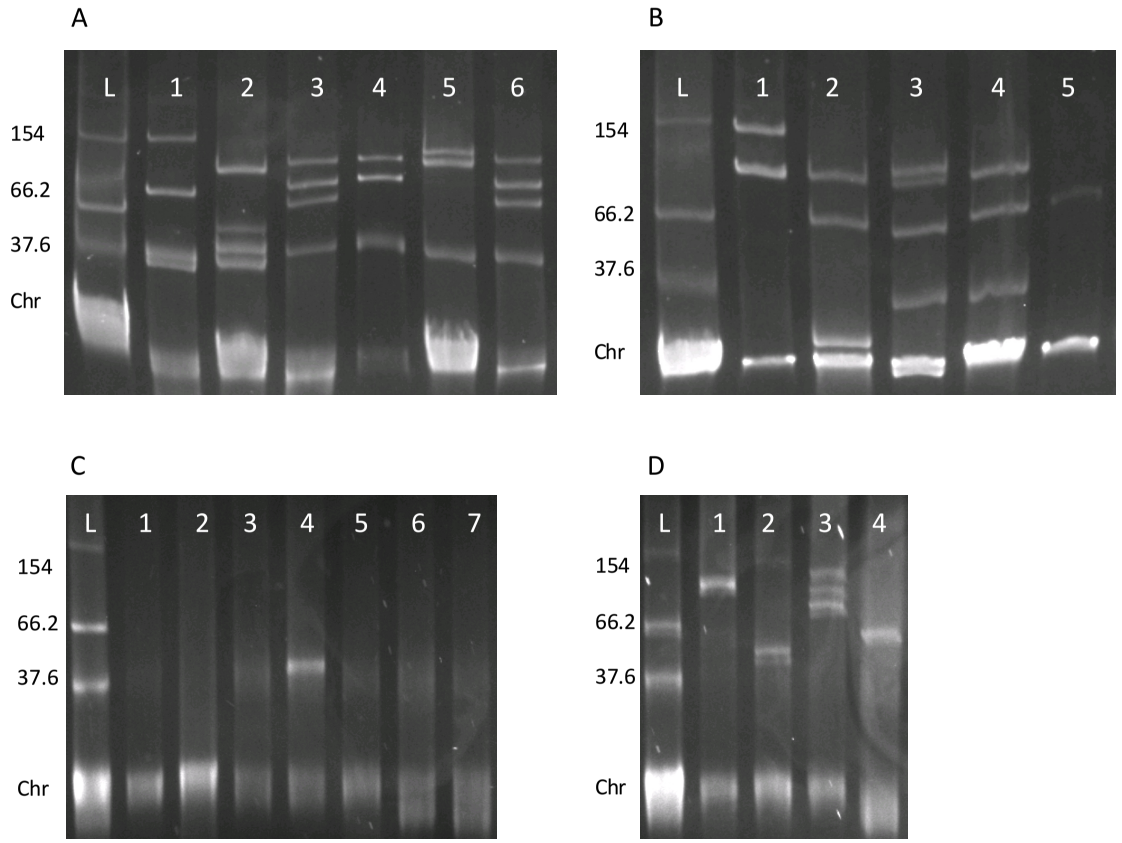


Figure 4.1: Plasmid profiles of all* Prunus*-infecting sequenced isolates and some out-groups for comparison. *The ladder was E. coli 39R861 marker strain which contains four plasmids of known sizes (three are shown: 37.6kb, 66.2kb and 154kb). A:* Psm *R1 (1: R1-5244, 2: R1-5300, 3: 9326, 4: 9629, 5: 9646, 6: 9657). B:* Psm *R2 and Ps 9643 (1: R2-5255, 2: R2-5260, 3: R2-leaf, 4: R2-sc214, 5: Ps-9643). C:* Pss *(1: syr9097, 2: syr9293, 3: syr9630, 4: 9644, 5: 9654, 6: 9656, 7: 9659). D: Out-groups and an additional* Psm *R2 strain sequenced later (1: R2-9095, 2: phas1448A, 3: avelBPIC631, 4: RMA1).*

The Illumina genomes were assembled using SPAdes (Bankevich *et al.* 2012). Most genomes assembled into ~140 contigs (mean = 142). The PacBio and Minion genomes were assembled using Canu (Koren *et al.* 2016). All PacBio-sequenced strains were assembled into complete chromosomal and plasmid contigs, with plasmid sizes corresponding to those sizes predicted in Figure 4.1. The putative plasmid contigs in the MinIon assembly were all the correct size, however the chromosome was split into two contigs (5.7Mb and 0.69Mb). To determine if the 0.69Mb contig was likely to be chromosomal, a whole genome alignment of the R1-5300 MinIon assembly with the R1-5244 PacBio assembly was performed using ProgressiveMauve (Darling *et al.* 2010). The alignment showed that this contig was homologous to a chromosomal region in R1-5244 (Figure S4.1).

4.4.2 Phylogenetics

For all further analyses, 79 additional genome sequences of strains were obtained from Genbank. These were mainly strains from *P. syringae* phylogroups 1-3. They included *Prunus*-infecting strains that had already been sequenced and various out-groups. Plant-pathogenic and environmental strains were included, as environmental strains often cluster phylogenetically with plant pathogens (Monteil *et al.* 2016).

4.4.2.1 MLST phylogeny

To determine the phylogenetic relationships between *Prunus*-infecting strains, first a Bayesian phylogenetic tree of 104 strains was constructed based on a concatenated 5650bp alignment of seven Multi-Locus Sequence Typing genes (*acnB*, *fruK*, *gapA*, *gltA*, *gyrB*, *pgi*, *rpoD*) (Figure 4.2). The alignment was then trimmed of any gaps. As these gene sequences were incomplete in particular strains, trimming led to a final alignment which only contained five genes (*acnB*, *fruK*, *gltA*, *gyrB*, *pgi).* The phylogeny placed the three *Prunus*-infecting clades into the three major *P. syringae* phylogroups (1, 2 and 3). *Psm* R1 and R2 formed monophyletic, highly homogenous clades within phylogroups 3 and 1 respectively. Both clades were closely related to other tree-infecting pathovars. *Psm* R1 was most closely related to a strain that has been described as *P.s* pv. *morsprunorum* but was isolated from apricot (*Prunus armeniaca*). Whereas, *Psm* R2 showed close similarity to the *P. syringae* pathovars *avellanae,*

actinidiae and *theae*. By contrast, strains of *Pss* did not form a monophyletic clade and were found across phylogroup 2. The strains were interspersed with a diverse range of strains including those isolated from bean and aquatic environments. The other cherry-infecting pathovars (*P.s* pv. *avii* and *P.s* pv. *cerasicola*) were found in phylogroups 1 and 3, although they were not closely related to either race of *Psm*. The plum leaf-wash isolate Ps 9643, which was shown to be non-pathogen in the tree-inoculations of cherry and plum (Chapter 3: Figure 3.3 and 3.4), was positioned in phylogroup 3 and was most closely related to strains from pathovars *avii*, *lachrymans* and *tomato*.

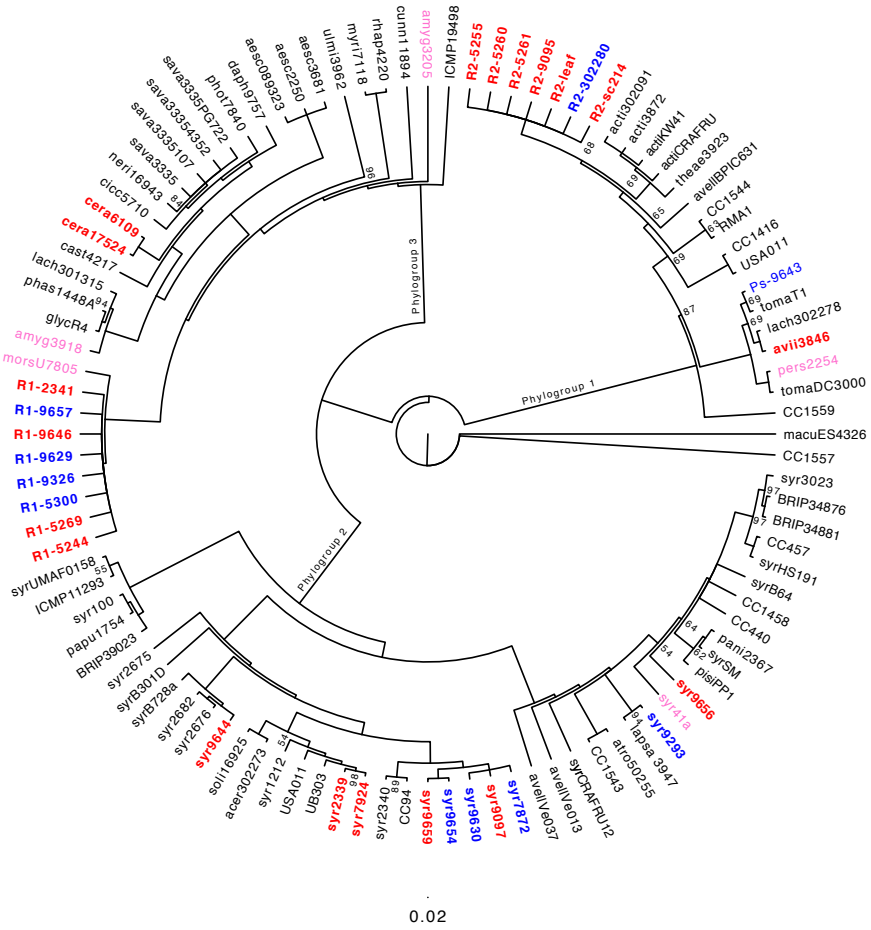


Figure 4.2: Bayesian phylogeny based on seven MLST house-keeping genes *(acnB, fruK, gapA, gltA, gyrB, pgi, rpoD). Isolates from* Prunus *are highlighted and coloured dependent on host of isolation with cherry isolates in red and plum isolates in blue. R1-9657 was originally isolated from cherry but due to pathogenicity tests showing it to possess a low virulence, it is coloured in blue as a non-cherry pathogen within Psm R1. Support values below 99% are shown for each node. Scale bar shows substitutions per site.*

4.4.2.2 Core-genome phylogeny

However, the use of MLST data was somewhat limited as it may not provide enough information to differentiate highly homogeneous strains and resolve complex branches with accuracy (Baltrus *et al.* 2014). Therefore, phylogenies were also produced using whole genome data. First, OrthoMCL was utilised to identify orthogroups present in all 104 strains. An initial core-genome phylogeny was constructed using all 104 strains, however this was found to give low support values on many branches, particularly in phylogroup 2 (Figure S4.2). The removal of two divergent strains (E54326 and CC1557) from the analysis, increased support values. The final tree was based on a 491,425bp concatenated alignment of 902 single-copy orthologous genes that were present in 102 strains. A maximum-likelihood phylogeny was built using RaxML with a GTR + Gamma model of evolution (Figure 4.3). This showed *Psm* R1 actually separates into two clusters of strains, which were not distinguishable using MLST data. The two clusters corresponded to cherry pathogens (strains 5244, 2341, 5269 and 9646) and those that were non-pathogenic on cherry (5300, 9657, 9629, 9326). This indicated that the two groups have diverged as there were allelic differences in their core genomes.

In the core genome phylogeny, the position of another strain designated as *P.s* pv. *morsprunorum* (U7805) isolated from apricot was no longer closely related to *Psm* R1. Instead it clustered with *P.s* pv. *aesculi* and *P.s* pv. *cerasicola*. This finding supports previous reports that MLST loci may not provide enough informative positions to resolve strain positions within each phylogroup accurately (Baltrus *et al.* 2014). The positions of the *Pss* strains within phylogroup 2 also differed between the MLST and core genome phylogenies.

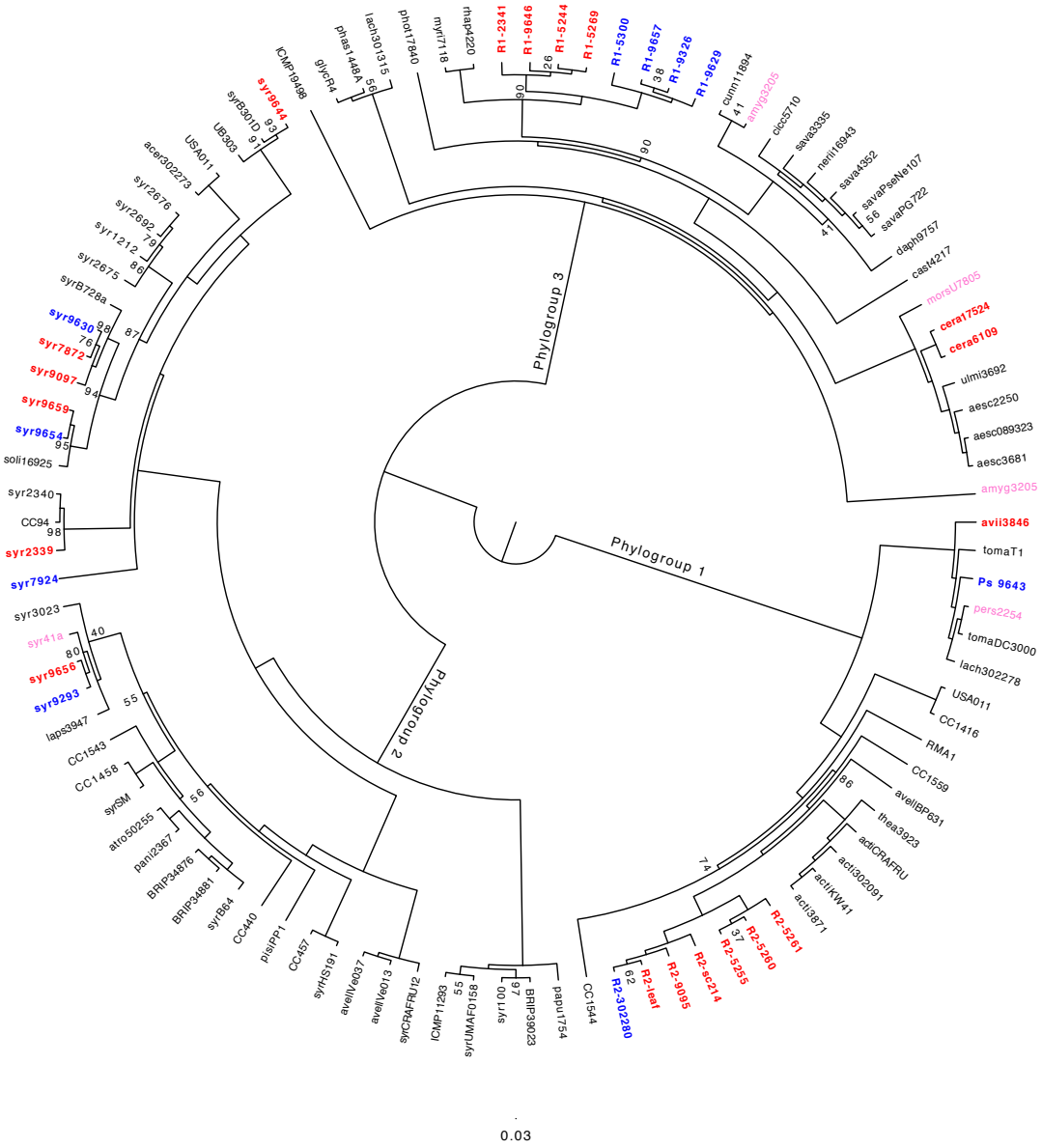


Figure 4.3: ML phylogeny based on 902 genes which represent the core genome of P. syringae*. Isolates from Prunus are highlighted and coloured dependent on host of isolation with cherry isolates in red and plum isolates in blue. R1-9657 was originally isolated from cherry but due to pathogenicity tests showing it to possess a low virulence, it is coloured in blue as a non-cherry pathogen within Psm R1. Support values below 99% are shown for each node. Scale bar shows substitutions per site.*

4.4.2.3 Reference sequence alignment phylogeny

As the two methods above gave different phylogenies, a third method was used for comparison. A whole genome phylogeny was constructed using an alignment

generated by mapping reads to a set of reference genomes with the program RealPhy (Bertels *et al.* 2014). This approach utilised the greatest amount of genomic data as it was not solely based on orthologous gene sequences. RealPhy kept both SNPs and non-polymorphic sites in the alignment used for phylogenetic reconstruction. To avoid the bias that could be introduced by using only one reference genome, three reference genomes (one from each of the three major *P. syringae* phylogroups) were used to produce whole-genome alignments. These alignments were merged to produce a final 1,502,260bp alignment. A Maximum-likelihood phylogenetic tree (Figure 4.4) was then constructed using RaxML.

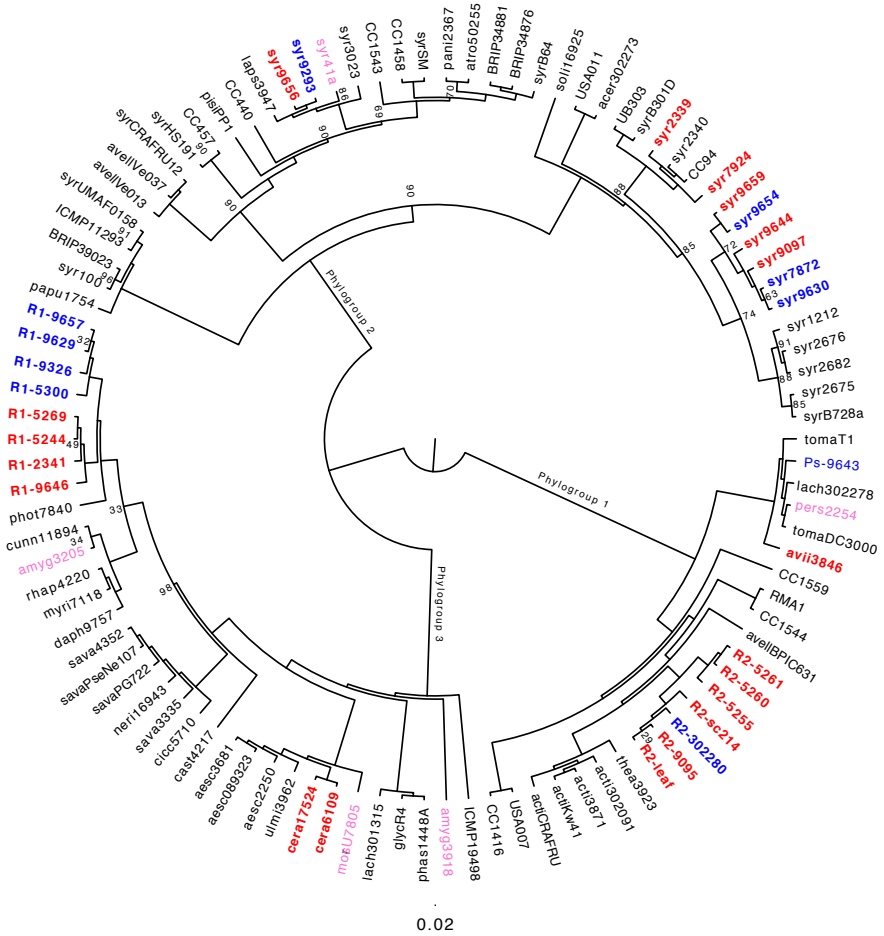


Figure 4.4: ML phylogeny based on the alignment generated using RealPhy*. Isolates from Prunus are highlighted and coloured dependent on host of isolation with cherry isolates in red and plum isolates in blue. R1-9657 was originally isolated from cherry but due to pathogenicity tests showing it to possess a low virulence, it is coloured in blue as a non-cherry pathogen within Psm R1. Support values below 99% are shown for each node. The scale bar shows substitutions per site.*

The overall topology matched the core genome tree and there was good support for each phylogroup. However, comparing both trees showed that within each phylogroup there was variation at the individual strain level, with the RealPhy and core genome tree placing strains of *Psm* R1 and *Prunus*-infecting *Pss* in different positions within their respective phylogroups. In particular, several of the *Prunus*-infecting *Pss* strains within phylogroup 2 (syr9097, syr9659, syr9654, syr9644, syr7872 and syr9630) formed a monophyletic clade, with a clade of Fabaceae (bean and pea) infecting strains as their close out-groups. By contrast, in the core genome tree the *Prunus* strains were interspersed across the phylogroup and the bean strain syrB728a was more closely related to *Prunus*-infecting syr9097 than other bean strains. Also, within phylogroup 3, the closest out-group to the *Psm* R1 clade varied depending on the tree used. The support values of the RealPhy phylogeny were generally lower than those in the core-genome tree, particularly in phylogroups 2 and 3. This suggested that some relationships were difficult to resolve.

4.4.3 Search for virulence factors

4.4.3.1 The *hrp* pathogenicity island

Various studies have shown type III effector proteins are crucially involved in pathogenicity and host specificity (Baltrus *et al.* 2011). First, all sequenced strains were confirmed to contain the *hrp* pathogenicity island and conserved effector locus (CEL) required for conventional Type III secretion (Alfano *et al.* 2000) . Figure 4.5 shows the CEL in the three PacBio sequenced genomes of *Psm* R1, R2 and *Pss*. The core Type III secretion genes were present such as *hrpA*, *hrpW*, *hrpL*, *hrcC, hrpRS* and *hrcV*. Core effectors such as *avrE1*, *hopM1* and *hopAA1* were also present. However, *hopAA1* was truncated in both *Psm* R1 and R2, which were missing the final 374bp and 395bp of this gene respectively. There were also differences in presence of other effectors e.g. *Psm* R2 possessed *hopN1* but lacked *hopA2* seen in both *Psm* R1 and *Pss*. Interestingly, Figure 4.6 shows that the CEL present in cherry-pathogenic strains of *Psm* R1 has undergone an inversion which has caused the truncation of both *hopAA1* and *hopM1*.

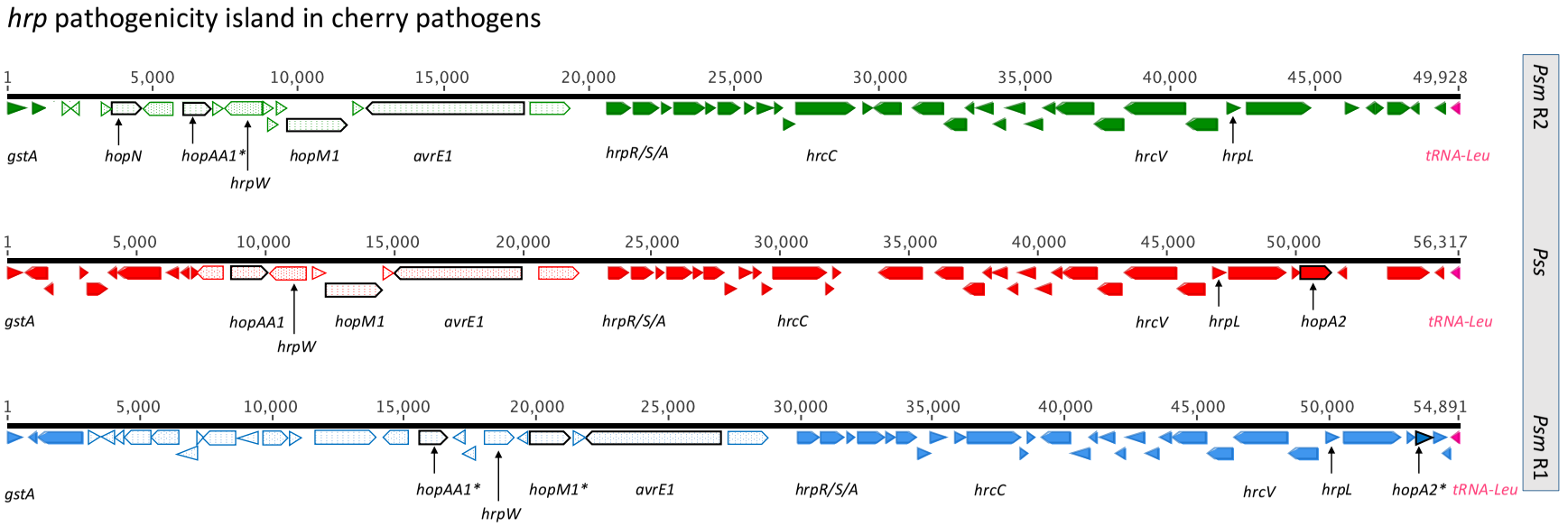


Figure 4.5: Organisation of the structural T3SS and conserved effector locus (CEL) of the three PacBio sequenced pathogens *(R2-leaf, syr9097 and R1-5244). Important structural genes and effectors in this pathogenicity island are labelled. * indicates an effector gene is disrupted. hopAA1 is truncated in* Psm *R2. Whilst in* Psm *R1, hopAA1 and hopM1 are disrupted due to an inversion event and hopA2 is pseudogenised due to a premature stop codon. Genes that form the CEL (as described in* P.s *pv* tomato *DC3000) are shaded. Effector genes are outlined in black.*

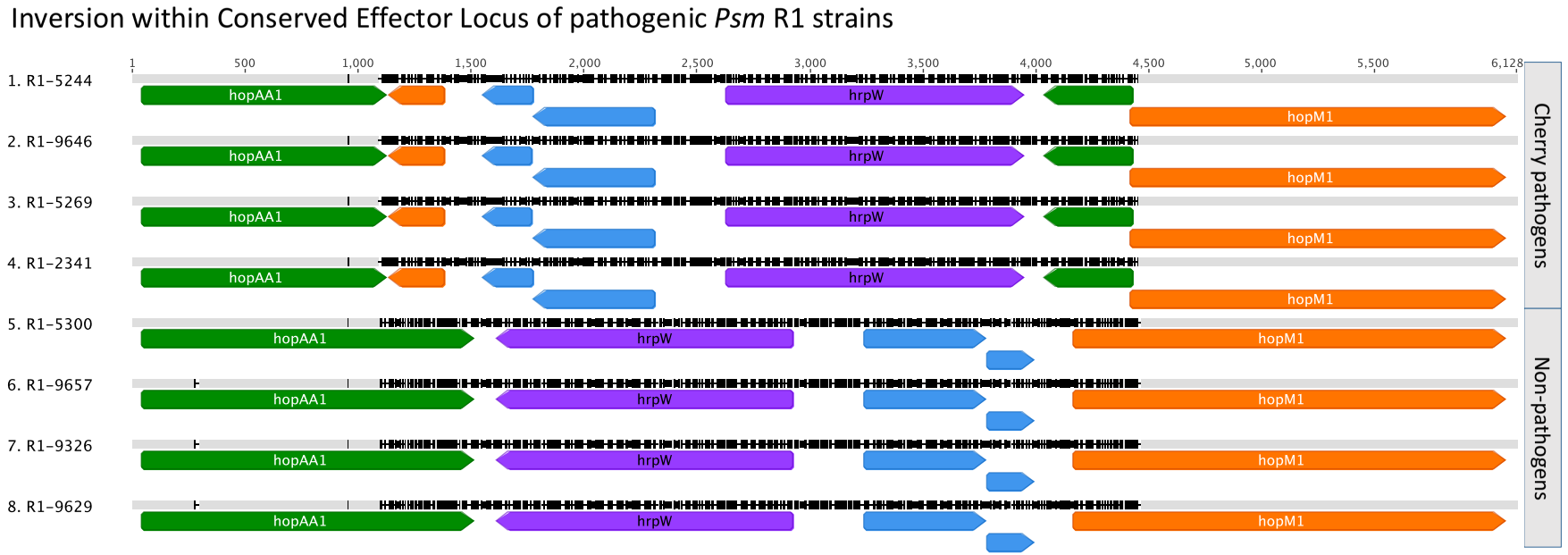
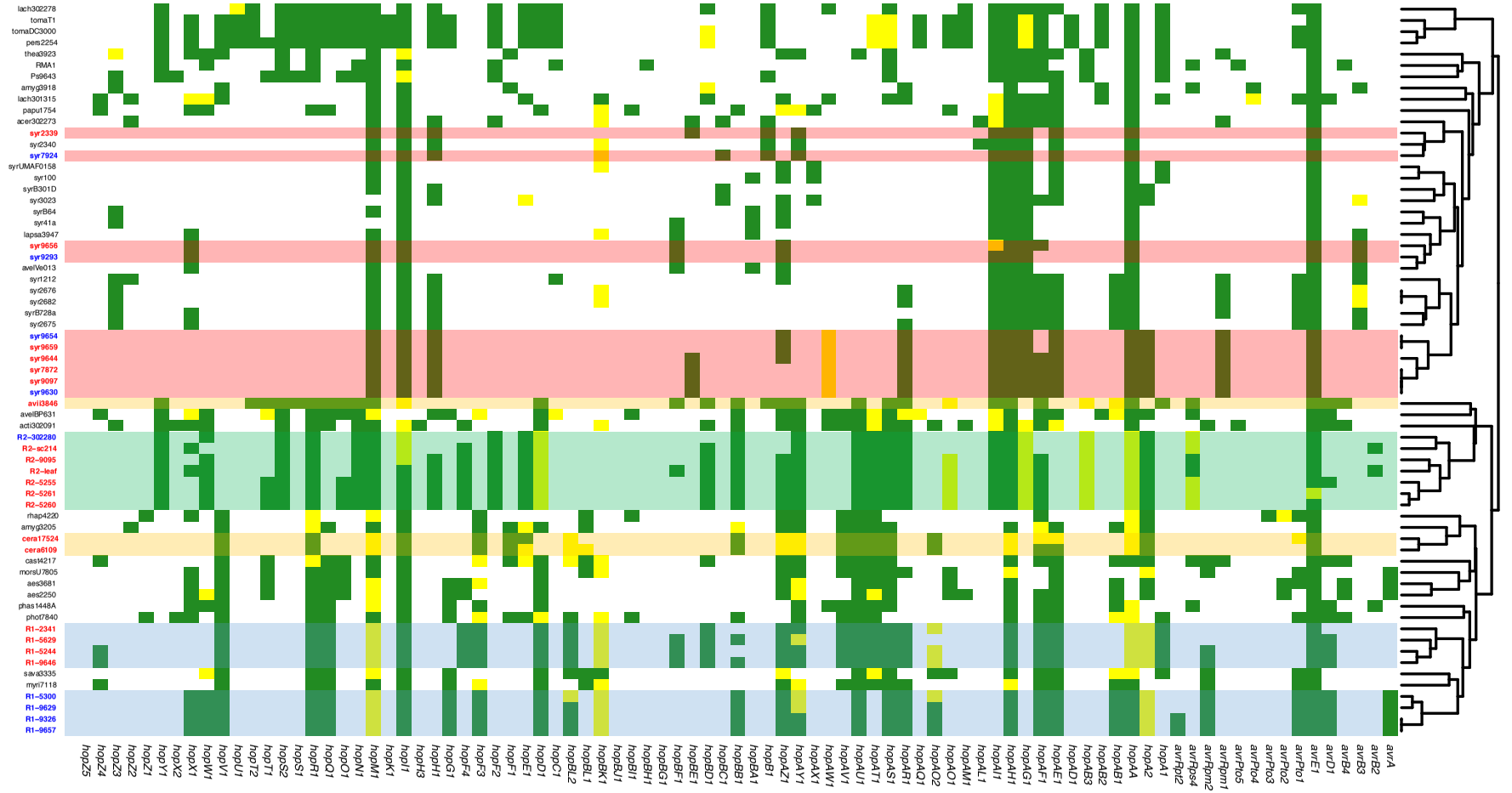


Figure 4.6: Organisation of the CEL in* Psm *R1, *showing an inversion in the four cherry pathogenic strains R1-5244, R1-9646, R1-5269 and R1-2341. An alignment was built of this region using Geneious. Similar sequences are in grey whereas dissimilar sequences appear in black. The effector genes* hopAA1 *and* hopM1 *have been disrupted by a putative sequence inversion.* hopAA1 *is coloured in green,* hopM1 *in orange and* hrpW *in purple. Other protein-encoding genes are in blue. The inversion of this sequence has meant that the end of hopAA1 and beginning of the hopM1 gene have swapped positions, putatively disrupting these effectors. The inversion means that hopM1 no longer possesses an upstream* hrp *box promoter.*

4.4.3.2 Type III effectors

Next, using tBLASTn was used to search 104 genomes for known *P. syringae* effectors (downloaded from pseudomonas-syringae.org). A recent study focusing on the effector family *hopF* (Lo *et al.* 2016) re-characterised some homologues into a new group (*hopF4*), so the analysis was adjusted to include this. Figure 4.7 shows a condensed heatmap of effector presence and absence in *Prunus*-infecting strains and their close out-groups (the full heatmap of 104 strains is presented in Figure S4.3. Strains are clustered based on their effector complement. The heatmap showed that there was a large amount of variation both between *Prunus*-infecting clades and also within the clades themselves in terms of effector presence. Only four effectors were found in all *Prunus* strains: *hopM1*, *hopAA1*, *hopAH1* and *avrE1*. However, *hopAA1* and *hopM1* are truncated in *Psm* R1 strains due to an inversion event and *hopAA1* is also truncated in *Psm* R2. Therefore, the functional status of these effectors is unknown. The status of *hopAH1* as an effector is uncertain, as it may encode a harpin involved in the translocation of effectors rather than an effector protein active within plant cells (Lindeberg *et al.* 2012). Table 4.3 shows a list of effectors in each PacBio/MinIon genome in order of appearance. Some effectors appear to be linked in these genomes (within 10kb), showing that they are likely inherited together.

Figure 4.7 (overleaf): Effector presence and absence heatmaps for* Prunus*-infecting strains and their close out-groups. *The heatmap was generated using the R gplots library. The green squares indicate presence of a full-length homolog of the effector gene whereas yellow squares indicate that the gene is disrupted or truncated in some way. The dendrogram clustered strains by similarity in effector gene presence and absence. Strains infecting* Prunus *are highlighted with cherry pathogens in red and plum pathogens in blue. The* Prunus*-infecting clades are highlighted with* Psm *R1 in blue,* Psm *R2 in green,* Pss *in red and the other pathovars (*P.s *pv.* avii *and* P.s *pv.* cerasicola*) in orange.*


lach302278
tomaT1
tomaDC3000
peis2254
thea3923
RMA1
Ps9643
amyg3918
lach301315
paput754
acer302273
syr2339
syr2340
syr7924
syrUMAF0158
syr100
syrB301D
syr3023
syrB64
syr41a
lapsa3947
syr9656
syr9293
avelVe013
syr1212
syr2676
syr2682
syrB728a
syr2675
syr9654
syr9659
syr9644
syr7872
syr9097
syr9630
avli3846
avelBP631
acti302091
R2−302280
R2−sc214
R2−9095
R2−leaf
R2−5255
R2−5261
R2−5260
rhap4220
amyg3205
cera17524
cera6109
cast4217
morsU7805
aes3681
aes2250
phas1448A
phot7840
R1−2341
R1−5629
R1−5244
R1−9646
sava3335
myri7118
R1−5300
R1−9629
R1−9326
R1−9657
hopZ5
hopZ4
hopZ3
hopZ2
hopZ1
hopY1
hopX2
hopX1
hopW1
hopV1
hopU1
hopT2
hopT1
hopS2
hopS1
hopR1
hopQ1
hopO1
hopN1
hopM1
hopK1
hopI1
hopH3
hopH1
hopG1
hopF4
hopF3
hopF2
hopF1
hopE1
hopD1
hopC1
hopBL2
hopBL1
hopBK1
hopBJ1
hopBI1
hopBH1
hopBG1
hopBF1
hopBE1
hopBD1
hopBC1
hopBB1
hopBA1
hopB1
hopAZ1
hopAY1
hopAX1
hopAW1
hopAV1
hopAU1
hopAT1
hopAS1
hopAR1
hopAQ1
hopAO2
hopAO1
hopAM1
hopAL1
hopAI1
hopAH1
hopAG1
hopAF1
hopAE1
hopAD1
hopAB3
hopAB2
hopAB1
hopAA
hopA2
hopA1
avrRpt2
avrRps4
avrRpm2
avrRpm1
avrPto5
avrPtoB4
avrPtoB3
avrPtoB2
avrPtoB1
avrE1
avrD1
avrB4
avrB3
avrB2
avrA

R1-5244	Contig	Effectors
Chromosome	tig0	*hopAZ1, hopA2, avrE1, hopM1, hopAA1*, hopZ4, hopAT1, hopQ1, hopD1, hopR1, hopF2, hopBL2, hopAV1, hopAO2, hopAY1, hopF3, hopAS1, hopI1, hopAE1, hopAF1-2, hopAU1, hopAH1, hopV1, hopAR1, hopBK1*
Plasmid	tig3	*hopAF1-1, hopBF1, avrD1, avrRpm2, hopBD1*
Plasmid	tig4	*hopA1*
Plasmid	tig5	-
Plasmid	tig6	-
R1-5300	**Contig**	**Effectors**
Chromosome	tig0	*hopAU1, hopAY1, hopF3, hopAF1-2, hopAE1, hopI1, hopBL2, hopAS1, avrPto1, avrRpm2, hopAO2, hopR1, hopD1, hopQ1, hopAB1, hopAA1, hopM1, avrE1*, hopA2, hopQ1-2, avrA1, hopAZ1, hopV1*
Chromosome	tig75	*hopW1, hopBK1, hopAR1*
Plasmid	tig46	-
Plasmid	tig65	*hopX1, hopBB1, hopG1*
Plasmid	tig84	-
Plasmid	tig113	*avrD1*
R2-leaf	**Contig**	**Effectors**
Chromosome	tig0	*hopY1, hopAS1, hopAT1, hopH1, hopF4, hopW1, hopR1, hopAH1-2, hopAI, hopN1, hopAA1, hopM1, avrE1*, hopF2, hopE1, hopA2, hopAH1-1, hopAH1-1**, hopAB3, avrRps4, hopS2, hopI1, hopAR1*
Plasmid	tig5	*hopAO1, hopAZ1, hopAY1*
Plasmid	tig4	*hopD, hopAU1*
Plasmid	tig8	*hopAF1-1, hopBF1*
Plasmid	tig8	*hopBB1, hopBD1*
Plasmid	tig9	*avrB2, hopX1*
syr9097	**Contig**	**Effectors**
Chromosome	tig0	*hopAG1, hopAI1. avrRpm1, hopAR1, hopI1, hopAE1, hopBE1, hopAF1, hopAH1, hopAW1, hopH1, hopA2, avrE1, hopM1, hopAA1**

Table 4.3: Effectors in PacBio/Minion genomes. *Effectors are listed in order of appearance on each assembly contig (chromosomal or plasmid). Where effectors were considered linked (within 10kb) they are underlined. * effectors within the conserved effector locus. ** this effector had multiple annotations.*

4.4.3.3 Number of effectors

Psm R1 and R2 possessed many more putative effector-encoding genes than *Pss*. On average strains of *Psm* R1 and R2 had 30 and 31 effectors respectively, whereas *Pss* strains possessed only 14 effector genes. Figure 4.8 shows a bar chart of number of effectors per strain. The reduced effector repertoire of *Pss* was representative of phylogroup 2 strains generally (Dudnik & Dudler 2014). The *Pss* strains all possessed a set of six core effector genes (*avrE1, hopAA1, hopAG1, hopAH1, hopI1 and hopM1*) and 5-10 additional effectors which varied between strains.

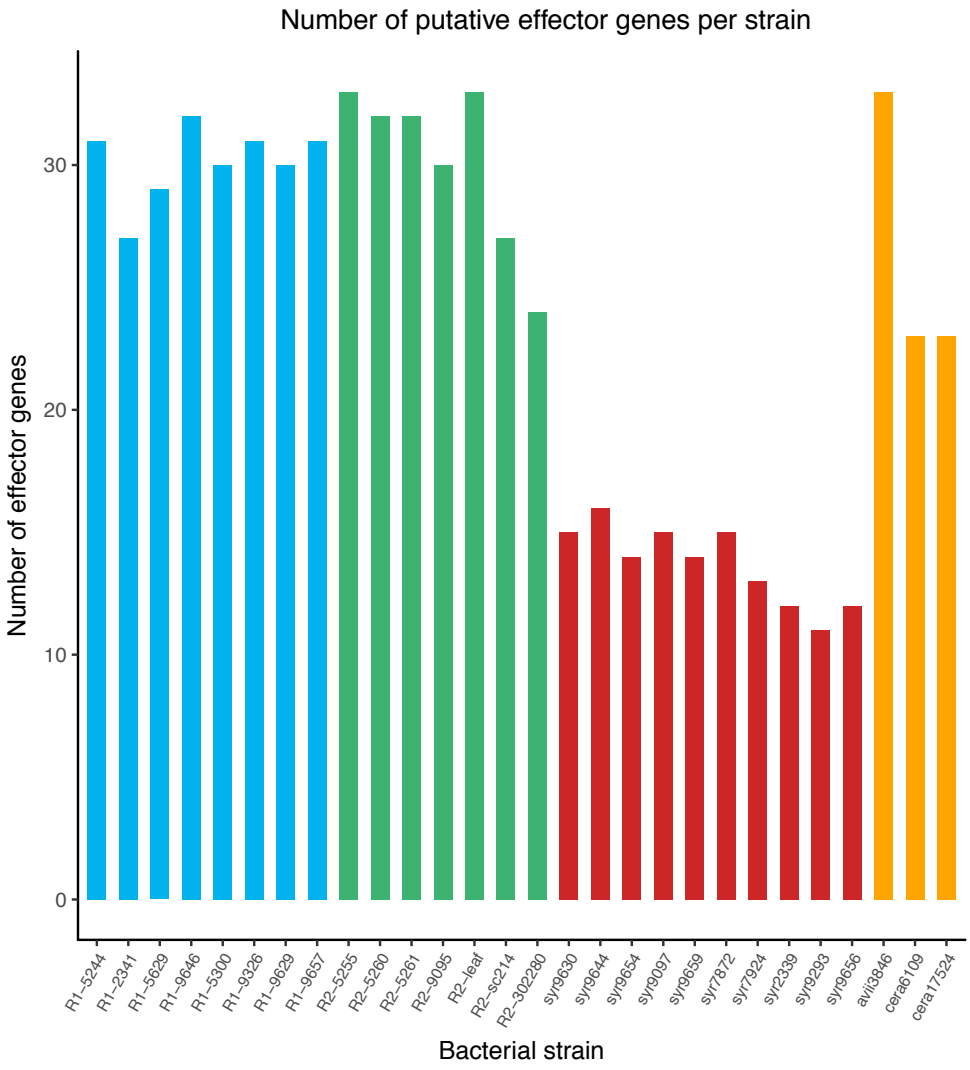


Figure 4.8: Barplot showing the number of putative effector genes identified in each* Prunus*-infecting strain. *The* Prunus*-infecting clades are highlighted with* Psm *R1 in blue,* Psm *R2 in green,* Pss *in red and the other pathovars (*P.s *pv.* avii *and* P.s *pv.* cerasicola*) in orange.*

4.4.3.4 Lack of a redundant effector group is linked to host specificity

Within *Psm* R1, although the strains formed a monophyletic clade based on core genes, they did not cluster together based on effector presence. Indeed, the two groups (pathogens and non-pathogens), had strikingly different effector complements indicating that their flexible genomes (plasmids and genomic islands) may differ significantly, despite a similar core genome. Table 4.4 lists the differential effectors in the two groups. Those present in pathogenic strains are potential candidate virulence genes, whilst those present in non-pathogenic strains could be avirulence genes detected by the cherry immune system so prevent these strains growing *in planta*. Interestingly, both *avrPto1* and *hopAB1* were absent from pathogenic strains. These two effectors have been shown to form a redundant effector group (REG) vital for pathogenicity of *P.s* pv. *tomato* DC3000 in *N. benthamiana* as they act early to prevent PAMP detection and interrupt immune responses (Kvitko *et al.* 2009; Cunnac *et al.* 2011) The DNA regions surrounding these effectors in *Psm* R1 were aligned. A 38kb region containing the *hopAB1* gene is absent from pathogenic *Psm* R1, with homologous sequences at either end. Whilst, a 2411bp region found in the non-pathogen R1-5300, which contained the *avrPto1* gene, has been replaced with a 2913bp region containing three mobile element proteins in the pathogenic strain (R1-5244) (Figure 4.9).

Pathogens	Non-pathogens
hopA1	*avrA1*
*hopAF1**	*avrPto1*
hopAT1	*avrRpt2*
hopAV1	*hopAA1*
hopBD1	*hopAB1*
hopBF1	*hopAO2*
hopF4	*hopG1*
hopZ4	*hopW1*
	hopX1

Table 4.4: Effector presence comparisons between cherry pathogenic and non-pathogen* Psm *R1 strains**. * Both strains possess hopAF1-1 but the pathogenic strains possess an additional divergent copy, present on a plasmid in R1-5244

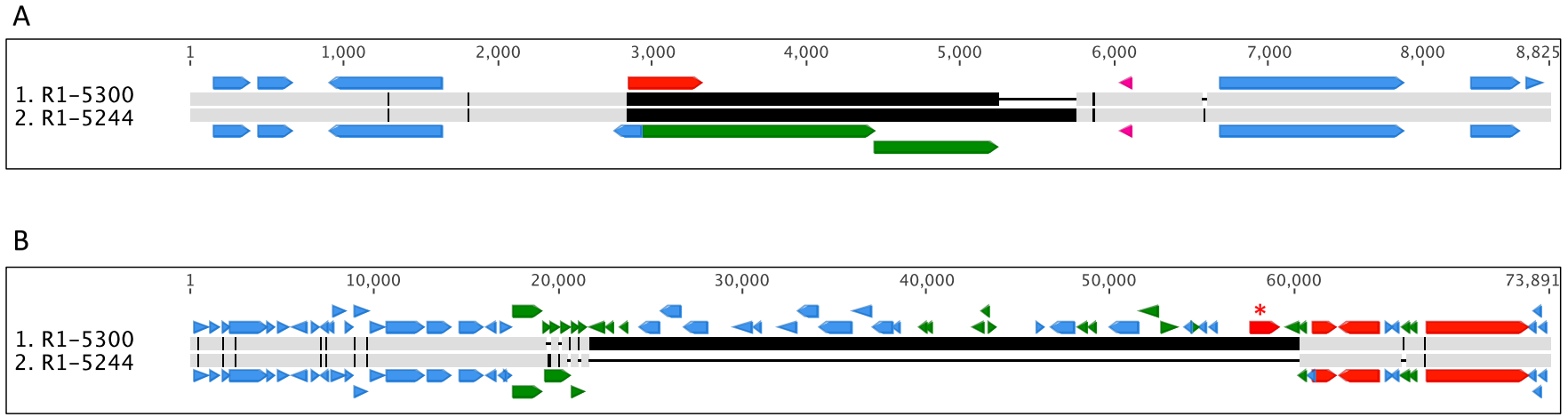
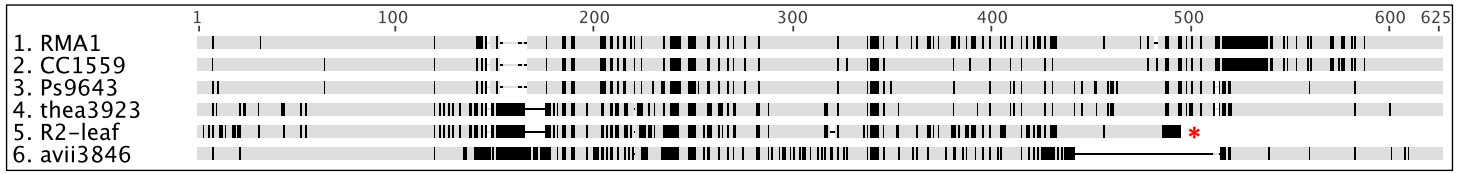


Figure 4.9: Alignment of the DNA regions surrounding the* avrPto1 *(A) and* hopAB1 *(B) *showing how these genes have been lost from* Psm *R1 cherry pathogens. Grey indicates sequence identity whereas black indicates divergence. The effector genes are coloured in red, whereas other CDS are in blue, mobile element genes are in green and tRNA genes are pink. In B, * indicates the location of* hopAB1 *in R1-5300, whilst the upstream effectors are* hopQ1, hopD1 *and* hopR1.

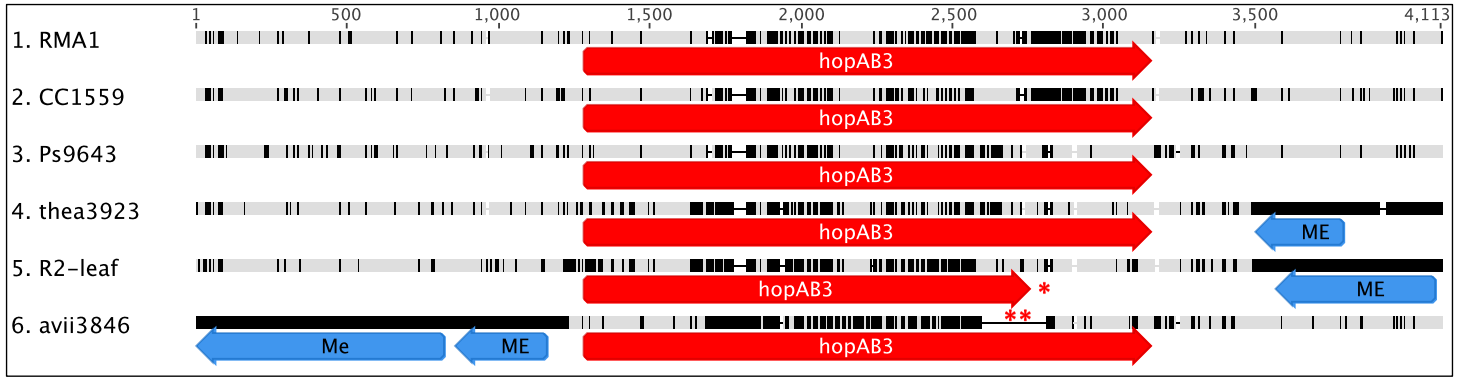
In *Psm* R2 there were no homologues of *avrPto*, however, there was a homologue of *hopAB* (*hopAB3*). A protein alignment indicated that it was truncated compared to other *P. syringae hopAB3* proteins (Figure 4.10A). DNA alignments (Figure 4.10B) showed this was caused by a 2bp insertion (GG) that has resulted in a frame-shift and truncation of the protein and putative pseudogenisation. Further analysis revealed that the truncation of *hopAB3* in *Psm* R2 meant that it lacked the E3 Ubiquitin Ligase domain. In addition, variation in the Pto-interaction domain (PID), which is where HopAB putatively interacts with the immune protein Pto, meant that this domain is not picked up in the *Psm* R2 homologue using the protein domain finder Interproscan (Figure 4.10C). The avrPto/hopAB REG was therefore potentially absent from both *Psm* R1 and R2. It was also absent from all *Prunus*-infecting *Pss* strains. To examine if this was common to other cherry pathogens, these genes were examined in *P.s* pv. *avii* and *P.s* pv. *cerasicola*. The gene *avrPto1* was truncated in *P.s* pv. *cerasicola*, whilst *P.s* pv. *avii* had two variants of *hopAB* (*hopAB1* and *hopAB3*). Both *hopAB* alleles were classed as potential pseudogenes as *hopAB1* was a truncated version of the full effector (also found in other strains such as *P.s* pv. *savastanoi* 3335) (Figure 4.11) and *hopAB3* had a deletion that disrupted the E3 Ubiquitin Ligase domain and again lacked homology to the PID of other HopAB3 proteins (Figure 4.10).

Figure 4.10 (overleaf): Truncation of* hopAB3 *in* Psm *R2. *A: Protein alignment of HopAB3. Grey indicates sequence identity whereas black indicates divergence. The Psm R2-leaf HopAB3 protein is truncated (*). B: Nucleotide sequence alignment of the* hopAB3 *gene and surrounding regions in strains including Psm R2. The hopAB3 gene has been truncated due to a GG insertion leading to a frameshift in Psm R2 (*). In avii3846 there is a deletion at the end of the gene (**).C: Diagram showing the location of key domains in the HopAB3 protein including the Pto-interacting domain (PID), BAK1-interacting domain (BAK1) and E3 Ubiquitin ligase (E3). In avii3846 the beginning of this domain is lost so it is outlined with a dotted line. Genes labelled ME were annotated as mobile elements.*

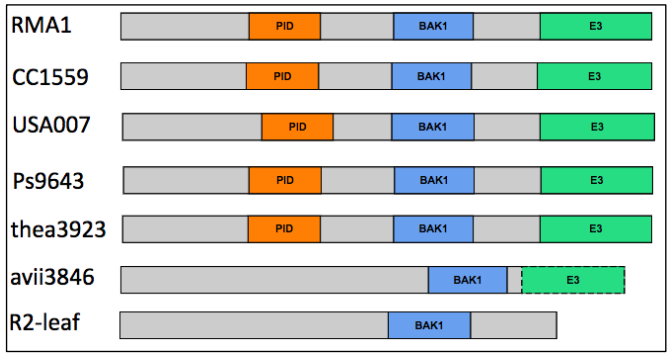
A


1
100
200
300
400
500
600
625
1. RMA1
2. CC1559
3. Ps9643
4. thea3923
5. R2-leaf
6. avii3846
*

B


1
500
1,000
1,500
2,000
2,500
3,000
3,500
4,113
1. RMA1
2. CC1559
3. Ps9643
4. thea3923
5. R2-leaf
6. avii3846
hopAB3
ME
Me
*
**

C


RMA1
CC1559
USA007
Ps9643
thea3923
avii3846
R2-leaf
PID
BAK1
E3

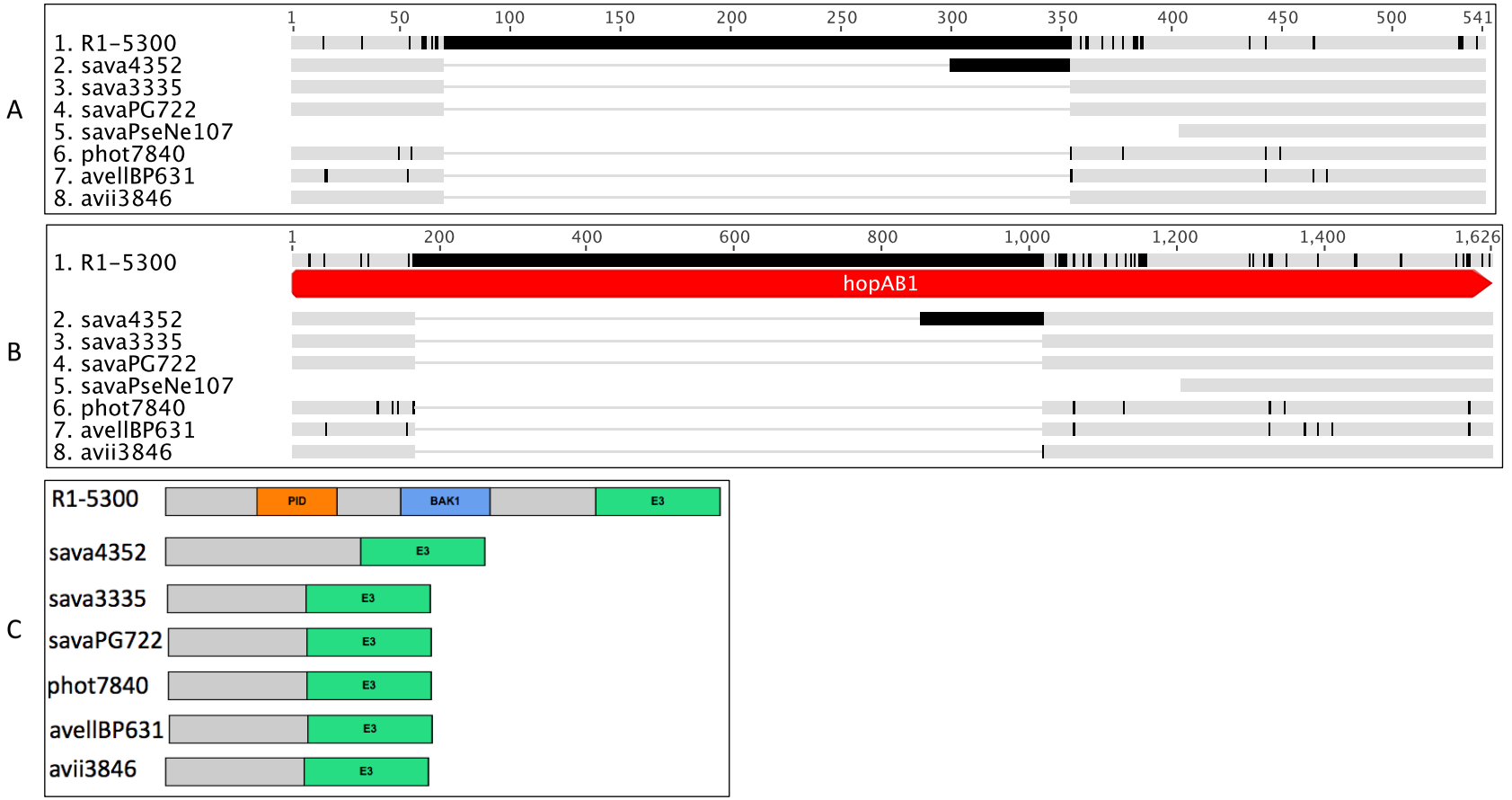


Figure 4.11: Some pathovars possess a truncated version of the* hopAB1 *gene. *A: Protein alignment of the* HopAB1 *protein. Grey indicates sequence identity whereas black indicates divergence. The full-length version from R1-5300 is presented, whilst the truncated versions of this gene have a large deletion in the middle. The truncated version has homology with the beginning and end of the full-length protein. B: The nucleotide alignment of the hopAB1 gene and surrounding regions. C: Diagrams of HopAB1 protein domains showing that the truncated version lacks the PID- and BAK1-interacting domains.*

This analysis showed that in all cherry pathogens the *avrPto*/*hopAB* REG was absent or when genes were present, they were truncated or disrupted in some way. A heatmap of the REG genes alone showed that in Phylogroup 2 (which contains *Prunus*-infecting *Pss*) this REG is uncommon, however in the other phylogroups most members were found to possess at least one REG member (Figure 4.12).

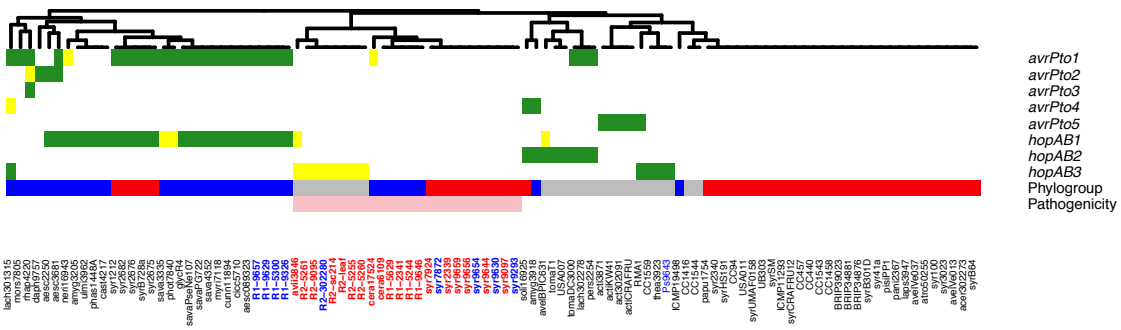


Figure 4.12: Heatmap showing the presence and absence of the* avrPto/hopAB *REG across the* P. syringae *complex. *The green squares indicate presence of a full-length homolog of the effector gene whereas yellow squares indicate that the gene is disrupted or truncated in some way. The phylogroup is indicated by either grey (phylogroup 1), red (phylogroup 2) or blue (phylogroup 3). Cherry pathogenicity is indicated by the pink squares. The dendrogram clustered strains by similarity of traits (gene presence, phylogroup and cherry pathogenicity). Strains isolated from cherry and plum are highlighted in bold, with cherry pathogens in red and plum pathogens in blue. The non-pathogenic plum isolate Ps 9643 is also highlighted in blue. Strain names in black represent other pathovars that don't infect cherry and plum.*

4.4.4 Gain and loss analysis of effector genes

The program GLOOME (Cohen *et al.* 2010) was then used to plot the gain and loss of effectors onto the core genome phylogenetic tree, focusing on the branches leading up to the cherry pathogenic strains. This approach highlighted significant events that have occurred during host specialisation. GLOOME took a binary matrix of effector presence and absence data and the core genome phylogeny and used maximum parsimony or stochastic mapping to infer when effectors were gained or lost. Both full-length and truncated effector gene homologues were considered to be present. Figures 4.13 and 4.14 illustrates the gain and loss of effectors on the branches leading to cherry pathogenic strains. Parsimony predicted that several key effectors have been

gained in multiple clades. These included members of *hopAF1*, *hopAO*, *hopAR1*, *hopA*, *hopAZ1*, *hopBB1*, *hopBD1*, *hopBF1*, *hopF*, *avrD1, avrRpm1, avrRpm2* and *avrB*. For example, *hopBB1* has been gained along the branches leading to *Psm* R1, *Psm* R2*, P.s* pv. *avii* and *P.s* pv. *cerasicola*. *hopAR1* has also been gained in *Psm* R1, *Psm* R2, *P.s* pv. *avii* and members of *Pss*. Aside from new effectors, pathogenic strains of *Psm* R1 had also gained an additional identical copy of the effector *hopAF1,* which is on a plasmid in the PacBio sequenced strain *Psm* R1 5244. The convergent acquisition of these candidate effectors suggests they may be important in cherry pathogenicity. As the use of long-read sequencing produced complete genomes for several representative cherry pathogens, the genomic locations of gained effectors could be determined. In *Psm* R1 and *Psm* R2, many of these gained effector genes were located on plasmids in the PacBio-sequenced strains, suggesting that plasmid-mobility is a key mechanism for gaining additional effector genes.

One putatively important effector, *hopBB1* is closely related to members of the *hopF* family and *avrRpm2* as they share their N-terminus domain (Lo *et al.* 2016). In addition to the significant acquisition of *hopBB1* homologues, cherry pathogens also gained members of *hopF* and *avrRpm2*. Pathogenic strains in *Psm* R1, R2 and *P.s* pv. *cerasicola* all possessed two different *hopF* homologues each (*hopF3* and *hopF4*, *hopF2* and *hopF4*, and *hopF1* and *hopF3* respectively). The wild cherry strain *P.s* pv. *avii* did not possess any *hopF* homologues but had gained *hopBB1*. By contrast, *Pss* strains that infect cherry lacked all *hopF* family members. The expansion of *hopF* and related effectors within cherry-infecting clades of *P. syringae* indicates that the functions of this effector family may play a key role in the establishment of this disease.

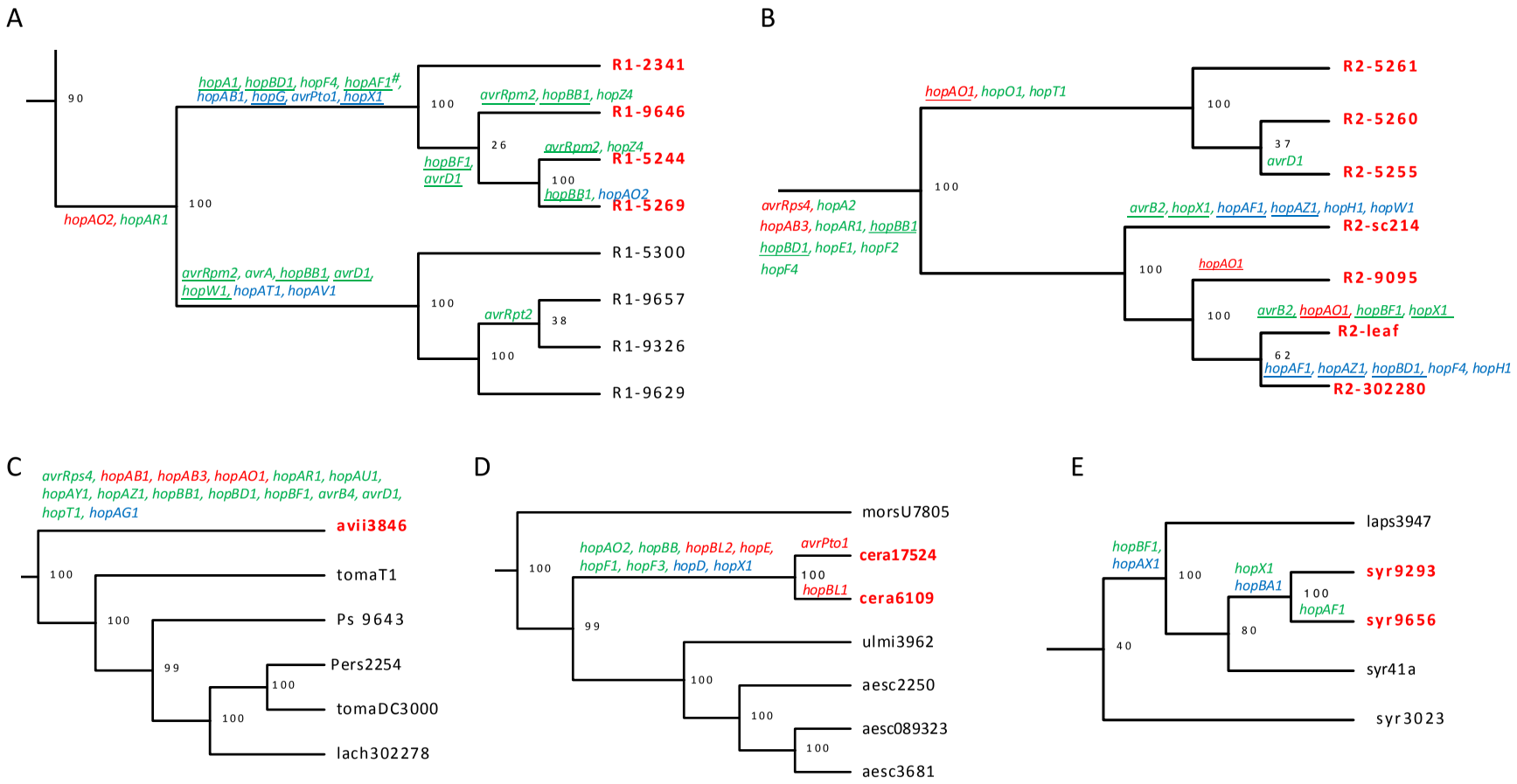


Figure 4.13: Phylogenetic trees with the gain and loss of effector genes occurring on each branch leading to cherry pathogenic strains. *This was based on the GLOOME analysis of the core genome phylogeny. Gained effectors are coloured in green, lost effectors in blue and those that have been gained but are pseudogenised are coloured in red. Where an effector gene is underlined it has been identified on a plasmid in one of the complete genomes. Cherry pathogenic strains are highlighted in red. A:* Psm *R1, B:* Psm *R2, C:* P.s *pv.* avii*, D:* P.s *pv.* cerasicola*, E: divergent* Pss *strains.*

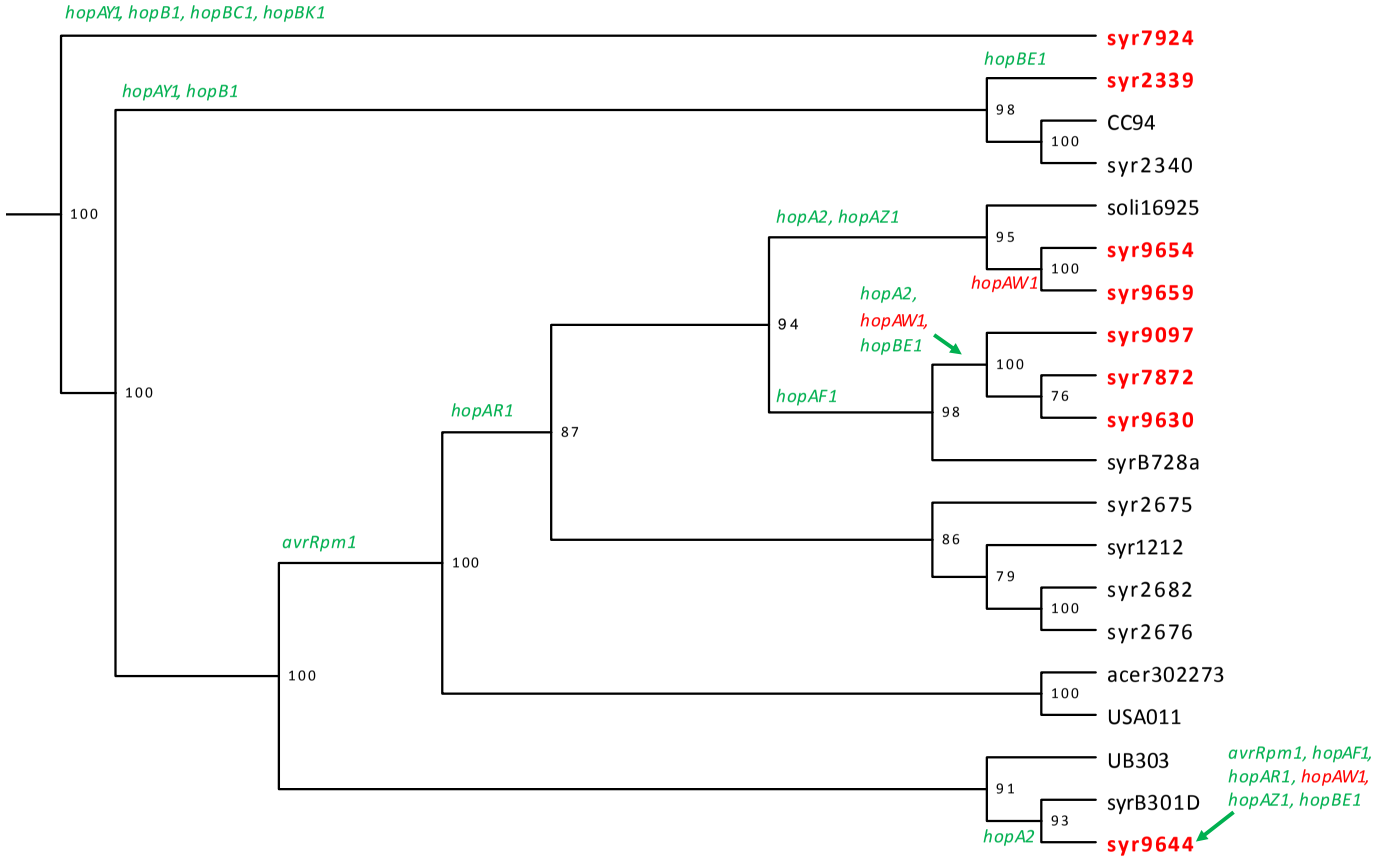


Figure 4.14: Phylogenetic trees with the gain and loss of effector genes occurring on each branch leading to cherry pathogenic* Pss *strains. *This was based on the GLOOME analysis of the core genome phylogeny. Gained effectors are coloured in green, lost effectors in blue and those that have been gained but are pseudogenised are coloured in red. Where an effector gene is underlined it has been identified on a plasmid in one of the complete genomes. Cherry pathogenic strains are highlighted in red.*

GLOOME also predicted the loss of effectors. On the branch leading to *Psm* R1 cherry pathogens four effectors had been lost (*hopAB1*, *hopG1*, *avrPto1* and *hopX1*). The loss of *hopX1* may be correlated with the gain of biosynthesis of the phytotoxin coronatine, as this effector has been shown to function redundantly with coronatine (Gimenez-Ibanez *et al.* 2014). Both *hopG1* and *hopX1* are located on a plasmid in the non-pathogenic strain *Psm* R1 (R1-5300) (see Table 4.3), so could have been lost together in the pathogenic lineage. In *Psm* R2, the strains R2-sc214 and R2-302280 had both lost several effectors. These include *hopAF1*, *hopAZ1* and *hopBD1.* These effector genes were located on plasmids in the PacBio strain R2-Leaf, indicating that these strains may have lost/not gained these plasmids. The main *Pss* clade within phylogroup 2 (Figure 4.14) did not experience any effector losses. As members of this phylogroup generally possessed fewer effector genes, they may be less likely to lose effectors as their repertoires contain only those genes whose products are core to pathogenicity.

Focusing on the *avrPto*/*hopAB* REG, GLOOME analysis revealed there has been loss of this group in *Psm* R1. Where the REG was present in other cherry-infecting clades, homologues were truncated or disrupted, indicating that either the REG is not required or functions in avirulence. Surprisingly, the GLOOME analysis revealed that the branches leading to *Psm* R2, *P.s* pv. *avii* and *P.s* pv. *cerasicola* all experienced gains in these effector groups, suggesting that these effectors have been gained in the ancestors of cherry pathogens.

GLOOME analysis was also performed using the RealPhy-based phylogeny as the topology had varied slightly to the core genome phylogeny (compare Figures 4.15 and 4.16). The results generally supported those produced using the core-genome tree, although exact branches on which gains and losses occurred varied slightly. When using the core-genome, there appeared to be convergent gain of particular effectors such as *hopAR1* and *hopAW1* in the cherry-infecting *Pss* (Figure 4.14). However, when using the RealPhy phylogeny, as most of these strains formed a monophyletic clade, these gains only occurred once in the phylogeny.

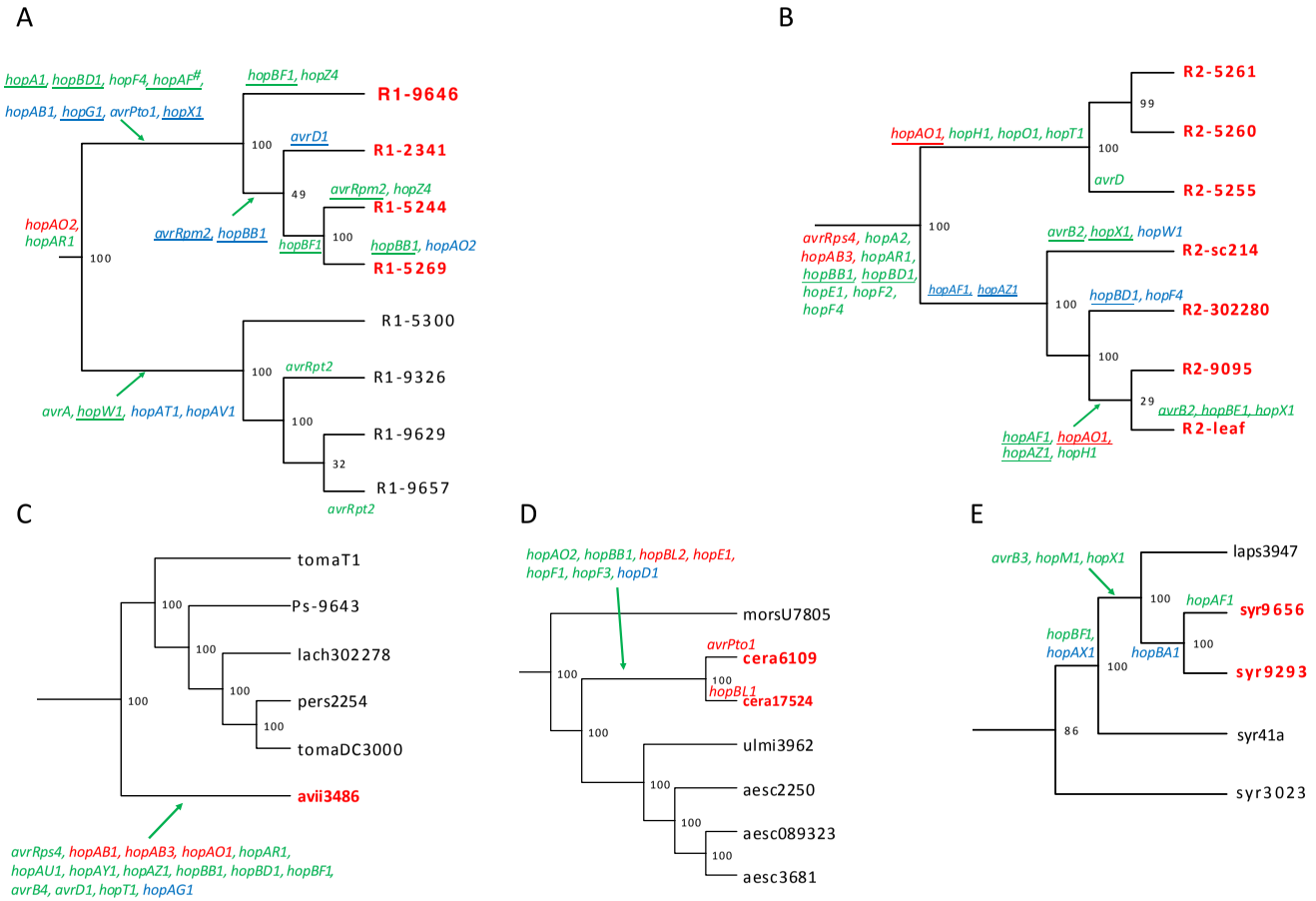


Figure 4.15: Phylogenetic trees with the gain and loss of effector genes occurring on each branch leading to cherry pathogenic strains. *This was based on the GLOOME analysis of the RealPhy phylogeny. Gained effectors are coloured in green, lost effectors in blue and those that have been gained but are pseudogenised are coloured in red. Where an effector gene is underlined it has been identified on a plasmid in one of the complete genomes. Cherry pathogenic strains are highlighted in red. A:* Psm *R1, B:* Psm *R2, C:* P.s *pv.* avii*, D:* P.s *pv.* cerasicola*, E: divergent* Pss *strains.*

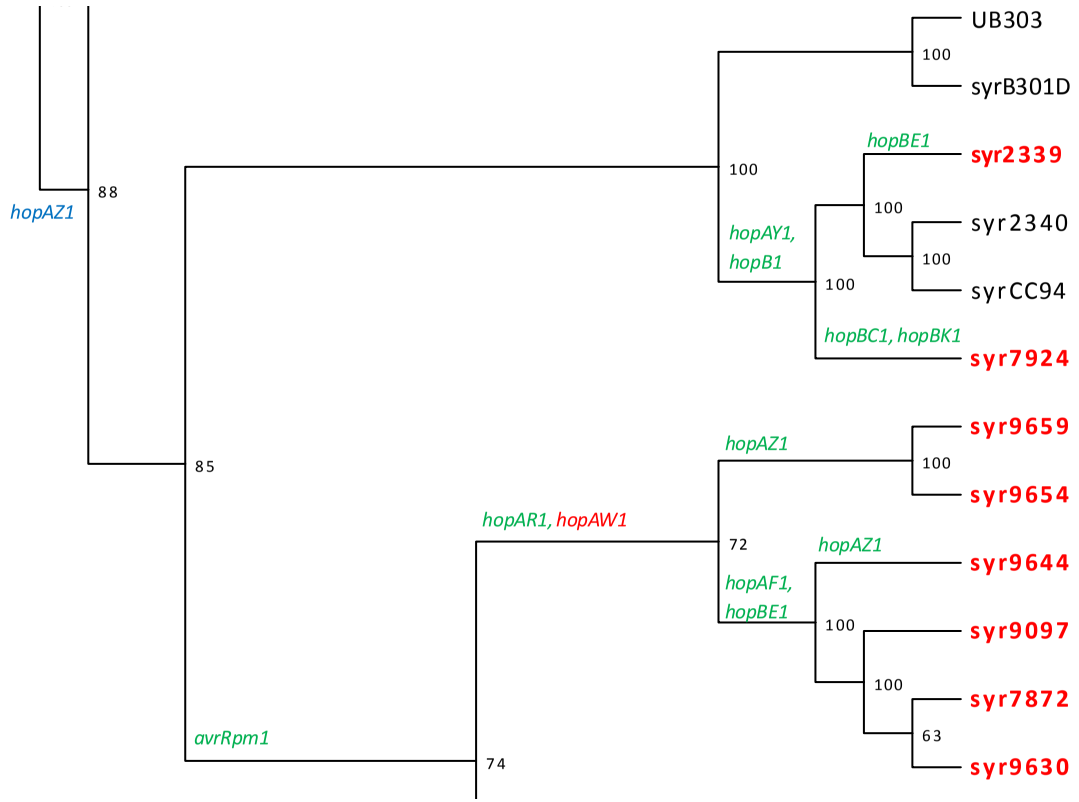


Figure 4.16: Phylogenetic trees with the gain and loss of effector genes occurring on each branch leading to cherry pathogenic* Pss *strains. *This was based on the GLOOME analysis of the RealPhy phylogeny. Gained effectors are coloured in green, lost effectors in blue and those that have been gained but are pseudogenised are coloured in red. Where an effector gene is underlined it has been identified on a plasmid in one of the complete genomes. Cherry pathogenic strains are highlighted in red.*

4.4.5 Association of effector families with cherry pathogenicity

GLOOME analysis provided an insight into the patterns of effector gain and loss leading to cherry pathogenicity. To further analyse whether the evolution of cherry pathogenicity is correlated with the presence of particular effectors, the program BayesTraits was used. BayesTraits took an input matrix of the presence and absence of two traits and a phylogenetic tree. It determined if the two characters (effector gene and cherry pathogenicity) have evolved independently or dependently. The program ran both models of evolution and from this a likelihood ratio was calculated for each effector and a p-value determined using a Chi-Squared test.

Using the core genome phylogeny, BayesTraitsV2 was ran 100 times. Effectors that were potential pseudogenes were designated as absent as they may not be functional *in planta*. Effectors that produced a p-value <0.05 in >90% of the runs were significantly correlated with cherry pathogenicity. Figure 4.17 shows the likelihood ratio of all effector genes. Candidate effectors whose presence/absence was significantly associated with cherry pathogenicity are highlighted. Those predicted to be associated with pathogenicity (present in cherry pathogens) included *hopBF1*, *hopAR1* and *avrD1*. Whereas, *hopAB1* was significantly lost in cherry pathogens based on the phylogenetic tree, meaning it could be a candidate avirulence gene. Some other effectors that were previously hypothesised to be important due to their presence in multiple pathogen clades, also had low p-values that were almost significantly correlation with cherry pathogenicity. These included *hopBB1* (p=0.07) and *avrRpm2* (p=0.06) which are putative virulence factors.

Figure 4.17 (overleaf): Barplot showing the output from BayesTraits analysis of effector gene and pathogenicity evolution using the core genome phylogeny. *The barplot shows the likelihood ratio for the correlation of each effector gene with pathogenicity. The values are obtained from means of 100 independent runs of the program with error bars showing standard error above and below the mean. Those effectors that were not significantly associated with pathogenicity are coloured in grey. Effector genes where presence of the gene was associated with pathogenicity ($p \leq 0.05$) are coloured in blue, whilst where absence of the gene was associated with pathogenicity ($p \leq 0.05$) the bar is coloured in red.*

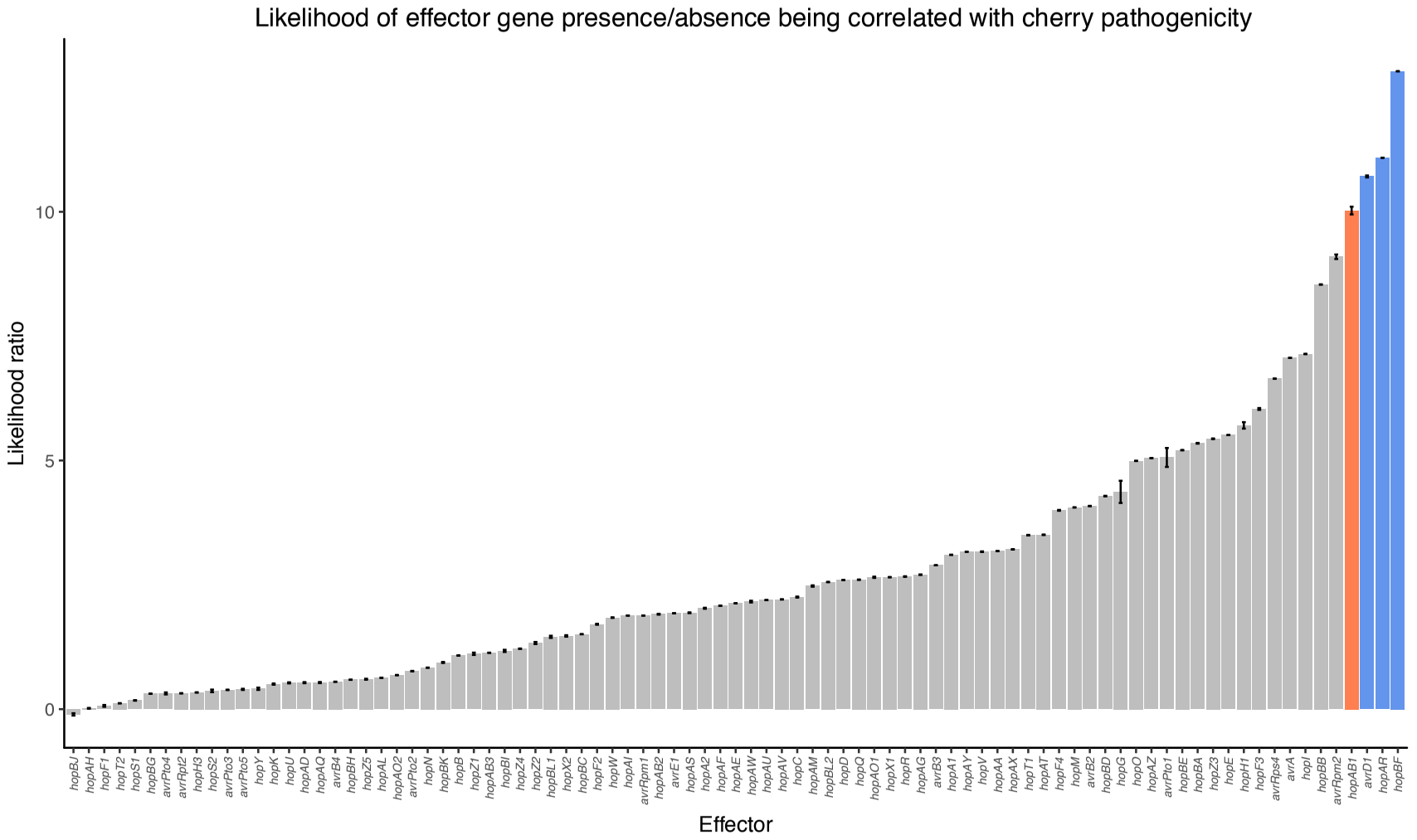
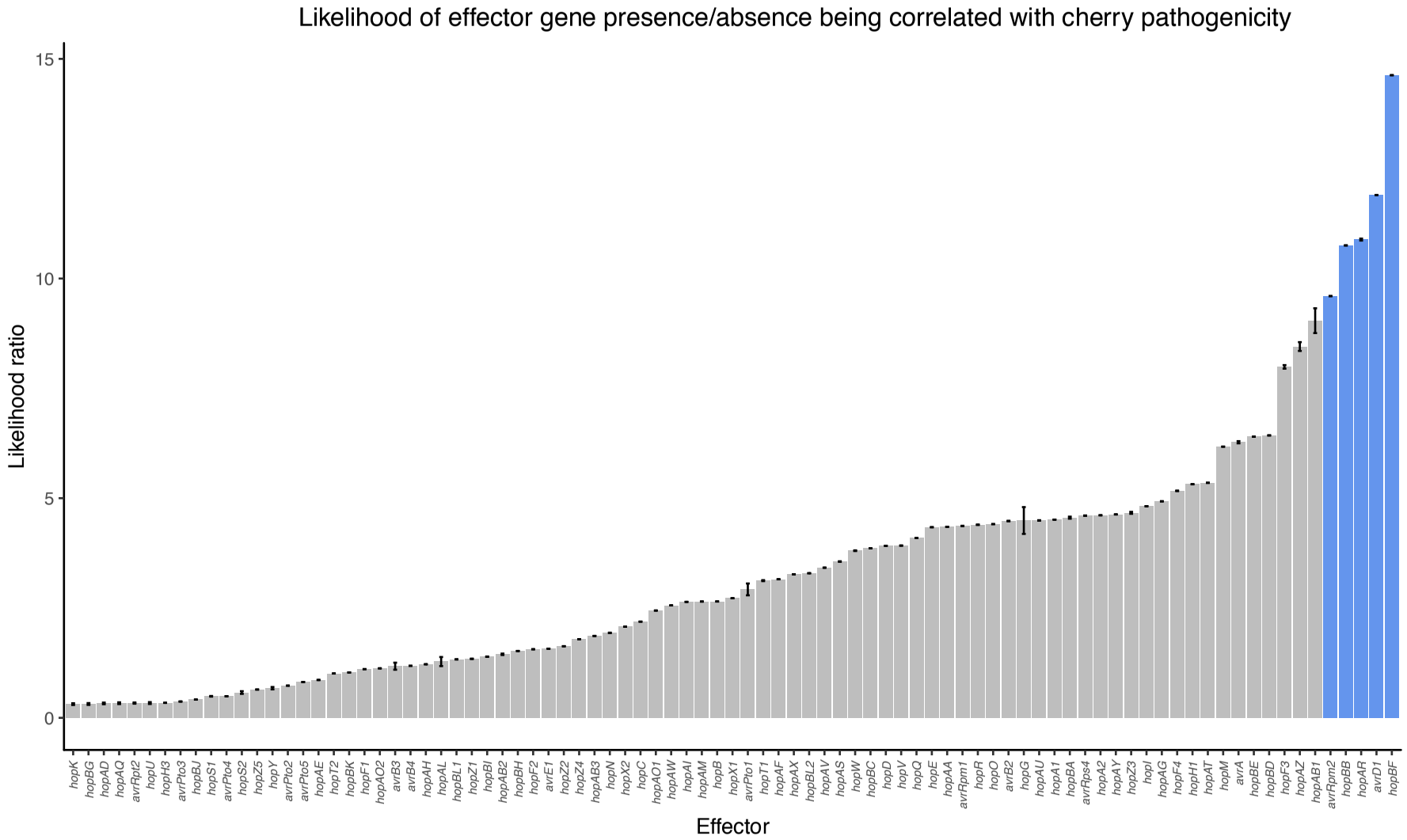

Likelihood of effector gene presence/absence being correlated with cherry pathogenicity
Likelihood ratio
0
5
10
Effector
hopBJ
hopAH
hopF1
hopT2
hopS1
hopBG
avrPto4
avrRpt2
hopH3
hopS2
avrPto3
avrPto5
hopY
hopK
hopU
hopAD
hopAQ
avrB4
hopBH
hopZ5
hopAL
hopAO2
avrPto2
hopN
hopBK
hopB
hopZ1
hopAB3
hopBI
hopZ4
hopZ2
hopBL1
hopX2
hopBC
hopF2
hopW
hopAI
avrRpm1
hopAB2
avrE1
hopAS
hopA2
hopAF
hopAE
hopAW
hopAU
hopAV
hopC
hopAM
hopBL2
hopD
hopQ
hopAO1
hopX1
hopR
hopAG
avrB3
hopA1
hopAY
hopV
hopAA
hopAX
hopT1
hopAT
hopF4
hopM
avrB2
hopBD
hopG
hopO
hopAZ
avrPto1
hopBE
hopBA
hopZ3
hopE
hopH1
hopF3
avrRps4
avrA
hopI
hopBB
avrRpm2
hopAB1
avrD1
hopAR
hopBF

Figure 4.18 shows results of the same analysis using the RealPhy tree. The same effectors (*hopBF1*, *avrD1*, *hopAR1*) were significantly associated with cherry pathogenicity. This analysis also predicted *hopBB1* and *avrRpm2* to be significantly correlated ($p<0.05$ in >90% of runs). The candidate avirulence effector *hopAB1* had a slightly higher p-value when using the Realphy phylogeny (mean $p=0.13$), so was not considered significantly associated with pathogenicity in this analysis.

Figure 4.18 (overleaf): Barplot showing the output from BayesTraits analysis of effector gene and pathogenicity evolution using the RealPhy phylogeny. *The barplot shows the likelihood ratio for the correlation of each effector gene with cherry pathogenicity. The values are obtained from means of 100 independent runs of the program with error bars showing standard error. Those effectors that were not significantly associated with pathogenicity are coloured in grey. Effector genes where presence of the gene was associated with pathogenicity ($p \leq 0.05$) are coloured in blue.*


Likelihood of effector gene presence/absence being correlated with cherry pathogenicity
Likelihood ratio
15
10
5
0
hopK hopBG hopAD hopAQ avrRpt2 hopU hopH3 avrPto3 hopBJ hopS1 avrPto4 hopS2 hopZ5 hopY avrPto2 avrPto5 hopAE hopT2 hopBK hopF1 hopAO2 avrB3 avrB4 hopAH hopAL hopBL1 hopZ1 hopBI hopAB2 hopBH hopF2 avrE1 hopZ2 hopZ4 hopAB3 hopN hopX2 hopC hopAO1 hopAW hopAI hopAM hopB hopX1 avrPto1 hopT1 hopAF hopAX hopBL2 hopAV hopAS hopW hopBC hopD hopV hopQ hopE hopAA avrRpm1 hopR hopO avrB2 hopG hopAU hopA1 hopBA avrRps4 hopA2 hopAY hopZ3 hopI hopAG hopF4 hopH1 hopAT hopM avrA hopBE hopBD hopF3 hopAZ hopAB1 avrRpm2 hopBB hopAR avrD1 hopBF
Effector

4.4.6 Origins of key effectors in cherry pathogens

To understand how key effectors identified using GLOOME and BayesTraits have evolved, the nucleotide sequences of each effector family were aligned using ClustalW2 and phylogenies were produced to infer the likely origin of gained effectors. Horizontal gene transfer on plasmids and genomic islands is a key mechanism for effector shuffling within *P. syringae* (Arnold & Jackson 2011). If effectors have been gained via horizontal transfer from another clade within the species complex, the effector gene phylogenies were predicted to be incongruent with the core genome phylogeny. The phylogenies showed that, based on sequence variation, many of the effectors predicted to be important for disease had experienced horizontal transfer between the cherry-infecting clades. Table 4.5 reports those effectors that appear to show horizontal gene transfer between cherry-infecting clades. There has been frequent exchange of effectors between *Psm* R1, R2 and *P.s* pv. *avii*. Effectors putatively shared between these clades were mostly plasmid-encoded in the PacBio strains and included *hopBB1*, *hopAY1*, *hopBD1*, *hopAF1*, *hopBF1* and *avrD1*. In R1-5244 several of these were encoded on one particular plasmid (Contig 2), whilst in R2-leaf they were distributed across two different plasmids (Contig 3 and 4).

Effector	Cluster	Region	Plasmid location
hopBB1	R1 - R2 – *P.s* pv. *avii*	Plasmid	tig4 (R2), tig65 (R1-5300)
hopAY1[#]	R1 - R2 – *P.s* pv. *avii*	Plasmid	tig1 (R2)
hopBD1	R1 - R2 – *P.s* pv. *avii*	Plasmid	tig2 (R1-5244), tig4 (R2)
hopAF1-1	R1 - R2 – *P.s* pv. *avii*	Plasmid	tig2 (R1-5244), tig3 (R2)
hopBF1	R1 - R2 – *P.s* pv. *avii*	Plasmid	tig2 (R1-5244), tig3 (R2)
avrD1	R1 - R2 – *P.s* pv. *avii*	Plasmid	tig2 (R1-5244)
hopAT1	R1 - R2	Genomic island	-
hopAR1	R2 - Phylogroup 2 *Pss*	Phage	-
avrRps4[#]	R2 – *P.s* pv. *avii*	Next to cluster of mobile elements	-
hopAO1[#]	R2 – *P.s* pv. *avii*	Plasmid	tig1 (R2)
hopAU1	R2 – *P.s* pv. *avii*	Plasmid	tig2 (R2)
hopAV1	R1 – *P.s* pv. *cerasicola*	In region near mobile elements and other effectors	-
hopD1[#]	R2 – *P.s* pv. *avii*	Plasmid	tig2 (R2)
hopE1	R2 – *P.s* pv. *cerasicola*	Genomic island	-
hopO1	R2 – *P.s* pv. *avii*	Next to cluster of mobile elements (next to *hopT1*)	-
hopT1	R2 – *P.s* pv. *avii*	Next to cluster of mobile elements (next to *hopO1*)	-

***Table 4.5: List of putative horizontal gene transfer events that have occurred between* Prunus*-infecting clades within* P. syringae.** *Where the effector gene is present in the PacBio or Minion sequenced strains it is indicated if it is residing on a plasmid. #: Effector gene is disrupted in some way so is labelled as a putative pseudogene.*

Focusing on the horizontally transferred effectors, the genomic regions that contained these genes were then aligned. These alignments and the corresponding phylogenetic trees are presented in Figures S4.5-S4.35. The effector genes were often flanked by mobile elements, which may have aided their insertion into the genome by recombination. Of particular interest is *hopAR1* as it significantly associated with cherry pathogenicity and present in the majority of *Prunus*-infecting strains. This effector is predicted to have been gained in *Psm* R1 and R2 via different phages and in the cherry *Pss* via a genomic island (Figure 4.19). Phylogenetic analysis showed that the *Psm* R2 homologues cluster with cherry *Pss* and other homologues from Phylogroup 2 strains. Aligning the surrounding region showed that *hopAR1* from *Psm* R2 is most similar to the *Pss* strain syr2675 which was isolated from bean, in which the gene is also present within a phage (Figure S4.14). Therefore, there is evidence for a horizontal transfer of *hopAR1* between *Psm* R2 and a member of Phylogroup 2.

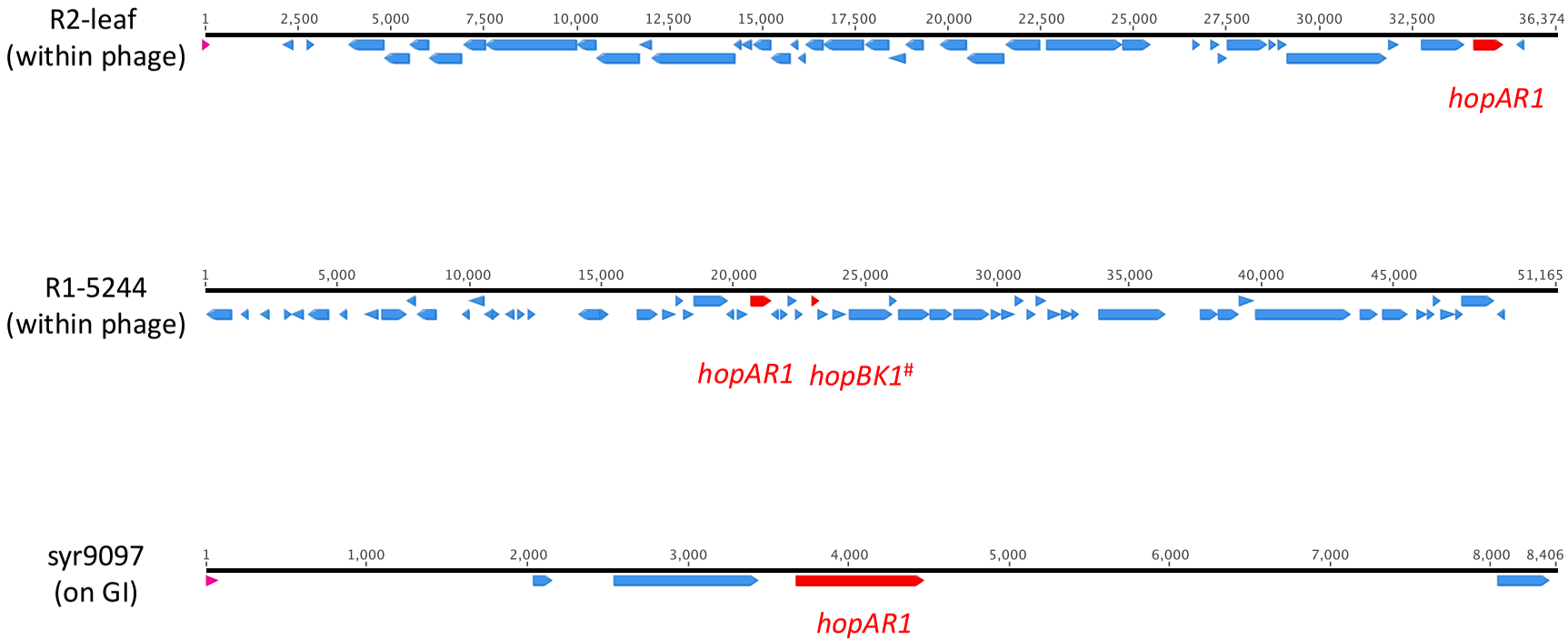


Figure 4.19: Genomic locations of the* hopAR1 *gene in the three PacBio sequenced pathogens. *The gene is located within phage sequences in* Psm *R1 and R2, whereas in syr9097 it is on a genomic island adjacent to a tRNA gene. Effector genes are coloured in red, other CDS in blue and tRNA genes in pink. # Indicates that hopBK1 is a pseudogene in this strain.*

4.4.7 Search for novel effectors

The initial search was based on identifying homologues of previously characterised effectors within *P. syringae*. To search for additional effectors within the sequenced strains that infect *Prunus,* the program Effective was used (Jehl *et al.* 2011). Given a set of protein sequences, Effective predicts candidate effectors based on the presence of eukaryotic-domains and signal peptides. All candidate sequences that scored ≥0.95 were then searched for an upstream *hrp*-box motif. Table 4.6 shows a list of candidate genes per strain with annotations. Many of these candidates were small hypothetical proteins and some appear to be false positives e.g. the identified multidrug and toxin extrusion proteins and transposases are unlikely to function as type III effectors. The fact that few new candidate effector genes were identified was not surprising as previous studies have carried out exhaustive searches of diverse strains of *P. syringae* to build the database of known effectors. It seems likely that much of the effector diversity has already been sampled (Baltrus *et al.* 2011).

Table 4.6 (overleaf): Novel candidate effector genes in sequenced Prunus strains *identified through a pipeline involving the program Effective and identification of the* hrp-*box motif. Annotations and domain names are presented where applicable.*

Strain	Protein	*hrp* box motif	Annotation	Homologs in	Domains
9646	peg.3383	GGAACCCACAAGCCCTGATGACCACATA	multidrug and toxin extrusion (MATE)	Various	Many
5244	peg.4522	TGAACCATCGGTGCACTCACAGCCAGCAA	sensor histidine kinase	Various	pfam03543
5300	peg.4628	GGAACCCATCAGCACAAGCGCACCACGA	macrolide-efflux	Various	cd06174
5300	peg.5268	GGAACCGAAACCGTAGGATTAGCCACTTA	mobile element transposase	varous	
9326	peg.4808	GGAACCCATCAGCACAAGCGCACCACGAA	macrolide-efflux	Various	cd06174
9326	peg.5576	GGAACCGAAACCGTAGGATTAGCCACTTA	mobile element transposase	Various	
9657	peg.644	GGAACCAAATATGTAGTTATGGTCACTCA	hypothetical	Various	None
9657	peg.4824	GGAACCCATCAGCACAAGCGCACCACGAA	macrolide-efflux	Various	cd06174
9657	peg.5554	GGAACCGAAACCGTAGGATTAGCCACTTA	mobile element transposase	Various	
9629	peg.634	GGAACCAAATATGTAGTTATGGTCACTCA	hypothetical	Various	None
9629	peg.4807	GGAACCCATCAGCACAAGCGCACCACGAA	macrolide-efflux	Various	cd06174
9629	peg.5574	GGAACCGAAACCGTAGGATTAGCCACTTA	mobile element transposase	Various	
Leaf	peg.1293	GGAACTGTTTTTCAGCATAAACCACATA	hypothetical protein	Ps *actinidiae*	PRK14954
Leaf	peg.2077	TGAACCATCGGTGCACTCACAGCCAGCAA	sensor DegS	Various	Many
sc214	peg.2390	TGAACCATCGGTGCACTCACAGCCAGCAA	sensor DegS	Various	Many
sc214	peg.127	GGAACTGTTTTTCAGCATAAACCACATA	hypothetical protein	Ps *actinidiae*	PRK14954
5255	peg.3142	TGAACCATCGGTGCACTCACAGCCAGCAA	sensor DegS	Various	Many
5255	peg.1922	GGAACTGTTTTTCAGCATAAACCACATA	hypothetical protein	Ps *actinidiae*	PRK14954
5260	peg.3935	TGAACCATCGGTGCACTCACAGCCAGCAA	sensor DegS	Various	Many
5260	peg.1968	GGAACTGTTTTTCAGCATAAACCACATA	hypothetical protein	Ps *actinidiae*	PRK14954
9643	peg.254	TGAACCATCGGTGCACTCACAGCCAGCAA	sensor DegS	Various	
9097	peg.1709	GGAACCTTGCCCGGCTTCCTGCCACTCA	peptidase	Various	Many
9097	peg.614	TGGAACTATTTTCCACGTTAAGCCACTCA	hypothetical	Ps *aceris*	None
9630	peg.2520	GGAACCTTGCCCGGCTTCCTGCCACTCA	peptidase2C	Various	Many
9654	peg.1754	GGAACCTTGCCCGGCTTCCTGCCACTCA	peptidase2C	Various	Many
9659	peg.1212	GGAACCTTGCCCGGCTTCCTGCCACTCA	peptidase2C	Various	Many
9644	peg.2907	GGAACCTTGCCCGGCTTCCTGCCACTCA	peptidase	Various	Many

4.4.8 Other virulence factors

Other virulence factors within the cherry pathogens that have been shown to be important in *P. syringae* pathogenicity were also identified. Toxins are a major class of secreted molecules that *P. syringae* uses to subvert host responses and cause necrosis (Bender *et al.* 1999). tBLASTn was used to search *P. syringae* genomes for known toxin synthesis genes. If all known biosynthesis genes were found (>70% QL and >40% ID), it was concluded that the toxin was putatively produced. Figure 4.20 shows that strains of *Pss* encoded genes for between two and four toxins, whereas *Psm* R1 strains encoded for only one (coronatine) and R2 did not possess any known toxin genes. All pathogenic *Psm* R1 strains had the biosynthesis gene clusters for coronatine, indicating that coronatine may be important for pathogenicity. The two gene clusters involved in coronatine biosynthesis were plasmid-borne in the PacBio strain R1-5244.

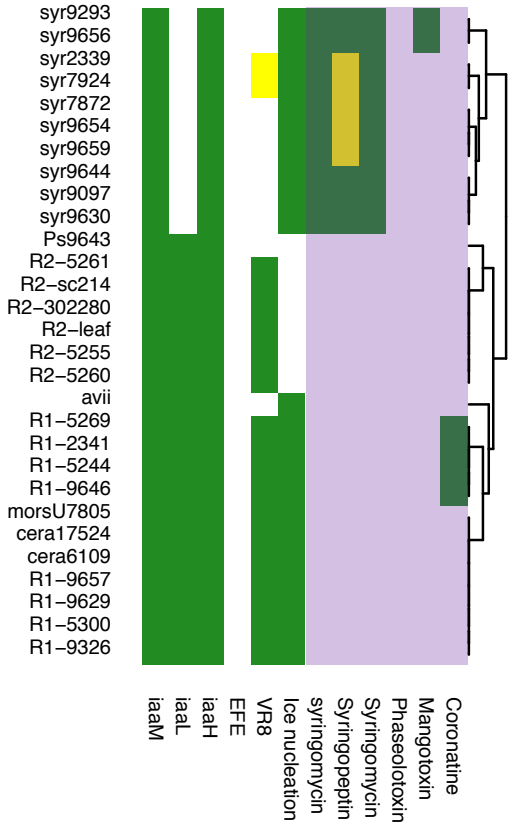


Figure 4.20: Heatmap showing the presence and absence of other virulence factors in the* Prunus*-infecting strains. *The green squares indicate presence of a full-length homologue of all genes involved whereas yellow squares indicate that at least one of the genes involved is disrupted or truncated in some way. Phytotoxin biosynthesis clusters are in the purple box. The dendrogram clustered strains by similarity in virulence gene presence and absence. It should be noted that in Psm R2-302280 has a contig break in the* iaaH *gene but it is assumed to be full-length in the heatmap.*

The genomes were also searched for homologues of genes involved in auxin production and inactivation. All strains possessed genes for auxin production (*iaaM* and *iaaH*), whilst *Pss* strains all lacked the auxin inactivation gene *iaaL*.

Pss strains all possessed at least two of the non-ribosomal peptide toxin gene clusters (syringomycin, syringolin and syringopeptin), with several strains possessing all three. Two divergent *Pss* strains, which belong to a separate clade within phylogroup 2, also possessed the biosynthesis genes for the antimetabolite mangotoxin. Further analysis of toxin presence across phylogroup 2 (Figure 4.21) showed that mangotoxin production is mostly limited to this particular clade of *P. syringae*. It has been previously reported that a negative correlation exists between the number of effector genes and toxin biosynthesis clusters (Hockett *et al.* 2014). To determine if this was true in *Prunus*-infecting strains, the number of effector genes was plotted against the number of toxin clusters. Figure 4.22 confirms that a negative correlation exists between effectors and toxins (p=-0.91), with *Pss* strains possessing only 11-16 effectors and 3-4 toxins. The other clades possessed 23-33 effectors and apart from *Psm* R1 coronatine-producing strains, all others possessed no toxin encoding clusters.

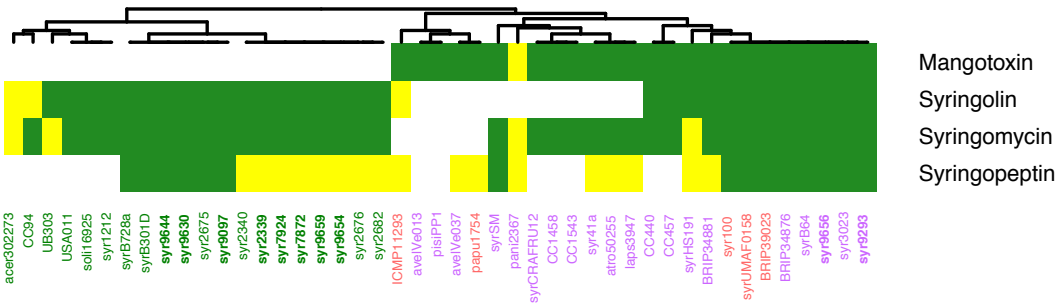


Figure 4.21: Heatmap showing the presence and absence of several phytotoxin biosynthesis clusters in phylogroup 2 strains*. The green squares indicate presence of a full-length homologue of all genes involved whereas yellow squares indicate that at least one of the genes involved is disrupted or truncated in some way. Strain names are coloured dependent upon the clade within phylogroup 2 and Prunus-infecting strains are in bold.*

Furthermore, genes involved in the degradation of aromatic compounds that have been suggested to be important for woody pathogens were identified in the sequenced *Prunus* strains (Rodríguez-Palenzuela *et al.* 2010; Bartoli *et al.* 2015; Nowell *et al.* 2016; Caballo-Ponce *et al.* 2016). A 15kb gene cluster named WHOP/VR8 that contains genes involved in anthranilate and catechol degradation was found in all strains of *Psm* R1 and R2 as well as *P.s* pv. *cerasicola* and *P.s* pv. *morsprunorum* U7805

isolated from apricot (Figure 4.20). However, most *Pss* stains contained no homologues of these genes, despite their ability to produce severe cankers in cherry. Although, two cherry *Pss* strains 2339 and 7924 possessed the catechol *catBCA* cluster.

Finally, the genomes were searched for a gene cluster involved in ice nucleation as this has been hypothesised to be important in the virulence of *P. syringae* as it allows the bacteria to create wounds through frost damage, providing an entry route (Lamichhane *et al.* 2014). Interestingly, all *Psm* R1, *Pss*, *P.s* pv. *cerasicola* and *P.s* pv. *avii* strains possessed genes involved in ice nucleation (*inaA*, *inaK*, *inaV* and *inaZ*), whereas *Psm* R2 lacked all the genes except *inaA* (Figure 4.20).

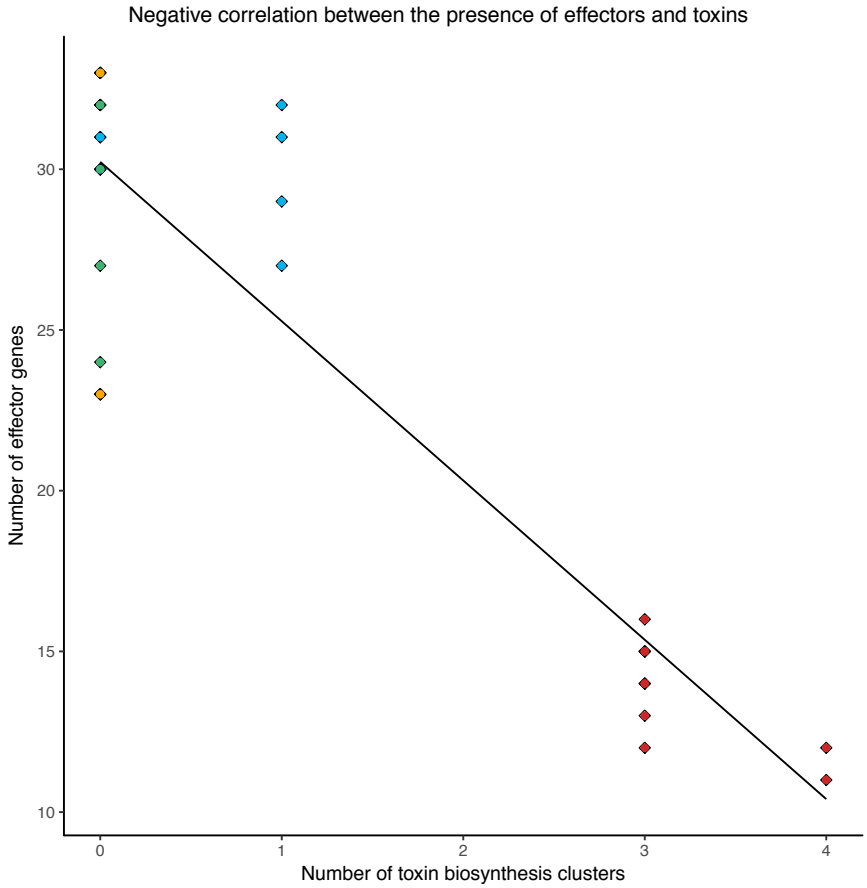


Figure 4.22: Scatterplot with line of correlation of the number of putative effector genes and toxin biosynthesis clusters for each* Prunus*-infecting strain*.*** *Values are coloured dependent on clade of* P. syringae *(*Psm *R1 in blue,* Psm *R2 in green,* Pss *in red and other cherry pathogens in orange).*

4.4.9 Orthology analysis

To analyse the whole set of potential genes within each genome that might contribute to pathogenicity, the program OrthoMCL was utilized, which identifies shared and unique genes based on orthology (Li *et al.* 2003).

Running OrthoMCL on the 104 genomes identified an overall pan genome of 10,998 orthogroups. Of these orthogroups, 7499 had two or more members whilst 3499 were strain-specific genes. Of the 7499 shared orthogroups, 2901 were present in multiple copies within some genomes (paralogous copies), whilst 4598 were in single copy. 2006 orthogroups were found to be present in all 104 strains and represent the 'core' *P. syringae* genome.

4.4.9.1 Identification of orthogroups on branches leading to cherry pathogens

To identify orthogroups that have been gained during the specialisation of cherry pathogens, GLOOME was ran on the orthogroup matrix using both the core genome and RealPhy phylogenies. Many of the gained orthogroups contained genes which were either hypothetical proteins or those associated with mobile elements e.g. plasmid structural genes, phage-associated genes and toxin-antitoxin systems. The gain of other orthogroups is presented in Figures 4.23 and 4.24. In all cherry-infecting clades there were gains of genes in metabolic subsystems. These included genes involved in fatty acid metabolism, degradation of aromatic compounds, alginate biosynthesis, pyruvate metabolism, gentisate degradation, urea degradation and sugar utilisation. Genes involved in resistance to heavy metals such as copper and arsenic were also gained in multiple clades. Furthermore, multiple cherry-infecting clades had gained genes involved in iron uptake, as well as the production of peptidoglycan, exopolysaccharide and bacteriocins. On the branch leading to *Psm* R1 cherry pathogens, gain of coronatine biosynthesis genes was observed.

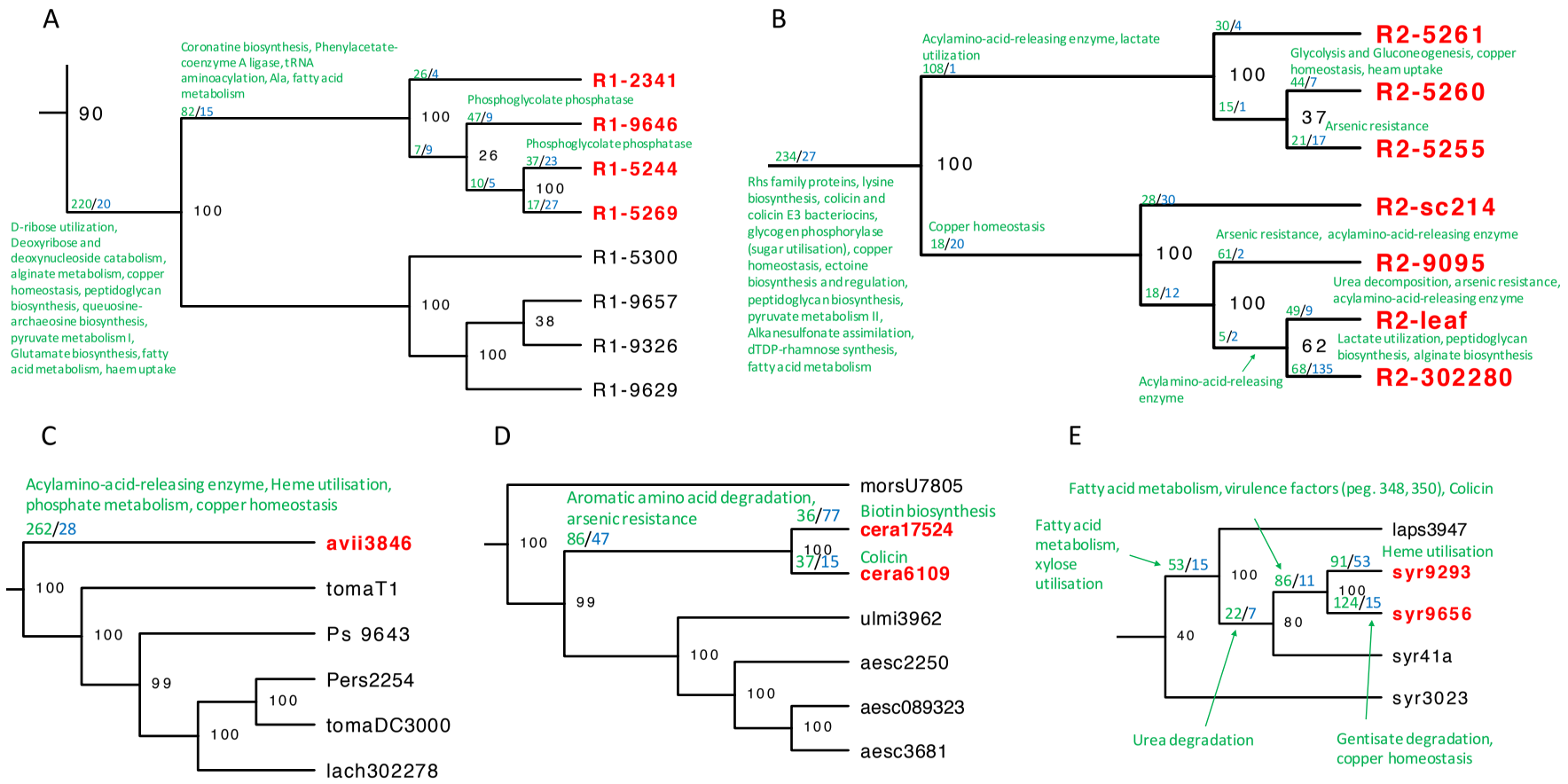


Figure 4.23: Phylogenetic trees with the gain and loss of particular orthogroups occurring on each branch leading to cherry pathogenic strains. This was based on the GLOOME analysis of the core genome phylogeny. *Gained genes are coloured in green. Cherry pathogenic strains are highlighted in red. The predicted number of gains and losses are presented for each branch with gains in green and losses in blue. A:* Psm *R1, B:* Psm *R2, C:* P.s *pv.* avii, *D:* P.s *pv.* cerasicola, *E: divergent* Pss *strains.*

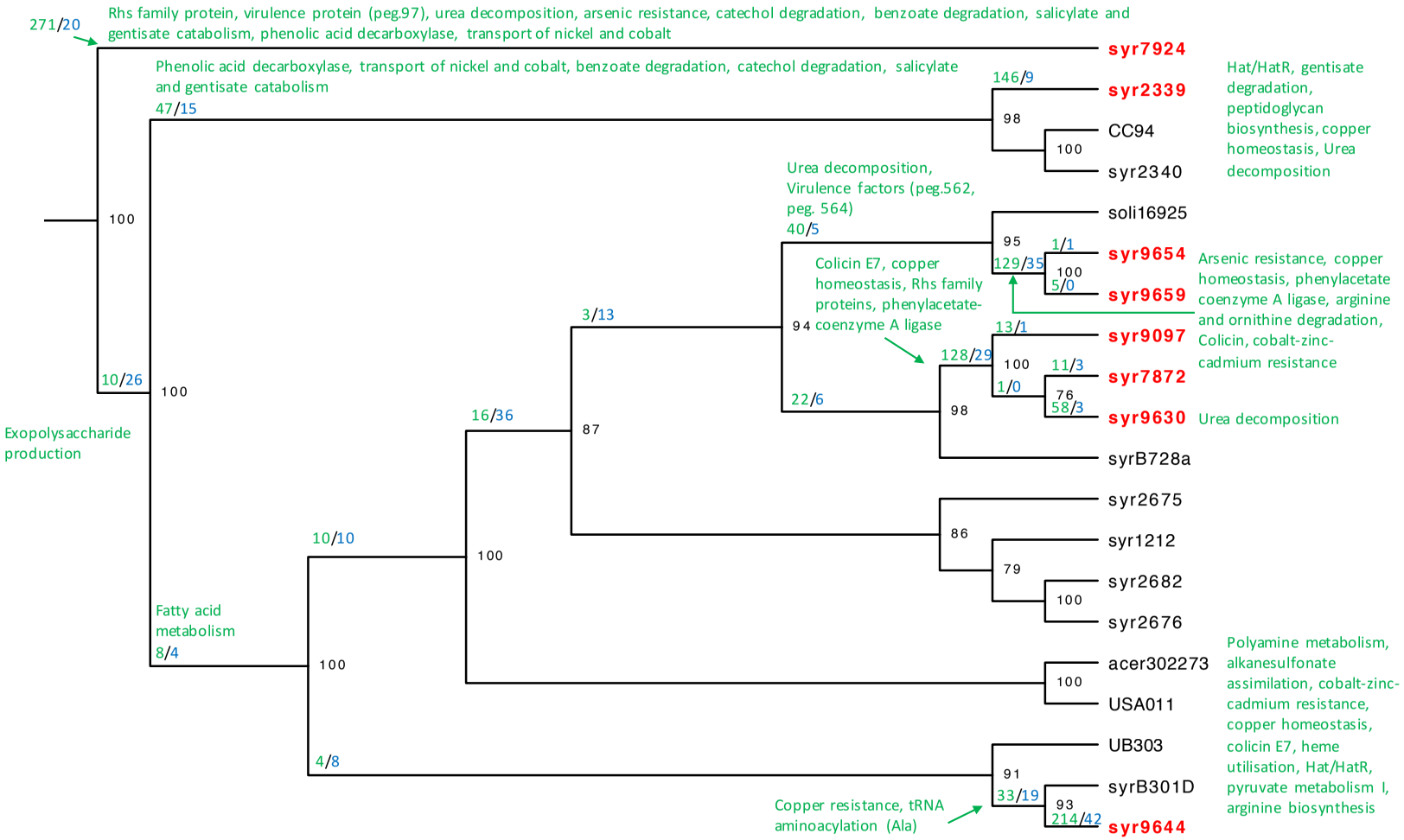


Figure 4.24: Phylogenetic trees with the gain and loss of particular orthogroups occurring on each branch leading to cherry pathogenic* Pss *strains, using the GLOOME analysis of the core genome phylogeny. *Gained genes are coloured in green. Cherry pathogenic strains are highlighted in red. The predicted number of gains and losses are presented for each branch with gains in green and losses in blue.*

The analysis using the RealPhy phylogeny (Figures 4.25 and 4.26) supported the same gains of these metabolic and virulence subsystems in each cherry clade, however the exact numbers and placement of orthogroup gains varied due to differences in tree topology. This was particularly apparent in the *Pss* clade where clustering of cherry-infecting strains led to single gains in particular subsystems on deeper branches rather than convergent gains on the branches leading to each cherry-infecting strain.

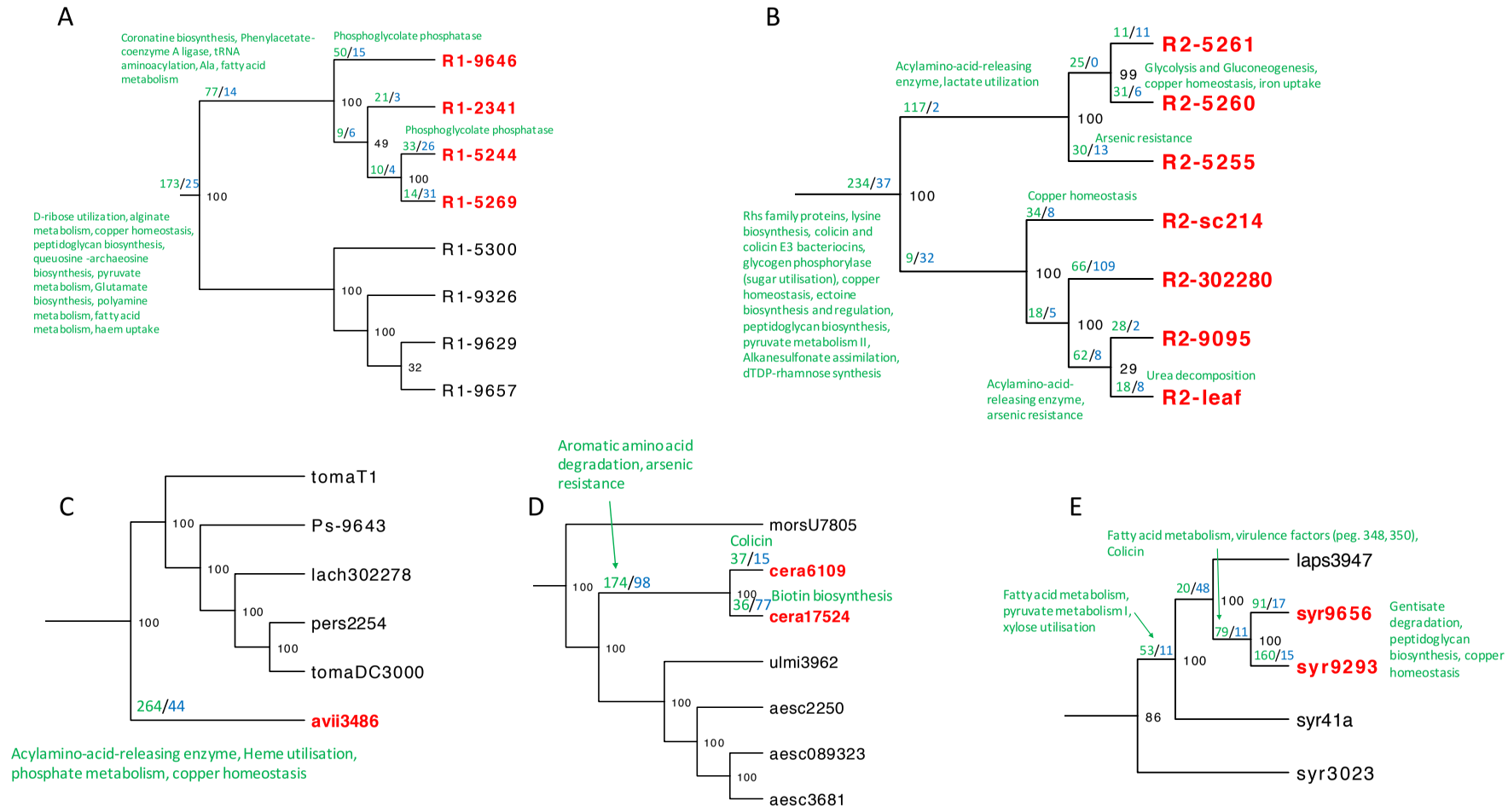


Figure 4.25: Phylogenetic trees with the gain and loss of particular orthogroups occurring on each branch leading to cherry pathogenic strains, using the GLOOME analysis of the RealPhy phylogeny. *Gained genes are coloured in green. Cherry pathogenic strains are highlighted in red. The predicted number of gains and losses are presented for each branch with gains in green and losses in blue. A:* Psm *R1, B:* Psm *R2, C:* P.s *pv.* avii*, D:* P.s *pv.* cerasicola*, E: divergent* Pss *strains.*

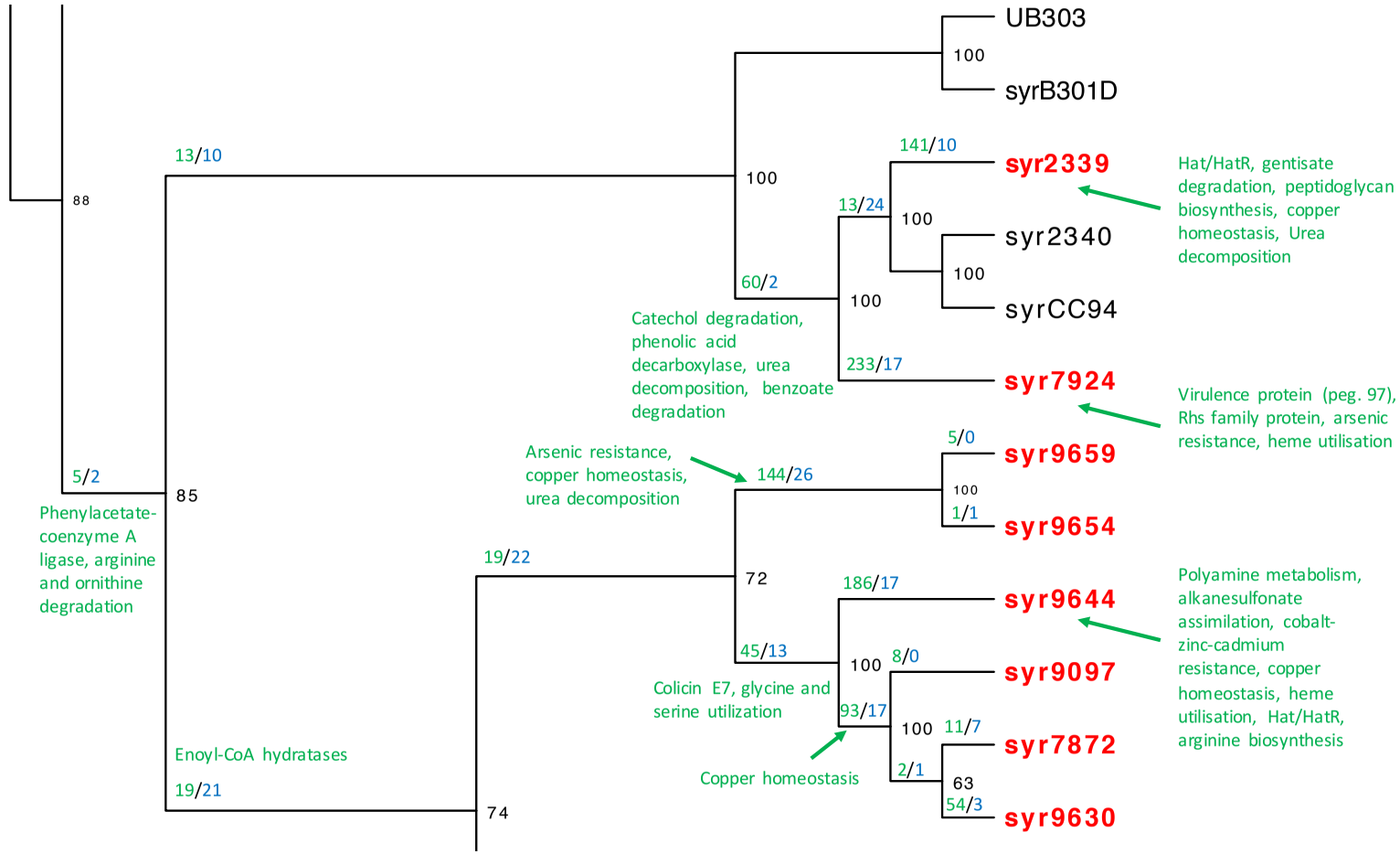


Figure 4.26: Phylogenetic trees with the gain and loss of particular orthogroups occurring on each branch leading to cherry pathogenic Pss strains. This was based on the GLOOME analysis of the Realphy phylogeny. *Gained genes are coloured in green. Cherry pathogenic strains are highlighted in red. The predicted number of gains and losses are presented for each branch with gains in green and losses in blue.*

4.4.10 Mobility of virulence components

Many of the factors described above are mobilised between bacteria on genomic islands. Genomic islands were identified using IslandFinder for the three PacBio-sequenced strains, the islands and their features are documented in Tables 4.7-4.9. Effectors and other virulence genes were identified upon islands in the strains. *Psm* R1 possessed 36 putative islands, notably containing the coronatine biosynthesis cluster, *hopF3*, *hopA1*, *hopAT1*, *hopBL2*, *hopAO2* and *hopAY1*. In *Psm* R2, effector genes located on islands included *hopAT1*, *hopD1*, *hopX1*, *hopAR1*, *hopE1*, *hopAE1*, *avrB2*, *hopAZ1*, *hopAF1* and *hopH1*. Finally, in *Pss hopAR1*, *avrRpm1* and *hopBE1* were found on genomic islands.

Genomic island	Contig no	Start	End	Length	Contig type	Annotations present
GI1	3	141566	168858	27292	Plasmid	**Coronatine biosynthesis gene clusters,** malonyl CoA-acyl carrier transacylase, hypothetical coding sequence, non-ribosomal peptide synthase, long-chain-fatty-acid--CoA ligase, crotonyl-CoA reductase, 3-oxoacyl-[acyl-carrier-protein] synthase, KASII (EC 2.3.1.179), mobile element, *cmaB*, 3-hydroxyacyl-[acyl-carrier-protein] dehydratase, *fabZ* form (EC 4.2.1.59)
GI2	0	230388	253754	23366	Chromosome	D-3-phosphoglycerate dehydrogenase, *traY*, *terA*, mobile element genes, ISPsy4, hypothetical coding sequences, *tnpA* transposase
GI3	6	12090	28296	16206	Plasmid	Type IV secretion complex, *parA*, resolvase, hypothetical coding sequences
GI4	0	3422270	3438178	15908	Chromosome	Hypothetical coding sequences, phage genes, 3'-phosphoadenosine 5'-phosphosulfate sulfotransferase (PAPS reductase)/FAD synthetase and related enzymes, transcriptional regulator, *luxR* family gene
GI5	0	5756201	5771435	15234	Chromosome	Hypothetical coding sequences, kinase/phosphorylase, nitrogen regulation, integrase, mobile element gene, thymidylate synthase (EC 2.1.1.45), Prolipoprotein diacylglyceryl transferase (EC 2.4.99.-), ISPsy4
GI6	0	161692	173708	12016	Chromosome	***hopF3***, hypothetical coding sequences, 2-Oxobutyrate oxidase, mobile element genes, phytanoyl-CoA dioxygenase, *umuC* DNA polymerase, peptide deformylase
GI7	0	5874146	5884973	10827	Chromosome	*clpB*, *impG*, *impC*, *impH*, hypothetical coding sequences, *impB*, *impF*
GI8	4	48480	59260	10780	Plasmid	***hopA1*****,** *schA*, hypothetical coding sequences, glycogen phosphorylase, resolvase, *parA*, mobile element genes, growth inhibitor, traR
GI9	0	1940026	1950562	10536	Chromosome	Aspartyl-tRNA synthetase, *tolB*, *tolA*, *ruvB*, *yebC*, *tolQ*, *ruvA*, *ruvC*, *ompA*, 4-hydroxybenzoyl-CoA thioesterase family active site, *tolR*, hypothetical coding sequences
GI10	0	196195	206565	10370	Chromosome	Aldehyde dehydrogenase, hypothetical coding sequences, D-alanyl-D-alanine carboxypeptidase, integrase
GI11	3	113418	123596	10178	Plasmid	Chloride channel protein, enolase, C4-dicarboxylate transporter/malic acid transport, *lysR* family transcriptional regulator, transposase, inorganic

Genomic island	Contig no	Start	End	Length	Contig type	Annotations present
						pyrophosphatase, arsenical-resistance *acr3*, hypothetical coding sequences
GI12	0	1034313	1043742	9429	Chromosome	*snf2* DNA/RNA helicase, Xaml DNA methyltransferase, plasmid-related protein, hypothetical coding sequences
GI13	0	1045112	1053907	8795	Chromosome	DNA topoisomerase III, abortive infection bacteriophage resistance protein, *pilL* from PFGI-1-like cluster, hypothetical coding sequences
GI14	0	1004317	1012695	8378	Chromosome	***hopAT1***, membrane-bound lytic murein transglycosylase B precursor, *parA*, hypothetical coding sequences, integral membrane protein, *traR*
GI15	0	3424575	3432915	8340	Chromosome	hypothetical coding sequences, prophage repressor, phage genes, *luxR*
GI16	0	5392587	5399641	7054	Chromosome	***hopBL2***, D-3-phosphoglycerate dehydrogenase, integrase, hypothetical coding sequences, mobile element genes, single-stranded DNA-binding protein, recombinase
GI17	0	2931170	2937567	6397	Chromosome	hypothetical coding sequences, phage genes, Clp proteases, FKBP-type peptidyl-prolyl cis-trans isomerases 1
GI18	0	125459	131647	6188	Chromosome	Permease, *lysR* transcriptional regulation, hypothetical coding sequences, *impD*
GI19	0	2379900	2385665	5765	Chromosome	Histidine kinase/response regulator, methyl-accepting chemotaxis protein I (serine chemoreceptor), recombinase, DNA methylase
GI20	0	4830800	4836556	5756	Chromosome	Retron-type RNA-directed DNA polymerase, hypothetical coding sequence, mobile element gene, Ni,Fe-hydrogenase I cytochrome b subunit
GI21	0	3790660	3796397	5737	Chromosome	HlyD family secretion protein, cyclolysin secretion ATP-binding protein , mobile element genes, hypothetical coding sequence
GI22	0	1067116	1072691	5575	Chromosome	tRNA-Pro-TGG, pyruvate/2-oxoglutarate dehydrogenase complex, integrase, DNA helicase in PFGI-1-like cluster, *parA*, hypothetical coding sequence
GI23	0	208395	213891	5496	Chromosome	***hopAO2***, *parA*, *umuC*, hypothetical coding sequences
GI24	0	5902525	5907999	5474	Chromosome	Hypothetical coding sequence, ADP-ribosylglycohydrolase family protein

Genomic island	Contig no	Start	End	Length	Contig type	Annotations present
GI25	0	1485251	1490204	4953	Chromosome	Hypothetical coding sequences
GI26	0	355673	360618	4945	Chromosome	*hopD1*, *hopQ1*, hypothetical coding sequences
GI27	0	6079445	6084266	4821	Chromosome	Tn*7*-like transposition proteins, transposase
GI28	4	9509	14294	4785	Plasmid	Permease, L-2-hydroxyglutarate oxidase, hypothetical coding sequences, *gntR*
GI29	0	2654000	2658702	4702	Chromosome	Hypothetical coding sequences
GI30	0	1376492	1381166	4674	Chromosome	CDP-glucose 4,6-dehydratase, hypothetical coding sequences, Glycosyltransferase, glucose-1-phosphate cytidylyltransferase
GI31	0	176079	180621	4542	Chromosome	***hopAY1***, hypothetic proteins, *secA*, resolvase, *tnpA* transposases, *yafQ* toxin
GI32	0	5378650	5383123	4473	Chromosome	Sensory box/GGDEF family protein, *nagC*, ABC-type phosphate/phosphonate transport system, periplasmic component, hypothetical coding sequences
GI33	6	1	3641	3640	Plasmid	Hypothetical coding sequences, *traL*
GI34	0	2302562	2306880	4318	Chromosome	Homolog of fucose/glucose/galactose permease, mobile element gene, hypothetical coding sequences, *deoR*
GI35	5	6584	10771	4187	Plasmid	*traL*, hypothetical coding sequence
GI36	3	134051	137948	3897	Plasmid	Resolvase, *traR*, hypothetical coding sequences, Helix-turn-helix motif CDS

Table 4.7: Genomic islands predicted in the* Psm *R1 5244. *The island locations, length and gene annotations are presented. Key virulence-associated genes are highlighted in bold.*

Genomic island	Contig no	Start	End	Length	Contig type	Annotations present
GI1	0	673197	735275	62078	Chromosome	lsuRNAs, ssuRNAs, tRNAs, 5S RNA, DNA-directed RNA polymerase subunits, excinuclease ABC subunit A, Type I restriction-modification system, *tefG*, hypothetical coding sequence, integrase, putative transport protein, translocase *secY* subunit, tyrosyl-tRNA synthetase, pantothenate kinase, biotin--protein ligase, bacterioferritin, mobile element genes
GI2	5	68674	102889	34215	Plasmid	*mobA*, hypothetical coding sequences, chloride channel, twitching-motility protein, *traO*, malic acid transport protein, *parA*, mobile element genes, enolase, surface antigen, *parB*, recombinase, LysR family transcriptional regulator, phage genes, transposase, IS1415, inorganic pyrophosphatase, arsenical-resistance *acr3*, *traH*, putative lipoprotein, DNA primase
GI3	0	951560	981100	29540	Chromosome	***hopAT1***, tRNAs, clpB, DNA topoisomerase III, pyruvate/2-oxoglutarate dehydrogenase complex, putative membrane protein, *parA*, DNA/RNA helicases, hypothetical coding sequences, abortive infection bacteriophage resistance protein, *tonB*, *pilL* in PFGI-1-like cluster, *traR*
GI4	0	291081	320543	29462	Chromosome	Type I restriction-modification system, HTH transcriptional regulator, hypothetical coding sequences, UDP-glucose dehydrogenase, mobile element genes, ISPsy4, XRE transcriptional regulator
GI5	0	6092650	6121111	28461	Chromosome	Hypothetical coding sequences, leucyl-tRNA synthetase, phage genes, integrase, mobile element genes, rubredoxin-NAD(+) reductase, NADH oxidase, DNA-binding protein HU-alpha, Rubredoxin protein
GI6	4	46881	70620	23739	Plasmid	***hopD1***, Mobile element genes, iguanylate cyclase/phosphodiesterase, hypothetical coding sequences, facilitator family permease, levansucrase, L-2-hydroxyglutarate oxidase, predicted glycoside hydrolase, ISPsy4, resolvase, *pilT*, *traR*, *gntR*, Helix-turn-helix motif protein, Type II secretory pathway, component *exeA*
GI7	0	4600070	4623091	23021	Chromosome	Rhs family proteins, mannuronan C-5-epimerase, mobile element genes, *madN*, oxidoreductase, *luxR*, 3-phosphoglycerate kinase, Lead, cadmium, zinc and mercury transporting ATPase
GI8	0	2030941	2050309	19368	Chromosome	Exonuclease *sbcC* and *sbcD*, L-arabonate dehydratase, TPR domain protein in aerotolerance operon, *batD*, nitrate/nitrite transporter, 3-

Genomic island	Contig no	Start	End	Length	Contig type	Annotations present
						carboxymuconate cyclase, *batA*, sugar transporter, *moxR* in aerotolerance operon, hypothetical coding sequences, *gntR* family protein
GI9	0	4810809	4828227	17418	Chromosome	UDP-glucose dehydrogenase, mobile element genes, IsPsy4, hypothetical coding sequences, N-acetylglucosamine-1-phosphate uridyltransferase
GI10	6	54527	69520	14993	Plasmid	Type IV secretion system complex, resolvase, hypothetical coding sequences
GI11	0	161395	176891	15496	Chromosome	Hypothetical coding sequences, large exoproteins involved in heme utilization or adhesion, diguanylate cyclase/phosphodiesterase, IsPsy4, mobile element genes
GI12	9	1	14148	14147	Plasmid	***hopX1***, RepA, mobile element genes, recombinase *xerC*, *marR* transcriptional regulator, Sorbitol-6-phosphate 2-dehydrogenase, *tetR* transcriptional regulator, *parA*, hypothetical coding sequences
GI13	9	1829	14148	12319	Plasmid	*repA*, mobile element genes, recombinase *xerC*, *marR* transcriptional regulator, *parA*, hypothetical coding sequences
GI14	0	5927071	5940007	12936	Chromosome	***hopAR1*,** ATPase, mobile element gene, hypothetical coding sequences, ISPsy4, hypothetical antitoxin
GI15	0	3506777	3519594	12817	Chromosome	Hypothetical coding sequences, mobile element genes
GI16	0	1797361	1808954	11593	Chromosome	tRNA, pyruvate/2-oxoglutarate dehydrogenase complex, integrase, *tolB*, *tolA*, hypothetical coding sequences, *motA* channel protein, queuosine biosynthesis proteins, PR repeat containing exported protein, mobile element genes, *tolR*
GI17	0	5051428	5062925	11497	Chromosome	Carbamoyl-phosphate synthase proteins, Cell division protein *ftsH*, *dnaJ*, Dihydropteroate synthase, 4-hydroxy-tetrahydrodipicolinate reductase, heat-shock protein, *greA*, RNA binding protein, hypothetical coding sequence
GI18	0	37921	49115	11194	Chromosome	RHS family protein, mobile element gene, Choline-sulftase, *lysR* transcriptional regulator proteins, hypothetical coding sequences
GI19	0	1819662	1829920	10258	Chromosome	***hopE1***, Secretory lipase precursor, recombinase *xerC*, mobile element gene, putative cytoplasmic protein, hypothetical coding sequences, peptide deformylase

Genomic island	Contig no	Start	End	Length	Contig type	Annotations present
GI20	0	5083008	5092454	9446	Chromosome	Methyl-accepting chemotaxis protein, mobile element gene, transposase, phage protein, integrase, hypothetical coding sequences, terminase
GI21	0	4480273	4488351	8078	Chromosome	Methylase, mobile element gene, ribonucleotide reductase, DNA/RNA helicase
GI22	0	3051175	3059157	7982	Chromosome	Conserved domain protein, D-alanyl-D-alanine carboxypeptidase, *lysR* transcriptional regulator, enoyl-CoA hydratase, hypothetical coding sequences
GI23	0	279507	287312	7805	Chromosome	Anthranilate synthase, *higA* antitoxin protein, HAD-superfamily hydrolase, hypothetical coding sequence, *asnC* transcriptional regulator
GI24	0	5878	13495	7617	Chromosome	DNA gyrase subunit B, hypothetical coding sequences, mobile element genes, *qbdB*, IsPsy4
GI25	0	997205	1004734	7529	Chromosome	***hopAE1***, Hypothetical coding sequences, mobile element genes, recombinase *xerC*, pentapeptide repeat family protein
GI26	0	3275431	3282907	7476	Chromosome	tRNAs, hypothetical coding sequences, nucleoside-diphosphate-sugar epimerase, lead, cadmium, zinc and mercury transporting ATPase
GI27	0	5522592	5529991	7399	Chromosome	Rhs family protein, putative cytoplasmic protein, hypothetical coding sequences
GI28	8	6270	12804	6534	Plasmid	Methyl-accepting chemotaxis protein, resolvase, DNA-damage inducible protein J, hypothetical coding sequence, *rlgA*
GI29	0	3961189	3967267	6078	Chromosome	GGDEF domain protein, *pfl*, mobile element gene, putative cytoplasmic protein, recombinase
GI30	9	11360	17435	6075	Plasmid	***avrB2***, *repA*, resolvase, *parA*, *yafQ* toxin protein, hypothetical coding sequence, mobile element gene
GI31	0	156390	161525	5135	Chromosome	Diguanylate cyclase/phosphodiesterase, alcohol dehydrogenase, mobile element genes, hypothetical coding sequence
GI32	0	2981573	2986639	5066	Chromosome	Phage integrase, mobile element genes, reductase, transcriptional regulator *hxlR*, UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate aminotransferase
GI33	5	24428	29491	5063	Plasmid	***hopAZ1***, mobile element genes, *tetR* transcriptional regulator, hypothetical coding sequences
GI34	8	6270	9478	3208	Plasmid	Resolvase, hypothetical coding sequences, DNA-damage-inducible protein J

Genomic island	Contig no	Start	End	Length	Contig type	Annotations present
GI35	4	73316	77991	4675	Plasmid	Mobile element gene, *impD*, invertase, hypothetical coding sequence
GI36	4	73316	77991	4675	Plasmid	***hopAF1***, mobile element gene, *impD*, invertase, hypothetical coding sequence
GI37	0	982664	987170	4506	Chromosome	***hopH1***, mobile element genes, putative exported protein, hypothetical coding sequence
GI38	0	6006150	6010562	4412	Chromosome	Probable phospholipase protein, Sel1-like repeat proteins, hypothetical coding sequence
GI39	0	1013986	1018362	4376	Chromosome	Permease, hypothetical coding sequence, mobile element genes, dipeptide-binding ABC transporter, periplasmic substrate-binding component
GI40	0	742887	747170	4283	Chromosome	Putative ATP-binding protein, hypothetical coding sequences

Table 4.8: Genomic islands predicted in the* Psm *R2 leaf. *The island locations, length and gene annotations are presented. Key virulence-associated genes are highlighted in bold.*

Genomic island	Contig no.	Start	End	Length	Contig type	Annotations present
GI1	0	3360963	3388897	27934	Chromosome	Mobile element genes, *ccoN*, *ccoO*, cytochrome bd2 subunit I, chloride channel protein, enolase, sulfite oxidase, malic acid transport protein, hypothetical coding sequences, surface antigen, *lsyR* transcriptional regulator, ABC-type amino acid transport periplasmic component, IS1415, inorganic pyrophosphatase, 4-oxalocrotonate tautomerase homolog
GI2	0	4396469	4419939	23470	Chromosome	*sugR*, hypothetical coding sequences, *secA*-related protein, mobile element genes
GI3	0	5627141	5645347	18206	Chromosome	Glutamine synthetase family protein, *oprD*, YD repeat protein, alkanesulfonates transport system proteins, mobile element genes, *tetR* transcriptional regulator, *secB*, tRNA (cytidine(34)-2'-O)-methyltransferase, rhodanese-related sulfurtransferase, Glutaredoxin3, organosulfonate utilization *ssuF*
GI4	0	2084110	2100013	15903	Chromosome	5-methylcytosine-specific restriction related enzyme, hypothetical coding sequences, phage genes, alpha/beta hydrolase family proteins, probable carboxyvinyl-carboxyphosphonate phosphorylmutase, *merR* transcriptional regulator
GI5	0	4088370	4103275	14905	Chromosome	tRNA-Arg-CCT, helicase/SNF2 family domain protein, *kefA* potassium efflux protein, hypothetical coding sequences, Mg/Co/Ni transporter *mgtE*, alginate biosynthesis protein
GI6	0	839431	852254	12823	Chromosome	tRNAs, *clpB*, integrase, Pyruvate/2-oxoglutarate dehydrogenase complex, *umuD*, mobile element genes, integrase
GI7	0	1043986	1055427	11441	Chromosome	Hypothetical coding sequences, enoyl-[acyl-carrier-protein] reductase [NADPH], enoyl-CoA hydratase, putative inner membrane protein
GI8	0	5723018	5733545	10527	Chromosome	EF hand domain protein, *vgrG*, serine/threonine protein kinase, hypothetical coding sequences
GI9	0	1607407	1617756	10349	Chromosome	Serine chemoreceptor protein, tryptophan 2-monooxygenase, Indoleacetamide hydrolase, methylcrotonyl-CoA carboxylase carboxyl transferase subunit, *umuC*, transcriptional regulator CDS
GI10	0	203793	214015	10222	Chromosome	Translation-disabling ACNase *rloC*, recombinase, hypothetical coding sequence
GI11	0	1594825	1604564	9739	Chromosome	Aspartyl-tRNA synthetase, *tolB*, *tolA*, *ruvB*, *yebC*, *tolQ*, *ruvA*, *ruvC*, *ompA*, *tolR*, 4-hydroxybenzoyl-CoA thioesterase family active site

GI12	0	3557992	3567051	9059	Chromosome	Alkaline phosphatase, probable glycosyltransferase
GI13	0	1860783	1869431	8648	Chromosome	ATP-dependent protease, *ppiD*, cell division trigger factor, Clp protease, *hxlR* transcriptional regulator, DNA binding protein HU-beta
GI14	0	2651560	2659530	7970	Chromosome	*impG*, *impA*, *impC*, *impH*, *impB*, hypothetical coding sequence
GI15	0	2491526	2499228	7702	Chromosome	Amidase clustered with urea ABC transporter and nitrile hydratase functions, *urtA*, *urtC*, *urtB*, *urtD*, DNA-binding heavy metal response regulator
GI16	0	1423031	1430702	7671	Chromosome	*norR*, flavohemoprotein, enoyl-[acyl-carrier-protein] reductase, transmembrane amino acid efflux protein, *yrdN*, hypothetical coding sequences
GI17	0	113328	120966	7638	Chromosome	ATPase involved in DNA repair, orphan protein, mobile element gene, hypothetical coding sequence
GI18	0	4873088	4880595	7507	Chromosome	Probable insecticidal toxin
GI19	0	2663735	2670726	6991	Chromosome	*icmF*-related protein, *impJ*, *impK*/*ompA*, GNAT family acetyltransferase, type IV secretion lipoprotein *vasD*, hypothetical coding sequence
GI20	0	860694	867101	6407	Chromosome	Long-chain-fatty-acid--CoA ligase, transmembrane efflux protein, hypothetical coding sequence
GI21	0	2647951	2653887	5936	Chromosome	membrane protein, *vGrG*, *impA*, Hcp secreted protein, hypothetical coding sequences
GI22	0	5375902	5380909	5007	Chromosome	Hypothetical coding sequences, peptide methionine sulfoxide reductase *msrA*
GI23	0	5684686	5689649	4963	Chromosome	***hopAR1***, toxin-antitoxin system, hypothetical coding sequences
GI24	0	2131656	2136485	4829	Chromosome	DNA recombinase, adenine-specific methyltransferase, lytic enzyme
GI25	0	36847	41401	4554	Chromosome	Rhs family protein, wall-associated protein
GI26	0	853909	858460	4551	Chromosome	***avrRpm1***, enoyl-CoA hydratase, mobile element gene, peptidase, hypothetical coding sequence
GI27	0	3301015	3305482	4467	Chromosome	Fe transport outer membrane receptor protein, amino acid ABC transporter permease protein, glutamate transport ATP-binding protein, hypothetical coding sequence
GI28	0	5051971	5056389	4418	Chromosome	Kinesin-related protein K4, helicase *hrpB*
GI29	0	1382245	1386646	4401	Chromosome	Chromosome segregation ATPase, hypothetical coding sequences
GI30	0	4478399	4482775	4376	Chromosome	***hopBE1***, hypothetical coding sequences

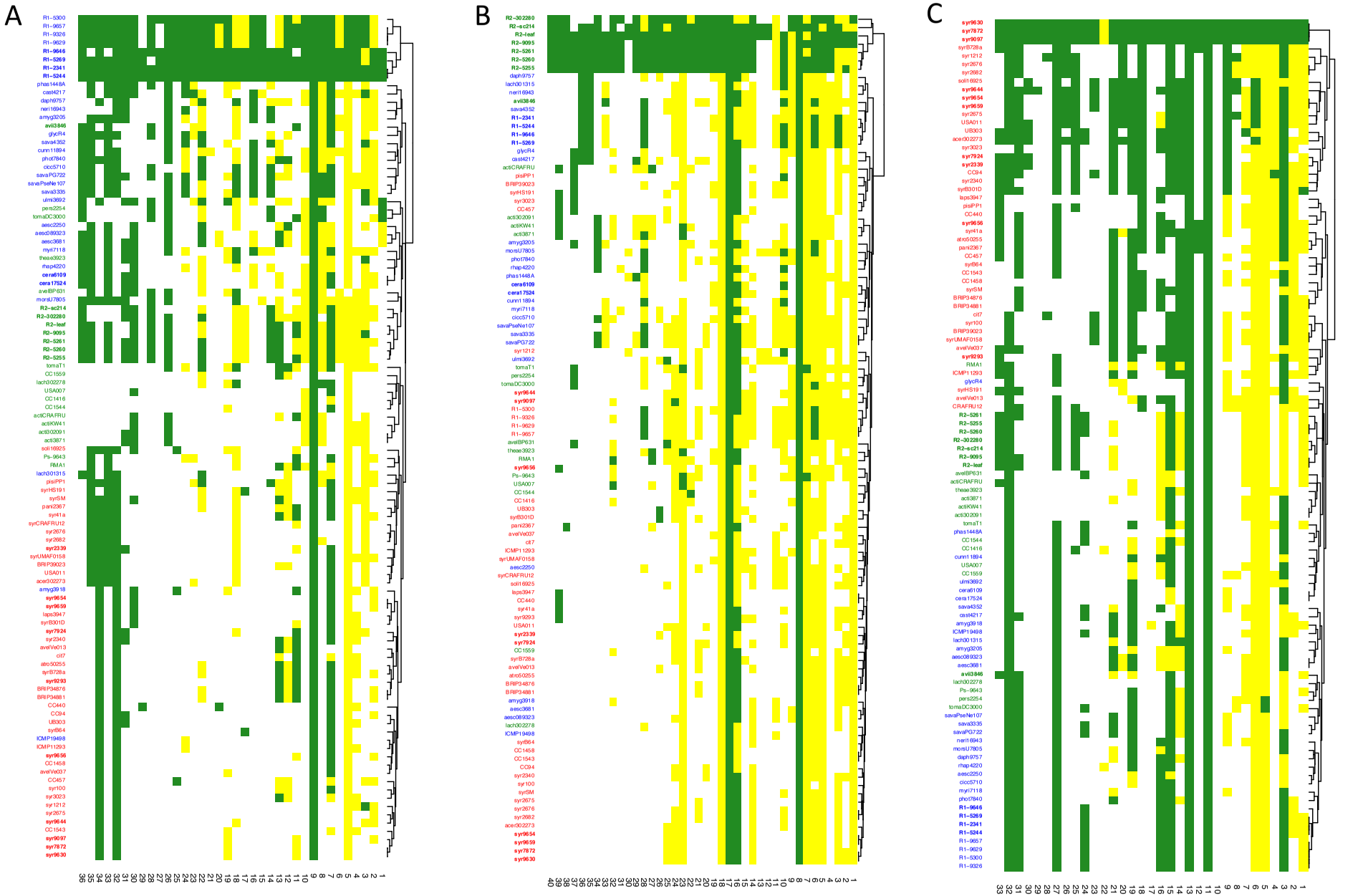
GI31	0	1330595	1334830	4235	Chromosome	OPT oligopeptide transporter, mobile element gene, Glutathione S-transferase, hypothetical coding sequences
GI32	0	4892211	4896412	4201	Chromosome	Peptide ABC transporter, *oppB* oligopeptide transport permease, Putative glutathione transporter, dipeptide transport system permease *dppC*, transcriptional regulator *ahyR*/*asaR* family, hypothetical coding sequence
GI33	0	794252	798297	4045	Chromosome	Hypothetical coding sequences, *muT*

Table 4.9: Genomic islands predicted in the* Pss *9097. *The island locations, length and gene annotations are presented. Key virulence-associated genes are highlighted in bold.*

The set of genomic islands found in these three strains were then searched for within other *P. syringae* to try and identify potential sources of transfer and see how common each island was within the species complex. The genomic islands queried were quite large, with lengths often exceeding 10kb, with the largest being 62kb in *Pss*. However, the set of query genomes were mostly incomplete and highly fragmented, meaning that a negative result could arise if the genomic island is cut off at the end of a contig. Therefore, each genomic island was split into 5000bp segments and BLASTn was used to search for each segment in the query genomes. Figure 4.27 shows heatmaps of putative genomic island presence in all strains used in this study.

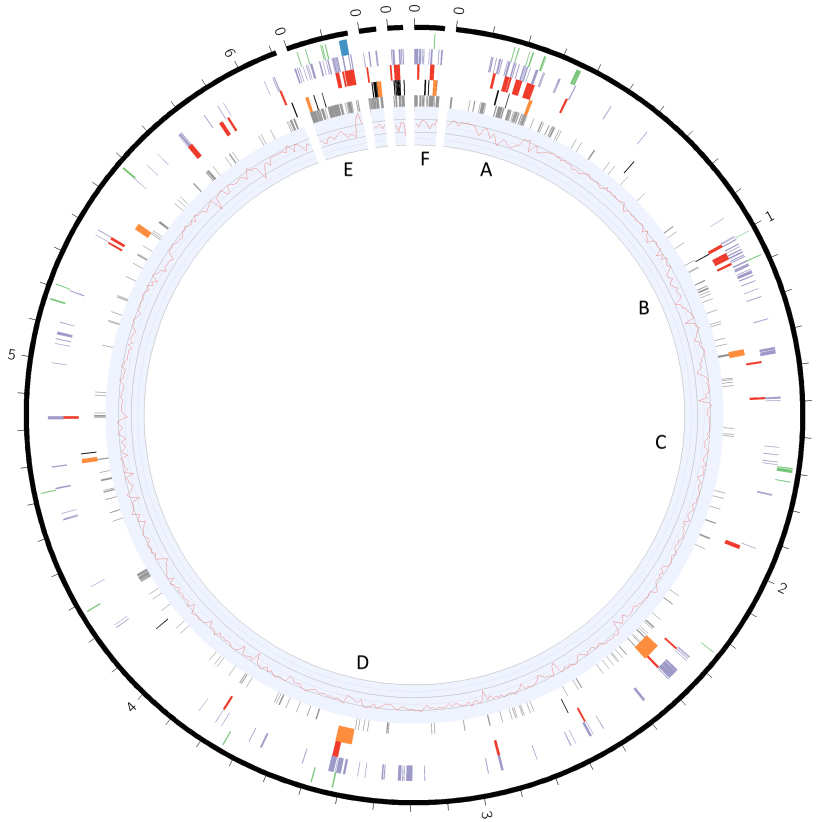
Within *Psm* R1, several islands were only present in strains pathogenic to cherry. These included three islands containing the coronatine biosynthesis cluster (GI1), *hopF3* (GI6) and *hopAT1* (GI14), indicating that these virulence genes have been gained on islands only in the pathogenic members of this clade. Most islands in *Psm* R1 produced significant hits in genomes across the *P. syringae* complex, particularly in phylogroups 1 and 3. In *Psm* R2, several islands were shared with other *Prunus*-infecting clades that contained effectors such as *hopAF1, hopAT1* and *hopD1*. Again, the heatmaps revealed that these islands were shared within members of Phylogroup 1 and 3, suggesting that mobilisation is occurring frequently between these two phylogroups. Finally, although most islands identified in *Pss* 9097 were commonly found across the species complex, several were only found in Phylogroup 2 strains, such as those containing *hopBE1*, *hopAR1* and *avrRpm2*, indicating that *Prunus* strains have gained these islands from exchange with other members of this phylogroup.

Figure 4.27 (overleaf): Heatmap showing the presence and absence of the genomic islands identified in the PacBio sequenced pathogens of (A)* Psm *R1, (B)* Psm *R2 and (C)* Pss *across the* P. syringae *complex. *To identify if the genomic islands were present they were split into 5000bp chunks and BLASTn was used to search each genome. The green squares indicate that all 5000bp sections were found in the query genome with at least 30% of the query sequence length covered. Yellow squares indicate that not all regions of the genomic island were found, meaning only parts of this island are present. The names of strains that infect Prunus are in bold. Each phylogroup is coloured with phylogroup 1 in green, 2 in red and 3 in blue.*


A
B
C

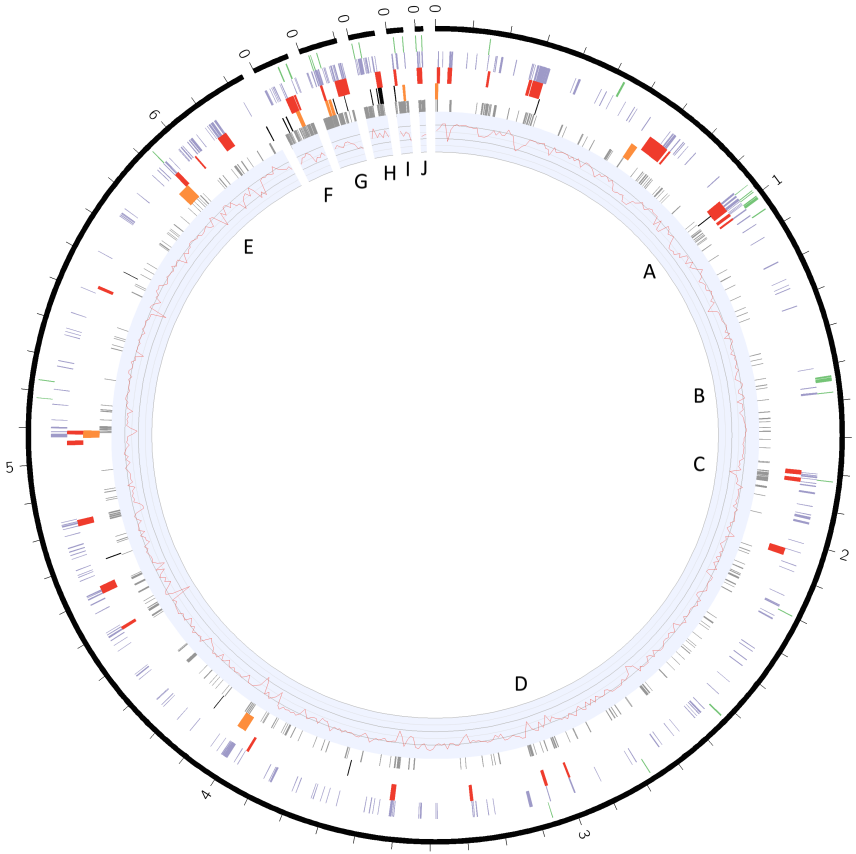
4.4.11 Genomic architecture of virulence-associated genes

To examine how virulence-associated genes clustered within each genome, a program called Circos was used (Krzywinski 2009). Given the location of particular genes Circos makes a plot of the genome. Figures 4.28, 4.29 and 4.30 show the Circos plots for the three PacBio sequenced cherry pathogens. Effector and toxin gene locations were plotted as well as genes belonging to orthogroups that have been gained in cherry-infecting lineages as identified using GLOOME. Genomic islands and mobile elements including phages and plasmid structural genes were plotted for comparison. Effectors and lineage-specific genes were often clustered together throughout the genomes and often present within genomic islands, which may contain phages and other mobile elements. In the two *Psm* races plasmids were enriched in lineage-specific genes and effectors. The key effectors identified as being important for disease (*hopAR1*, *hopBF1*, *hopBD1*, *hopBB1* and *avrRpm2*) were located in highly mobile regions of the genome in all three strains.



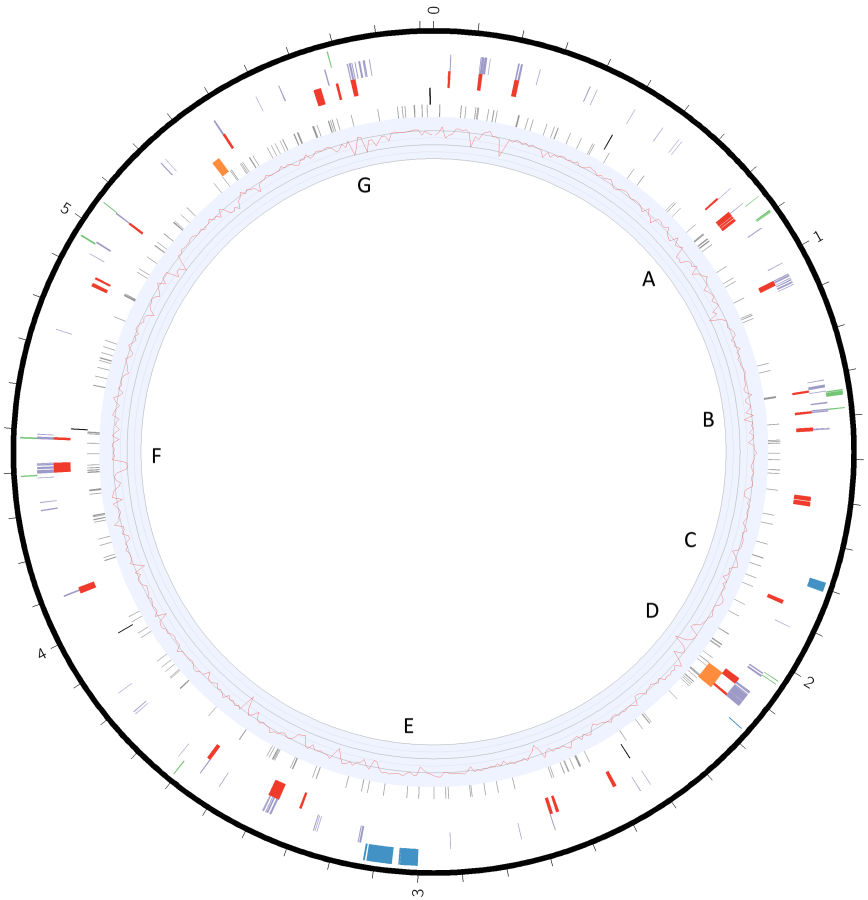
A. *hopF3, hopAY, hopAO2, hopAV, hopBL2, hopF4*
B. *hopAT1, hopZ4*
C. CEL
D. *hopAR1*
E. *avrRpm2, hopBD, hopBF, avrD,* coronatine biosynthesis
F. *hopA1*

Figure 4.28: Circular visualisation of the genome of Psm R1-5244 produced using Circos. *The contigs corresponding to chromosome and four plasmids are on the outer circle. Going inwards, the layers correspond to 1: Effector genes (green) and phytotoxin genes (blue). 2: Genes identified using GLOOME to be gained on the phylogenetic branches leading to this strain. 3: Genomic islands identified using IslandViewer3. 4: Plasmid-structural genes and phage genes identified using PHAST. 5: Insertion sequences identified using ISFinder. 6: GC content, with the middle line indicating 50%.*



A. *hopH1, hopAT1, hopW1, hopR1, hopF4, hopAH1, hopAI1*
B. CEL
C. *hopE1*
D. *avrRps4*
E. *hopAR1*
F. *hopD1, hopAU*
G. *hopAO1, hopAZ1, hopAY*
H. *hopBF1, hopAF1*
I. *hopBD1, hopBB1*
J. *hopX1, avrB2*

Figure 4.29: Circular visualisation of the genome of* Psm *R2-leaf produced using Circos. *The contigs corresponding to chromosome and five plasmids are on the outer circle. Going inwards, the layers correspond to 1: Effector genes (green) and phytotoxin genes (blue). 2: Genes identified using GLOOME to be gained on the phylogenetic branches leading to this strain. 3: Genomic islands identified using IslandViewer3. 4: Plasmid-structural genes and phage genes identified using PHAST. 5: Insertion sequences identified using ISFinder. 6: GC content, with the middle line indicating 50%.*



A. *avrRpm1, hopAG1, hopAI1*
B. CEL
C. Syringolin
D. *hopH1, hopAW1*
E. Syringomycin, syringopeptin
F. *hopAF1, hopBE1*
G. *hopAR1*

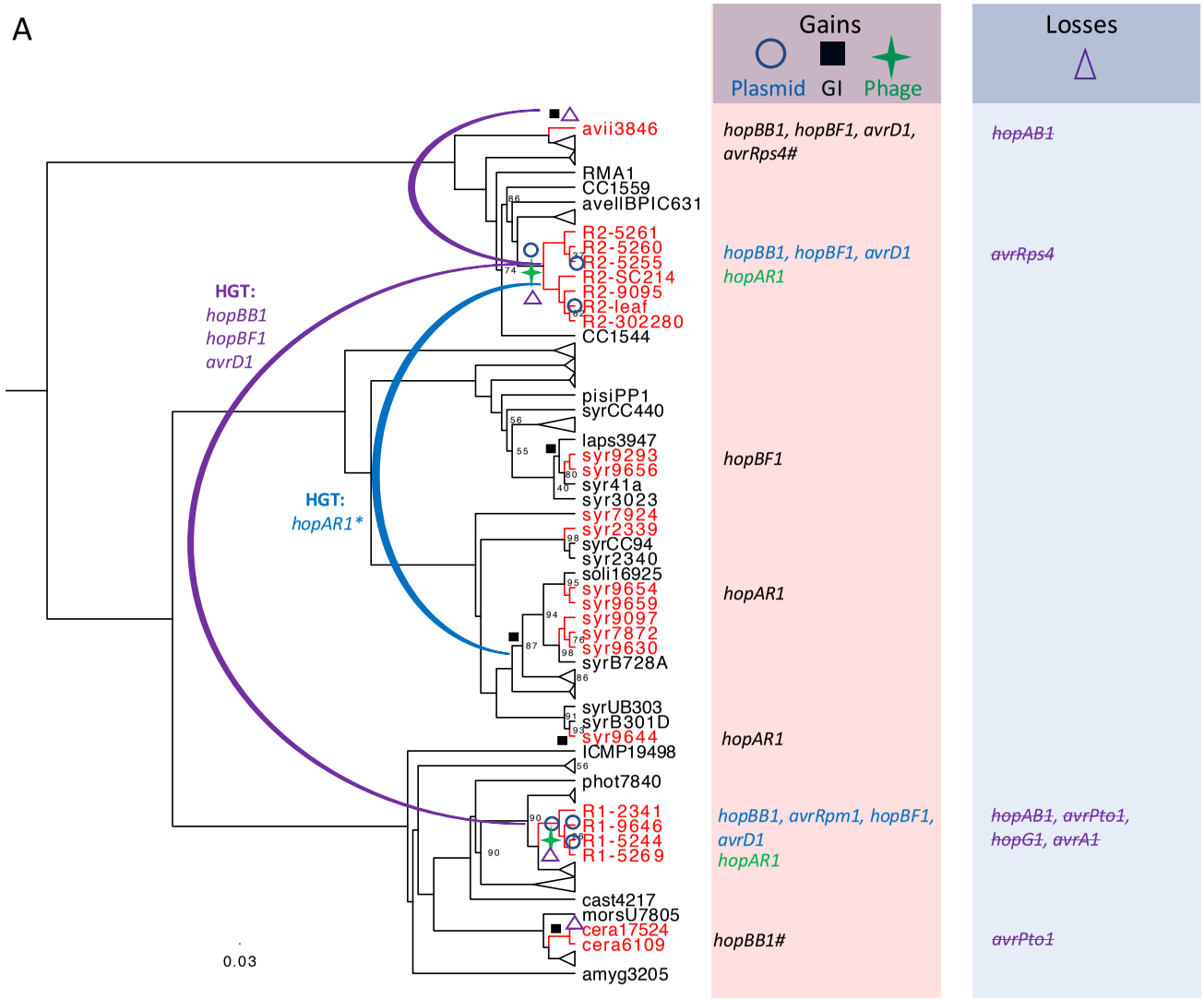
Figure 4.30: Circular visualisation of the genome of* Pss *9097 produced using Circos. *The chromosomal contig is presented on the outer circle. Going inwards, the layers correspond to 1: Effector genes (green) and phytotoxin genes (blue). 2: Genes identified using GLOOME to be gained on the phylogenetic branches leading to this strain. 3: Genomic islands identified using IslandViewer3. 4: Plasmid-structural genes and phage genes identified using PHAST. 5: Insertion sequences identified using ISFinder. 6: GC content with the middle line indicating 50%.*

4.4.12 Integrating results into an evolutionary model of pathogenicity

To integrate all of the effector analysis performed in this study, Figure 4.31 was constructed. This figure focuses on the gain and loss of T3 effector genes across the core genome phylogeny that were statistically associated with cherry pathogenicity in the Bayestraits analysis. During adaptation, multiple clades have gained the effector genes *hopBB1*, *hopBF1* and *avrD1*, often through horizontal gene transfer. Whereas, the effector gene *hopAR1* has been gained independently within bacteriophage sequences or on a genomic island. These effectors are therefore important virulence candidates. Another important event during the evolution of pathogenicity for cherry is the loss of effector genes such as *hopAB1*, which could be putative avirulence factors.

Figure 4.31: Gain and loss of genes statistically associated with cherry pathogenicity on the core genome phylogeny. *A: The core genome phylogeny is presented with strains that can cause disease on cherry in red. Clades containing no cherry pathogens were collapsed. Bootstrap supports below 99% are presented and the scale bar shows substitutions per site. Genes that have been gained and lost are presented in separate columns. Gene gains were divided into plasmid-borne in blue (based on presence of this gene on a plasmid in the PacBio genome sequenced representative of this clade), on a genomic island in black and in a phage sequence in green. Gene losses are in purple and striked through. Different events are plotted onto the specific phylogenetic branches they are predicted to have occurred, using the shapes presented at the column headers. #As* P.s *pv.* cerasicola *and* P.s *pv.* avii *do not have complete genome sequences, it is unknown if these genes are plasmid-borne or chromosomal. The predicted horizontal gene transfer (HGT) events are presented with lines connecting clades. The effector genes* hopBB1*,* hopBF1 *and* avrD1 *have been transferred between* Psm *R1,* Psm *R2 and* P.s *pv.* avii*. *A HGT has occurred of* hopAR1 *between* Psm *R2 and the clade containing cherry and bean strains of* P.s *pv.* syringae*. Genes included in this diagram were chosen as they scored p<0.2 in the Bayestraits analysis. B: Average p.values obtained from 100 runs of the analysis for these effectors are presented in a table.*

A



B

Effector	Mean p.value
avrA1	0.13260308
avrD1	0.03016452
avrPto1	0.07811563
avrRpm2	0.06121665
avrRps4	0.15591279
hopAB1	0.0387609
hopAR1	0.02559804
hopBB1	0.07373386
hopBF1	0.0121589
hopF3	0.18060954
hopG1	0.15869616
hopH1	0.10103587
hopI1	0.12861408

4.5 Discussion

4.5.1 Phylogenetic analysis

Genomic analysis of bacterial populations has the potential to reveal insights into the evolutionary origins of pathogenicity and host specialisation. To make accurate predictions about these complex evolutionary processes, the phylogenetic relationships between the *Prunus*-isolated strains and other strains within the *P. syringae* complex were determined. Three methods were utilised to construct phylogenies based on different amounts of genomic data. All methods placed the different *Prunus*-infecting strains into their respective phylogroups, with *Psm* R2 and *P.s* pv. *avii* in phylogroup 1, *Pss* in phylogroup 2 and *Psm* R1 and *P.s* pv. *cerasicola* in phylogroup 3. This analysis confirmed that the ability to infect *Prunus* species has evolved multiple times across the *P. syringae* complex. Such convergence has been observed with other pathovars. For example, *P.s* pv. *lachrymans* which infects cucumber has members in phylogroups 1 and 3 (Baltrus *et al.* 2011) and strains infecting hazelnut include members of *P.s* pv. *avellanae* and *P.s* pv. *syringae* in phylogroups 3 and 2 (O'Brien *et al.* 2012). The concept of a pathovar was invented before DNA sequencing and phylogenetic analysis (Misaghi & Grogan 1969). Despite its utility in pathogen identification, it seems quite misleading as it provides no information on genetic similarity. For example, the phylogenetic analysis showed *Psm* R1 and R2 fall into separate phylogroups and are therefore distantly related despite being described as the same pathovar.

The three phylogenetic methods involved using a concatenated alignment of seven MLST genes, a concatenated alignment of 902 core genes representing the core genome and a reference mapping approach. All three methods supported similar branching of the major phylogroups, however the placement of strains within each phylogroup varied between trees. The MLST method, although widely used to identify *P. syringae* isolates, was only based on seven genes and therefore may have lacked the information required to make accurate distinctions at the strain level. For example, when more genes were included in the core genome tree the cherry and plum pathogens within *Psm* R1 clustered into distinct clades, due to allelic differences

in the core genome alignment not found in the MLST genes. Using more core genes provided additional informative sites to build phylogenies and therefore should be used when genomic data is available. The RealPhy mapping method produced a similar tree to the core genome. This method differs from the two above as it builds an alignment by mapping reads to a reference genome. It therefore contained both coding and non-coding DNA sequence and the alignment was three times larger than the concatenated core genome alignment. This approach may, however, be biased as by using a reference genome it could imply greater similarity of divergent strains towards the reference, three different references were therefore used to reduce this bias.

The placement of individual strains within each phylogroup differed between the RealPhy and core genome trees, particularly in phylogroups 2 and 3 containing *Pss* and *Psm* R1 strains. In both phylogenies, most of the *Prunus*-infecting clades were monophyletic within their respective phylogroups. However, in the core genome phylogeny *Pss* strains were highly interspersed with other strains isolated from different hosts within phylogroup 2. This suggested that, if the other members of this phylogroup are unable to infect *Prunus,* the pathogenicity phenotype has convergently evolved several times. By contrast, the RealPhy phylogenetic tree clustered most of the *Prunus*-infecting *Pss* strains into three coherent clades. This tree would fit the model of evolution of cherry pathogenicity arising multiple times in this phylogroup 2, however less frequently than in the core genome tree. The RealPhy tree clearly clustered strains within one lineage of phylogroup 2 into a *Prunus*-infecting clade including syr9097 and a Fabaceae (bean/pea) -infecting clade which included the bean pathogen syrB728a. Whereas, in the core genome these strains were interspersed. The RealPhy alignment therefore included some host-specific patterns not observed using just the core genome. Focusing on the other phylogroups, the out-group to *Psm* R1 varied between the two trees, with pathovars *myricae* and *rhaphiolepidis* being closest in the core genome tree, whilst *P.s* pv. *photiniae* was closest in the RealPhy tree. The host species of pathovars *photiniae* and *rhaphiolepidis* are both in the same plant family as *Prunus* (Rosaceae), so could be suitable hosts from which a host shift to *Prunus* could have occurred. Therefore, it is difficult to

conclude which tree more accurately describes the genetic relationships between *Prunus*-infecting strains and other *P. syringae*. The RealPhy tree utilised more genetic information to build the tree, but as it contains non-core regions it may have been more prone to the effects of recombination. If recombination has occurred frequently between strains occupying the same ecological niche, such as the *Prunus*-infecting strains, they would appear more similar in this tree, as is seen within the *Pss* clade. The support values on the RealPhy tree were also slightly lower than those in the core genome tree, implying that this tree's topology was influenced by conflicting phylogenetic patterns.

4.5.2 Search for candidate effectors involved in cherry pathogenicity

To determine the virulence factors involved in cherry pathogenicity, the complete effector repertoires within each strain were identified. With the large amount of genomic data available for *P. syringae*, a large comparative analysis of effector presence and absence could be conducted. A BLAST approach was utilised to identify effector gene homologues within each strain. One limitation of this approach was that due to the fragmentation of incomplete genomes into hundreds of contigs, the results often had to be manually checked to ascertain if a gene truncation was due to a contig break or real variation in the sequence. However, with the increasing use and improved accuracy of long-read sequencing technologies such problems may be lessened in the future. *Prunus*-infecting clades all possessed a conserved effector locus and the core type III effector genes *avrE1*, *hopAA1*, *hopI1*, *hopAH1* and *hopM1* (Baltrus *et al.* 2011), although not all of these were full-length with a functional *hrp* box promoter. In general, the *Pss* strains of phylogroup 2 had much smaller effector repertoires than the other *Prunus*-infecting clades. There was a great amount of diversity in effector presence between the different clades and also within them. In particular, within *Psm* R1 there has been substantial shuffling of effectors, with divergence of the cherry and plum pathogen effector repertoires, potentially mirroring divergences in host immunity.

With this presence and absence matrix, statistical methods could be used to look at gain and loss of each gene on the phylogeny as well as correlating presence/absence

with pathogenicity. All *Prunus*-infecting clades experienced substantial gains in effector genes. Compared to the other clades *Pss* experienced few effector losses (Figures 4.13 and 4.14). As *Pss* has a reduced effector repertoire consisting of mainly core genes it is understandable that effector loss would not occur as this could lead to reduced virulence. By contrast, the other *Prunus*-infecting clades have larger effector gene repertoires where individual genes may function redundantly and therefore gene loss can occur as it would likely have minimal impact on virulence (Cunnac *et al.* 2011).

The GLOOME and BayesTraits analyses revealed 12 candidate virulence-associated effector genes that have been gained in multiple cherry-infecting clades of *P. syringae*, four of which were statistically associated with ability to cause the disease. For the Bayestraits correlation, a binary score of pathogenicity or non-pathogenicity was used. This was based on both the pathogenicity data obtained for the sequenced strains and the assumption that out-groups isolated from other plant species are non-pathogenic to cherry. This analysis could therefore be improved by obtaining clear pathogenicity data for the additional out-groups.

One effector encoding gene (*hopAR1*) was gained in *Psm* R1, R2, *Pss* and *P.s* pv. *avii* as well as being significantly associated with pathogenicity. This effector was present in 18 of the 24 cherry-infecting strains analysed so may be an important virulence factor. This effector was originally identified in bean-infecting *P.s* pv. *phaseolicola* 1302A as an avirulence gene (*avrPphB*) which is expressed from a genomic island and detected by the *R3* resistance protein (Jenner *et al.* 1991; Pitman *et al.* 2005). However, the gene was not found on a similar genomic island in any of the cherry strains, and appears to be within prophage sequences in *Psm* R1 and R2. HopAR1 is a cysteine protease which targets several receptor-like kinases (RLKs), such as BIK1 and the PBS1-like proteins PBL1 and PBL2, to interfere with the PTI response (Zhang *et al.* 2010). Also, despite its role in avirulence on resistant bean cultivars, the *hopAR1* gene has been shown to be retained at a low level in bacterial populations during passaging on the resistant host so it is likely to be important for virulence on susceptible host plants (Neale *et al.* 2016). As the majority of cherry-infecting strains have gained the *hopAR1* gene, this effector could play a similar role in PTI suppression in cherry.

Members of the *hopF* family (*hopF* alleles, *hopBB1* and *avrRpm2*) were also gained in multiple cherry-infecting clades, with *hopBB1* and *avrRpm2* being significantly associated with pathogenicity. A striking expansion of this effector family appears to have occurred in cherry-infecting clades. For example, pathogenic *Psm* R1 strains possess *hopF3*, *hopF4*, *hopBB1* and *avrRpm2*, several of which have been gained on the branches leading to these strains. All members of this diverse effector family share an N terminus and putative myristoylation sites important for effector localization to the plant cell membrane (Lo *et al.* 2016). Two effectors (HopF1 and HopF2) have been widely studied in model species and interact with various membrane-localised host proteins to cause virulence or avirulence. These effectors both have an ADP-ribosyltransferase (ADP-RT) domain, which has been shown for HopF2 to be involved in the ADP-ribosylation of various host proteins. In tomato and *Arabidopsis*, HopF2 appears to target various immune pathways. It interferes with PTI by blocking the phosphorylation of RLKs such as BIK1 which is induced by flagellin-perception (Wu *et al.* 2011; Zhou *et al.* 2014). It also interacts with the MKK5 protein involved in MAPK signalling (Wang *et al.* 2010). HopF2 has also has been shown to prevent ETI by interacting with the *Arabidopsis* protein RIN4 and is involved in stomatal immunity (Hurley *et al.* 2014). However, the other members of the family have different domains following the N-terminus, so their virulence functions may vary. The gain and selective retention of several members of this family in the cherry-infecting strains suggests that they may play various roles in virulence to prevent bacterial detection at the plant cell membrane.

In comparison with these virulence-associated effectors, effectors in the *hopAB* and *avrPto* families were absent from, or putatively pseudogenised in all cherry-infecting strains. BayesTraits also identified the absence of *hopAB1* as being significantly correlated with pathogenicity. These effector families were confirmed to form a redundant effector group (REG) vital for virulence in the model strain *P.s* pv. *tomato* DC3000 (Δ*hopQ1-1*) – N. *benthamiana* pathosystem (Kvitko *et al.* 2009; Cunnac *et al.* 2011). These two effector families function redundantly to suppress the early PTI response and have been shown to be important for virulence in the model systems

tomato, *Arabidopsis* and bean (Lin & Martin 2005; De Torres *et al.* 2006; Jackson *et al.* 1999). Five different alleles of *avrPto* and three alleles of *hopAB* have been identified in the *P. syringae* complex. The AvrPto and HopAB proteins are not very similar at the amino acid level, with HopAB being a much larger, multi-domain effector with a E3-Ubiquitin ligase at its C terminus. However, both effectors target early PTI signalling pathways by interfering with the functions of Pathogen Recognition Receptors (PRRs), such as FLS2 and EFR in *Arabidopsis* (Xiang *et al.* 2008; Xiang *et al.* 2011; Göhre *et al.* 2008). They also both trigger ETI by interacting with the serine-threonine kinase R protein Pto in tomato to trigger NB-LRR Prf-mediated immunity (Ronald *et al.* 1992; Kim *et al.* 2002).

Across the *P. syringae* complex, all non-cherry, plant-associated members of phylogroups 1 and 3 included in this study possessed at least one full-length member of this REG. However, the REG was uncommon in phylogroup 2, where it was only found in Fabaceae-infecting strains and *P.s* pv. *solidagae*, a pathogen of *Solidago altissima* (Figure 4.12). Alignments were created of the genomic regions in *Psm* R1 that contain the *hopAB1* and *avrPto1* genes. These revealed that in the cherry pathogenic strains both effectors have been lost due to the movement of mobile elements (Figure 4.9). In *Psm* R2 the *hopAB3* gene was found to be sufficiently divergent that the Pto-interacting domain was not picked up by Interproscan (Quevillon *et al.* 2005). It was also truncated meaning that the protein does not have the E3 ubiquitin ligase domain, which has been shown to be important in virulence functions (Rosebrock *et al.* 2007). *P.s* pv. *avii* and *P.s* pv. *cerasicola* also contained disrupted members of the REG which if translocated may function differently *in planta* to the full-length versions. The HopAB3 allele in *P.s* pv. *avii* also lacked the Pto-interacting domain and the E3 ubiquitin ligase was disrupted. Functional analysis would be required to determine if these truncated homologues are translocated and how they function *in planta*. The complete loss of both *avrPto1* and *hopAB1* in *Psm* R1 and the divergence and truncation of REG members in *Psm* R2, *P.s* pv. *avii* and *P.s* pv. *cerasicola* suggests that these effectors may be under strong selective pressures due to detection by the cherry immune system. The lack of a Pto-interacting domain in the HopAB proteins putatively produced by *Psm* R2 and *P.s* pv. *avii* also suggests that

these protein sequences may have diverged to avoid Pto recognition. It would be interesting to see if these effectors still function in the virulence of cherry-infecting strains. If this REG is not required for disease in cherry (as both *Psm* R1 and *Pss* lack both families), are other effectors involved in suppressing early PTI responses that this REG is clearly required for in model systems?

Finally, as there were two alternative phylogenies (the core genome and RealPhy trees), the GLOOME and Bayestraits analyses were ran using both to see if the results differed. The main differences were apparent in which branches each effector was gained/lost. In phylogroup 2 effectors such as *hopAR1*, *hopAW1* and *avrRpm1* appear to be gained multiple times when using the core genome tree as the *Prunus*-infecting strains are distantly related. Whereas, in the RealPhy tree the acquisition of these effectors only happens once as the strains for a monophyletic group. As the gene trees of these three effectors indicated that all *Prunus*-infecting *Pss* alleles were highly similar (Figure S4.36), the independent acquisition of these effectors during evolution is unlikely. Instead acquisition is occurring at a population level, as strains with diverse genetic backgrounds recombine and thus shuffle parts of their accessory genome.

4.5.3 Horizontal transfer has been important in the acquisition of key effectors

Many of the key effectors identified as being significantly associated with cherry pathogenicity showed evidence of horizontal transfer between the cherry-infecting clades. In particular, despite being distantly related *Psm* R1, R2 and *P.s* pv. *avii* commonly shared effectors that were on native plasmids within the PacBio-sequenced strains. This suggested that transfer of effectors via plasmids occurs frequently between phylogroups 1 and 3, allowing bacteria to shuffle their repertoire of effector genes. Selective retention of these plasmids within bacterial populations would then occur if the genes gained, such as those encoding effectors improve fitness in that particular niche (Guttman, 2009). In comparison, all of the *Prunus*-infecting *Pss* strains were found to lack plasmids, apart from syr9644 which possesses one plasmid (Figure 1). Previous studies that surveyed bacteria from *Prunus* orchards found that most *Pss* strains possess no plasmids (Liang *et al.* 1994). Instead, effectors within *Pss* were often located within genomic islands on the chromosome (Figure 4.30).

To determine if effectors were commonly found in mobile genomic islands, putative genomic islands were identified in the PacBio sequenced strains of *Psm* R1, R2 and *Pss*. To determine the likely source of these genomic islands, other strains of *P. syringae* were searched for homologous sequences (Figure 27). Homologous sequences of the *Psm* R1 genomic islands were found in many strains in both phylogroups 1 and 3 suggesting that these islands have been mobilised between the two phylogroups. The *Psm* R2 genomic islands were mostly lineage-specific, with some being found in members of the other phylogroups. Whereas, the majority of *Pss* genomic islands were commonly found in members of phylogroup 2, suggesting that *Pss* has recombined frequently with members of its own phylogroup to share genomic islands.

Therefore, as the *Prunus*-infecting clades occupy the same ecological niche, occurring together on cherry tissue, members of *Psm* R1, R2, *P.s* pv. *avii* and *P.s* pv. *cerasicola* may have frequently interacted to share important virulence effectors (as discussed by Liang *et al.* 1994). Strains of *Pss* appear to be genetically isolated as they have not shared effectors with the other clades. Further studies may reveal if differences in recombination machinery or genetic incompatibility mechanisms reduces the capacity of *Pss* to recombine and accept plasmids from the other phylogroups. Or perhaps, differences in life strategy mean that there is no selection pressure for *Pss* to gain any new effectors as it possesses other virulence factors, notably a suite of toxins, that allow it to achieve similar virulence levels *in planta*.

4.5.4 Other virulence factors are important for cherry pathogenicity

The genomes were also searched for the presence of phytotoxin biosynthesis clusters. Phytotoxins within *P. syringae* are generally host non-specific molecules that directly damage plant cells and are important in the production of disease symptoms such as chlorosis and necrosis (Bender *et al.* 1999). They are not secreted by the type III secretion system like the proteinaceous effectors and are not fully required for *P. syringae* pathogenicity, but do increase virulence *in planta* (Dudnik & Dudler 2014). The chlorosis-inducing toxin coronatine has been gained on a plasmid in pathogenic

Psm R1 strains. Coronatine mimics a protein in the jasmonic acid pathway to promote down-regulation of the salicylic acid signalling pathway, which is an important defence against *P. syringae* (Grant & Jones 2009). The gain of this toxin in the pathogenic strains may be one of the crucial differences that makes these strains virulent on cherry compared to other members of *Psm* R1. Interestingly, GLOOME showed that these strains have lost the gene for *hopX1*, an effector which functions redundantly with coronatine to reduce salicylic-acid signalling (Gimenez-Ibanez *et al.* 2014). The cost associated with producing both virulence factors could have led to the selective loss of *hopX1* from these strains. In non-pathogenic *Psm* R1 (R1-5300), *hopX1* is found on a plasmid along with another effector gene *hopG1* that has been lost in pathogenic strains, so it is likely that these linked effectors have been lost together during specialisation towards cherry.

The necrosis-inducing lipodepsipeptides syringopeptin and syringmycin and proteasome-inhibiting toxin syringolin A were found exclusively in phylogroup 2 strains (Scholz-Schroeder *et al.* 2001; Schellenberg *et al.* 2010; Baltrus *et al.* 2011). All *Prunus*-infecting *Pss* strains possessed two or three of these toxin biosynthesis clusters. The large necrosis observed on cherry fruits infected with *Pss* strains (see Chapter 3) is likely to be due to these necrosis-inducing toxins as their contribution has been demonstrated in previous studies (Scholz-Schroeder *et al.* 2001; Dudnik & Dudler 2014). Two divergent *Pss* strains also possessed the antimetabolite toxin mangotoxin which interferes with the plant's ability to produce ornithine/arginine (Arrebola *et al.* 2009). All *Pss* strains therefore possess at least two toxin biosynthetic clusters, whilst having substantially reduced effector repertoires. This observation has led to the hypothesis that a phenotypic trade-off exists between effectors and toxins, with members of phylogroup 2 retaining only a small set of core effectors and relying more on phytotoxins in their pathogenicity strategy (Baltrus *et al.* 2011; Hockett *et al.* 2013).

In addition to phytotoxins, genomes were searched for other virulence factors such as genes involved in the degradation of compounds such as the phenolics catechol and anthanilate found in woody plant tissues (Rodríguez-Palenzuela *et al.* 2010; Caballo-

Ponce *et al.* 2016). These genes were present in all *Prunus*-infecting clades apart from *P.s* pv. *avii* and some members of *Pss*. The convergent acquisition of these metabolic pathways suggests that they may improve fitness of bacterial populations during the woody-tissue lifecycle stage by allowing them to either utilize or reduce the toxicity of compounds specific to this niche. However, the lack of these pathways in strains of *P.s* pv. *avii* and *Pss* suggests that they are not essential for the ability to cause bacterial canker on *Prunus*. In particular, *Pss* strains were shown to be highly pathogenic during wound and field experiments (Chapter 3) even without possessing such genes. The contributions made by these metabolic pathways may underlie more subtle differences in the members of each clade's ability to persist *in planta*. Crosse and Garrett (1966), observed that *Psm* R1 survived in cankers for a longer time than *Pss*. Perhaps this increased persistence is linked to genes involved in woody-tissue adaptation.

Ice nucleation genes were also common in *Prunus*-infecting strains. The production of ice nucleation proteins is hypothesised to increase virulence on woody host species as ice nucleation promotes frost damage which provides an entry route for bacteria (Lamichhane *et al.* 2014). As with the wood degradation gene clusters, several *Prunus*-infecting clades possessed ice nucleation genes, so they could be important in promoting bacterial virulence on cherry.

Finally, GLOOME was utilised to look at the whole set of genes possessed by *P. syringae* and see which genes had been gained on the phylogenetic branches leading to cherry pathogens. With the large number of genomes available in online databases and the pathogenicity data collected for *Prunus*-infecting strains, this un-biased approach would allow the identification of novel pathways that are important in disease. GLOOME revealed that genes involved in various metabolic pathways have been gained in multiple cherry clades, such as the degradation of aromatic compounds, sugar-uptake, urea decomposition and haem utilisation. Plotting the gained genes onto the PacBio genomes (Figures 4.28-4.30) revealed that many were on genomic islands, so have likely been mobilised from other bacteria.

4.5.5 Linking genomics to host specialisation

Host specialisation onto cherry has occurred in multiple clades of *P. syringae*. Within each of the three major phylogroups pathogenicity has arisen at least twice, with strains using both shared and distinctive virulence strategies to cause disease.

The strategy employed by all clades apart from *Pss* appears to involve using a large effector repertoire to subvert host immunity. Effector evolution in *P. syringae* is highly dynamic with genes being gained, lost and mutated across the phylogeny (Guttman 2009; Sarkar *et al.* 2006). This rapid evolution allows bacteria to shuffle their effector set to adapt to a new host's immune system (Lindeberg *et al.* 2012) and as such has likely contributed toward the host shift of these clades onto cherry. Members of phylogroups 1 and 3 are distantly related. The pairwise average nucleotide identity (Figure S4.4) between members of the two phylogroups was on average 87.5%, showing they should be classed as separate 'species' (Kim *et al.* 2014). However, despite this divergence, cherry-infecting strains have gained many of the same effector genes and there is evidence for horizontal transfer between them.

The fact that, *Pss* has a much smaller effector repertoire than *Psm*, may allow it to occupy a wider host range as it is less likely to possess avirulence-associated effectors. As an evolutionary strategy *Pss* may therefore be more of a generalist pathogen using a small group of core effectors to suppress immunity and its non-host specific toxins to maximise virulence. This reduced effector complement is common across phylogroup 2. However, some lineages such as *P.s* pv. *papulans* and syrHS191 possessed over 20 effector genes (Figure S4.3), but lacked most of the toxin biosynthesis clusters which means they could be reverting to a more effector-driven virulence strategy.

One hypothesis is that with such a small effector repertoire *Pss* may be unable to remain biotrophic in cherry for as long as the *Psm* clades and therefore relies on its toxins to cause plant cell death. This idea is supported by the rapid symptom development observed when cherry leaves are inoculated with *Pss* (Chapter 3, Figure 3.12). It has been shown in grasses that *Pss* is able to multiply and reach similar

population sizes when its phytotoxin genes have been deleted (Dudnik & Dudler 2014). It would therefore be expected that *Pss* is unable to survive *in planta* without its effectors as these are required for the initial suppression of plant immunity. However, it would be interesting to determine how long *Pss* can persist in cherry tissues without its repertoire of phytotoxins.

Host-specificity appears to haved occured to a degree within the *P.s* pv. *syringae* clades of phylogroup 2. *Pss* strains were originally characterised as belonging to *P.s* pv. *syringae* due to their pathogenicity for lilac (Young, 1991). Most *Pss* strains are therefore both pathogenic to lilac and their homologous host (Yessad-Carreau *et al.* 1994). Some studies have identified host-specificity occurring in strains isolated from bean, pear and grasses (Yessad-Carreau *et al.* 1994; Rezaei & Taghavi 2014; Gross 1977). However, pathogenicity assays on apricot and almond showed that strains from various other hosts were able to incite disease (Rezaei & Taghavi 2014). This study found that *Pss* isolated from plum are pathogenic on cherry and have similar virulence levels to cherry isolates (Chapter 3). Comparative genomics showed cherry and plum isolates are closely related and have similar effector repertoires, meaning that the host range of these *Pss* strains is not limited to one host species. When comparing the *Prunus* strains with strains from other hosts, the *Prunus* strains were interspersed with strains isolated from Fabaceae in the core genome phylogeny. However, the RealPhy phylogeny, which is built using a larger alignment, clustered Fabaceae and *Prunus* strains into separate lineages. Examination of the effector repertoires showed distinct differences that could be involved in host specificity. For example, the bean strains all possessed *hopAB1* and *avrPto1*. Previous studies have shown the importance of *hopAB1* in the virulence of *P.s* pv. *phaseolicola* on bean (Jackson *et al.* 1999; De Torres *et al.* 2006). These *Pss* strains may have therefore gained this effector to maximise disease on bean. These effectors are absent from all cherry-infecting *P. syringae* so are either not required for disease or may be avirulence factors. If they are avirulence factors this would prevent the growth of Fabaceae strains on cherry, thus restricting their host range. In comparison, the *Prunus* strains have gained several effectors not found in bean strains, such as *hopA2*, *hopBE1* and *hopAW1*. Functional analysis could

be used to confirm if these particular effectors are important in disease and thus have been gained during specialisation towards *Prunus*.

Interestingly, the different cherry-infecting lineages within *Pss* have slight differences in their virulence factor repertoires which reflect their convergent adaptation towards *Prunus*. One clade (containing syr9097) possessed pathogenicity-associated effectors such as *hopAR1*, *hopAW* and *hopBE1* and possessed the phytotoxin gene clusters syringomycin, syringopeptin and syringolin. Another clade (containing syr9654) lacked syringopeptin and had slight differences in its effector repertoire (it had gained *hopAZ1,* but lacked *hopBE1* and *hopAF1*). A related clade (containing syr2339) yet again had a divergent effector repertoire, but possessed some of the genes involved in degradation of compounds found in wood. The final *Pss* clade (containing syr9656) varied in effector content and possessed the gene cluster for mangotoxin production. Therefore, effectors and other virulence factors may be functioning redundantly in these strains to enable them to cause the same disease outcome, with different effector repertoires.

4.5.6 Conclusions

In conclusion, comparative genomics revealed that host shifts onto cherry have occurred many times within the *P. syringae* species complex, with the different clades using both common and distinct mechanisms of pathogenicity. These clades use different virulence strategies to cause the same disease, illustrating how versatile members of the *P. syringae* complex are as plant pathogens. Genetic diversity within *P. syringae* has enabled the convergent evolution of disease and the dynamic nature of bacterial genomes, indicates that there may be other unknown clades of *P. syringae* with the capacity to infect and grow within cherry trees.

Using the knowledge of common effectors and other virulence factors involved in this disease, genomics could be used to predict the likelihood of a particular strain being pathogenic based on gene content. The field of genomics is already being applied to plant pathogens by identifying pathogen-specific DNA regions to inform the design of DNA-based field assays (Ruinelli *et al.* 2016; Getaz *et al.* 2017). However, these are

usually tailored to detect only a particular clade, an approach that would be limited when various pathogen groups are involved, as in bacterial canker. The use of whole-genome sequencing and bioinformatics tools to model the likelihood of pathogenicity provides a more robust method less affected by pathogen diversity. As the cost and ease of sequencing improve in the future, this may become a promising tool for both diagnostics and the predication of both plant and human disease epidemics (Reuter *et al.* 2013; Whiteside *et al.* 2016).

With the increasing number of genome sequences available and improvement of sequencing technologies, hypotheses can be made about complex evolutionary processes occurring in bacterial populations. Several candidate virulence and avirulence effector genes were identified in cherry pathogens. These could form the basis of further study into the immune responses triggered by these pathogens and how differential recognition of pathogen effector proteins is involved in both host and non-host resistance to canker. With this knowledge about the various distinct clades that are involved in bacterial canker, it may be possible to tailor breeding and control measures towards the dominant clades causing disease in particular geographical regions, and use shared 'core' effectors to search for broad resistance. Therefore, by improving knowledge of these plant pathogens, genomics-informed strategies could be used in the future to reduce the incidence of this chronic and sometimes devastating canker disease.

Chapter 4: Supplementary results

Effector	Strain	Protein length	*hrp* box	Comments
avrA1	5300	906	GAAACCGAAACGGCGTTGCTTGCCACACA	Acquired in *Psm* R1 on genomic island
avrA1	9326	906	GAAACCGAAACGGCGTTGCTTGCCACACA	
avrA1	9629	906	GAAACCGAAACGGCGTTGCTTGCCACACA	
avrA1	9657	906	GAAACCGAAACGGCGTTGCTTGCCACACA	
avrB2	leaf	352	GGAACCGTTCTGCAACTCATGCCACTAA	On a genomic island
avrB2	sc214	352	GGAACCGTTCTGCAACTCATGCCACTAA	
avrB3	9293	323	No *hrp* box found	On a genomic island
avrB3	9656	323	No *hrp* box found	
avrB4	RMA1	320	GGAACCGAACACCAATTTTTATCCACTCA	On a genomic island
avrD1	5244	311	GGAACCAAATCCGTCCCAAAGGCCACACA	
avrD1	5255	311	GGAACCAAATCCGTCCCAAAGGCCACACA	Putative HGT from *Psm* R1 isolates
avrD1	5300	311	GGAACCAAATCCGTCCCAAAGGCCACACA	Divergent from other members of *Psm* R1
avrD1	9326	311	GGAACCAAATCCGTCCCAAAGGCCACACA	
avrD1	9629	311	GGAACCAAATCCGTCCCAAAGGCCACACA	
avrD1	9646	311	GGAACCAAATCCGTCCCAAAGGCCACACA	
avrD1	9657	311	GGAACCAAATCCGTCCCAAAGGCCACACA	
avrE1	5244	1712	GGAACCCTCCATTATCAGGTGCCACTCA	
avrE1	5255	1801	GGAACCCGCTGGCATTGCATGCTACTCA	
avrE1	5260	1801	GGAACCCGCTGGCATTGCATGCTACTCA	
avrE1	5300	1712	GGAACCCTCCATTATCAGGTGCCACTCA	
avrE1	9097	1820	GGAACCGGTGCCTTCCCCATGCCACACA	
avrE1	9293	1789	GGAACCGGTGCCTTTCCGATGCCACACA	
avrE1	9326	1712	GGAACCCTCCATTATCAGGTGCCACTCA	
avrE1	9629	1712	GGAACCCTCCATTATCAGGTGCCACTCA	

Effector	Strain	Protein length	*hrp* box	Comments
avrE1	9630	1820	GGAACCGGTGCCTTCCCCATGCCACACA	
avrE1	9643	1795	TGAGTGGCATGCAATGCCAGCGGGTTCC	
avrE1	9644	1820	GGAACCGGTGCCTTCCCCATGCCACACA	
avrE1	9646	1712	GGAACCCTCCATTATCAGGTGCCACTCA	
avrE1	9654	1820	GGAACCGGTGCCTTCCCCATGCCACACA	
avrE1	9656	1789	GGAACTGGTGCCTTTCCCATGCCACACA	
avrE1	9657	1712	GGAACCCTCCATTATCAGGTGCCACTCA	
avrE1	9659	1820	GGAACCGGTGCCTTCCCCATGCCACACA	
avrE1	RMA1	1801	GGAACCCGCTGGCATTGCATGCCACCCA	
avrE1	leaf	1801	GGAACCCGCTGGCATTGCATGCTACTCA	
avrE1	sc214	1801	GGAACCCGCTGGCATTGCATGCTACTCA	
avrpto1	5300	162	GGAACCGATCTGCCCCCGATGACCACTCA	On a genomic island
avrpto1	9326	162	GGAACCGATCTGCCCCCGATGACCACTCA	
avrpto1	9629	162	GGAACCGATCTGCCCCCGATGACCACTCA	
avrpto1	9657	162	GGAACCGATCTGCCCCCGATGACCACTCA	
avrpto5	RMA1	158	TAAGTGGTCGCGGGGACCCGACTGGTTCC	
avrRpm1	9097	228	GGAACTCATTAGCTCTCGAATGCCACAGA	
avrRpm1	9630	228	GGAACTCATTAGCTCTCGAATGCCACAGA	
avrRpm1	9644	228	GGAACTCATTAGCTCTCGAATGCCACAGA	
avrRpm1	9654	228	GGAACTCATTAGCTCTCGAATGCCACAGA	
avrRpm1	9659	228	GGAACTCATTAGCTCTCGAATGCCACAGA	
avrRpm2	5244	223	GGAACCAAATATGTAGTTATGGTCACTCA	
avrRpm2	5300	223	GGAACCAAATATGTAGTTATGGTCACTCA	
avrRpm2	9326	223	GGAACCAAATATGTAGTTATGGTCACTCA	
avrRpm2	9629	223	GGAACCAAATATGTAGTTATGGTCACTCA	
avrRpm2	9646	223	GGAACCAAATATGTAGTTATGGTCACTCA	

Effector	Strain	Protein length	*hrp* box	Comments
avrRpm2	9657	223	GGAACCAAATATGTAGTTATGGTCACTCA	
avrRps4	5255	204	No *hrp* box found	Missing beginning of gene
avrRps4	5260	204	No *hrp* box found	Missing beginning of gene
avrRps4	RMA1	221	GGAACTGCCGGCTAGTGCGGCTCCACTAA	Potentially non-functional as no KRVY domain
avrRps4	leaf	221	GAGACAGTACACTAGTGCGGCTCCACTAA	Potentially non-functional as no KRVY domain
avrRps4	sc214	185	No *hrp* box found	Missing beginning of gene and contig
avrRpt2	9326	258	GGAACTATTTCGAACTCTGCGACCACTCA	On a genomic island
avrRpt2	9657	258	GGAACTATTTCGAACTCTGCGACCACTCA	
hopA1	5244	375	ATAACCTGTTGGAGCGAAGCTCGCCTATGA	*hrp* box too long
hopA1	9643	380	TGAGTGGAGCCTCTGCTCGAGGCGGTTCC	
hopA1	9646	375	GCAACCCTTGGTATTCGACGACACA	
hopA1	RMA1	380	TGAGTGGGGCCTCTGCTCGAGGCGGTTCC	
hopA2	5244	147	TGAACCCGATACAAACACGTTTCTCTAA	Putative stop codon in middle of gene
hopA2	5255	377	GGAACCGCCCCAGGCAGAACTCCCACTCA	Divergent from other pathovars
hopA2	5260	377	GGAACCGCCCCAGGCAGAACTCCCACTCA	Divergent from other pathovars
hopA2	5300	147	No *hrp* box found	Putative stop codon in middle of gene
hopA2	9097	382	GTCCATTCGTACAGTGTTCAGGGCACCGA	
hopA2	9326	147	CGATGCAACCCTTGGTATTCGACGACACA	Putative stop codon in middle of gene
hopA2	9629	147	CGATGCAACCCTTGGTATTCGACGACACA	Putative stop codon in middle of gene
hopA2	9630	382	GTCCATTCGTACAGTGTTCAGGGCACCGA	
hopA2	9644	382	GTCCATTCGTACAGTGTTCAGGGCACCGA	
hopA2	9646	147	CGATGCAACCCTTGGTATTCGACGACACA	Putative stop codon in middle of gene
hopA2	9654	382	GTCCATTCGTACAGTGTTCAGGGCACCGA	
hopA2	9657	147	CGATGCAACCCTTGGTATTCGACGACACA	Putative stop codon in middle of gene
hopA2	9659	382	GTCCATTCGTACAGTGTTCAGGGCACCGA	
hopA2	leaf	377	GGAACCGCCCCAGGCAGAACTCCCACTCA	Divergent from other pathovars

Effector	Strain	Protein length	*hrp* box	Comments
hopA2	sc214	377	GGAACCGCCCCAGGCAGAACTCCCACTCA	Divergent from other pathovars
hopAA1	5244	344	GGAACCGTCAACGGATCCGGGACCACACA	Inversion causing pseudogenisation
hopAA1	5255	344	GGAACCGTCAACCAATCCGGGACCACACA	Truncation, missing end
hopAA1	5260	344	TGTGTGGTCCCGGATTGGTTGACGGTTCC	Truncation, missing end
hopAA1	5300	486	GGAACCGTCAACGGATCCGGGACCACACA	
hopAA1	9097	483	GGAACCCCCGGCACTGTTGCGACCACACA	
hopAA1	9293	485	GGAACCCCCGGCACCCTTGCGACCACACA	
hopAA1	9326	482	GGAACCGTCAACGGATCCGGGACCACACA	
hopAA1	9629	482	TGTGTGGTCCCGGATCCGTTGACGGTTCC	
hopAA1	9630	483	GGAACCCCCGGCACTGTTGCGACCACACA	
hopAA1	9643	486	GGAACCGTCAACCGATCCGGGACCACACA	
hopAA1	9644	483	GGAACCCCCGGCACTGTTGCGACCACACA	
hopAA1	9646	344	GGAACCGTCAACGGATCCGGGACCACACA	Inversion causing pseudogenisation
hopAA1	9654	483	GGAACCCCCGGCACTGTTGCGACCACACA	
hopAA1	9656	485	GGAACCCCCGGCACCCTTGCGACCACACA	
hopAA1	9657	482	TGTGTGGTCCCGGATCCGTTGACGGTTCC	
hopAA1	9659	483	GGAACCCCCGGCACTGTTGCGACCACACA	
hopAA1	RMA1	485	TGTGTGGTCCCGGATCGGTTGACGGTTCC	
hopAA1	leaf	344	GGAACCGTCAACCAATCCGGGACCACACA	Truncation, missing end
hopAA1	sc214	344	GGAACCGTCAACCAATCCGGGACCACACA	Truncation, missing end
hopAB1	5300	541	GGAACCATTTATGTCCGATCGCCAACTCA	
hopAB1	9326	541	GGAACCATTTATGTCCGATCGCCAACTCA	
hopAB1	9629	541	GGAACCATTTATGTCCGATCGCCAACTCA	
hopAB1	9657	541	GGAACCATTTATGTCCGATCGCCAACTCA	
hopAB3	5255	440	GGAACGCTTTCCAACTTTGTTGCCACATA	Premature stop codon
hopAB3	5260	440	GGAACGCTTTCCAACTTTGTTGCCACATA	Premature stop codon
hopAB3	9643	579	GGAACTCTTTCCTGCTCTTTTTCCACACA	

Effector	Strain	Protein length	*hrp* box	Comments
hopAB3	RMA1	579	GGAACTCTTTCCTGCTCTTTTGCCACACA	
hopAB3	leaf	440	GGAACGCTTTCCAACTTTGTTGCCACATA	Premature stop codon
hopAB3	sc214	440	GGAACGCTTTCCAACTTTGTTGCCACATA	Premature stop codon
hopAE1	5244	912	GGAACCGTTCCGGGCCGGACGACACTCA	
hopAE1	5300	912	GGAACCGTTCCGGGCCGGACGACACTCA	
hopAE1	9097	912	GGAACCGTTCCGGGCCGGACGACACTCA	
hopAE1	9326	912	GGAACCGTTCCGGGCCGGACGACACTCA	
hopAE1	9629	912	GGAACCGTTCCGGGCCGGACGACACTCA	
hopAE1	9630	912	GGAACCGTTCCGGGCCGGACGACACTCA	
hopAE1	9643	912	GGAACCGTTCCGGATCGAGCGCCACTTA	
hopAE1	9644	912	GGAACCGTTCCGGGTCGGACGACACTCA	
hopAE1	9646	912	GGAACCGTTCCGGGCCGGACGACACTCA	
hopAE1	9654	912	GGAACCGTTCCGGGCCGGACGACACTCA	
hopAE1	9657	912	GGAACCGTTCCGGGCCGGACGACACTCA	
hopAE1	9659	912	GGAACCGTTCCGGGCCGGACGACACTCA	
hopAF1	5244	280	GGAACTTTTTCTTGCCCGCTACCACCCA	Possible transfer R1-R2 on genomic island
hopAF1	5255	280	GGAACTTTTTCTTGCCCGCTACCACCCA	
hopAF1	5260	280	GGAACTTTTTCTTGCCCGCTACCACCCA	
hopAF1	9646	280	GGAACTTTTTCTTGCCCGCTACCACCCA	
hopAF1	leaf	280	GGAACTTTTTCTTGCCCGCTACCACCCA	
hopAF1-2	5244	283	GGAACCGCTGAAGAGTTTTAGCCACTCA	
hopAF1-2	5300	283	GGAACCGCTGAAGAGTTTTAGCCACTCA	
hopAF1-2	9097	280	GGAACCGCTGAAGAGTTTTAGCCACTCA	
hopAF1-2	9326	283	GGAACCGCTGAAGAGTTTTAGCCACTCA	
hopAF1-2	9629	283	GGAACCGCTGAAGAGTTTTAGCCACTCA	
hopAF1-2	9630	280	GGAACCGCTGAAGAGTTTTAGCCACTCA	
hopAF1-2	9643	284	GGAACCAATGAAGAGCTTGAGCCACTCA	

Effector	Strain	Protein length	*hrp* box	Comments
hopAF1-2	9644	280	GGAACCGCTGAAGAGTTTTAGCCACTCA	
hopAF1-2	9646	283	GGAACCGCTGAAGAGTTTTAGCCACTCA	
hopAF1-2	9656	286	GGAACCGCTGAAGAGTTTTAGCCACTCA	
hopAF1-2	9657	283	GGAACCGCTGAAGAGTTTTAGCCACTCA	
hopAF1-2	RMA1	284	TGAGTGGTTAAAACTCTTCAGTGGTTCC	
hopAG1-1	9097	715	GGAACCGCACAGCACACTGCGACACTCA	
hopAG1-1	9293	715	GGAACCGCATAGCACACTGCGACACTCA	
hopAG1-1	9630	715	GGAACCGCACAGCACACTGCGACACTCA	
hopAG1-1	9643	708	GGAACCCGATGACACAAGGCGACACTCA	
hopAG1-1	9644	715	GGAACCGCACAGCACACTGCGACACTCA	
hopAG1-1	9654	715	GGAACCGCACAGCACACTGCGACACTCA	
hopAG1-1	9656	715	GGAACCGCATAGCACACTGCGACACTCA	
hopAG1-1	9659	715	GGAACCGCACAGCACACTGCGACACTCA	
hopAG1-1	RMA1	707	GGAACCCGATGACACAAGGCGACACTCA	
hopAG1-2	5255	222 start, 485 end	GGAACCCGATGACACAAGGCGACACTCA	disrupted in middle with Insertion Sequence
hopAG1-2	5260	222 start, 485 end	GGAACCCGATGACACAAGGCGACACTCA	disrupted in middle with Insertion Sequence
hopAG1-2	leaf	222 start, 485 end	GGAACCCGATGACACAAGGCGACACTCA	disrupted in middle with Insertion Sequence
hopAG1-2	sc214	222 start, 485 end	GGAACCCGATGACACAAGGCGACACTCA	disrupted in middle with Insertion Sequence
hopAH1_RAST	5255	372	No *hrp* box found	This is the RAST annotation
hopAH1_RAST	5260	372	No *hrp* box found	This is the RAST annotation
hopAH1_RAST	leaf	372	No *hrp* box found	This is the RAST annotation
hopAH1_RAST	sc214	372	No *hrp* box found	This is the RAST annotation
hopAH1_RAST	9643	404	No *hrp* box found	This is the RAST annotation
hopAH1_RAST	RMA1	392	No *hrp* box found	This is the RAST annotation
hopAH1-1	5244	416	GCAGCCACTTTTCGCAAAGAGCCCACAGA	Next to RAST annotation
hopAH1-1	5255	415	GCAGCCACTTTTCGCACAGAGCCCACAGA	Next to RAST annotation
hopAH1-1	5260	415	GCAGCCACTTTTCGCACAGAGCCCACAGA	Next to RAST annotation

Effector	Strain	Protein length	*hrp* box	Comments
hopAH1-1	9097	380	GAAACCCTGTACAAAAAACCGTTCACCCA	Next to RAST annotation
hopAH1-1	9293	380	GAAACCCTGCACAAAAACCCGTTCACCCA	Next to RAST annotation
hopAH1-1	9326	416	GCAGCCACTTTTCGCAAAGAGCCCACAGA	Next to RAST annotation
hopAH1-1	9629	416	GCAGCCACTTTTCGCAAAGAGCCCACAGA	Next to RAST annotation
hopAH1-1	9630	380	GAAACCCTGTACAAAAAACCGTTCACCCA	Next to RAST annotation
hopAH1-1	9643	415	GGAATAGCACGCCGACAGCCGCCACCACA	Next to RAST annotation
hopAH1-1	9644	380	GAAACCCTGTACAAAAAACCGTTCACCCA	Next to RAST annotation
hopAH1-1	9646	416	GCAGCCACTTTTCGCAAAGAGCCCACAGA	Next to RAST annotation
hopAH1-1	9654	380	GAAACCCTGTACAAAAAACCGTTCACCCA	Next to RAST annotation
hopAH1-1	9656	380	GAAACCCTGCACAAAAACCCGTTCACCCA	Next to RAST annotation
hopAH1-1	9657	416	GCAGCCACTTTTCGCAAAGAGCCCACAGA	Next to RAST annotation
hopAH1-1	9659	380	GAAACCCTGTACAAAAAACCGTTCACCCA	Next to RAST annotation
hopAH1-1	RMA1	415	GCAGCCACTTTTCGCACAGAGCCCACAGA	Next to RAST annotation
hopAH1-1	leaf	415	GCAGCCACTTTTCGCACAGAGCCCACAGA	Next to RAST annotation
hopAH1-1	sc214	415	GCAGCCACTTTTCGCACAGAGCCCACAGA	Next to RAST annotation
hopAH1-2	5255	255	GAAACCCTGCGAAAAACCGGTTCACCCA	Part of endoglucanase annotation
hopAH1-2	5260	255	GAAACCCTGCGAAAAACCGGTTCACCCA	Part of endoglucanase annotation
hopAH1-2	leaf	255	GAAACCCTGCGAAAAACCGGTTCACCCA	Part of endoglucanase annotation
hopAH1-2	sc214	255	GAAACCCTGCGAAAAACCGGTTCACCCA	Part of endoglucanase annotation
hopAH1-2	9643	255	GGAATAGCACGCCGACAGCCGCCACCACA	Part of endoglucanase annotation
hopAH1-2	RMA1	255	GGAACGGCACGCCGACAGCCGCCGCCACA	Part of endoglucanase annotation
hopAI1	5255	261	GAATATCAAGGAGGAGTCGGTGAGACTCA	
hopAI1	5260	261	GAATATCAAGGAGGAGTCGGTGAGACTCA	
hopAI1	9097	267	GGAACCTGTTTTACCTGTCGACCGCTCA	Truncated at beginning
hopAI1	9293	267	GGAACCTGATTTGTCTGTTGACCACTCA	
hopAI1	9630	267	GGAACCTGTTTTACCTGTCGACCGCTCA	Truncated at beginning
hopAI1	9643	261	TGAATATCAAGGAGGAGTCGGTGAGACTC	

Effector	Strain	Protein length	*hrp* box	Comments
hopAI1	9644	267	GGAACCTGTTTTACCTGTCGACCGCTCA	Truncated at beginning
hopAI1	9654	267	GGAACCTGTTTTACCTGTCGACCGCTCA	Truncated at beginning
hopAI1	9656	78	GGAACCTGATTTGTCTGTTGACCACTCA	Disruption due to stop codon in middle
hopAI1	9659	267	GGAACCTGTTTTACCTGTCGACCGCTCA	Truncated at beginning
hopAI1	RMA1	261	TGAGTGGGGCCTCTGCTCGAGGCGGTTCC	
hopAI1	leaf	261	GAATATCAAGGAGGAGTCGGTGAGACTCA	
hopAI1	sc214	261	GAATATCAAGGAGGAGTCGGTGAGACTCA	
hopAO1	5255	303	No *hrp* box found	Beginning disrupted by mobile element protein
hopAO1	5260	303	No *hrp* box found	Beginning disrupted by mobile element protein
hopAO1	5300	220	No *hrp* box found	Truncated beginning of gene
hopAO1	9326	466	GGAACCCCACAAGCATTTAAGACCACGTA	
hopAO1	9629	466	GGAACCCCACAAGCATTTAAGACCACGTA	Contig break: Sequencing confirmed full gene
hopAO1	9657	466	GGAACCCCACAAGCATTTAAGACCACGTA	
hopAO1	leaf	303	No *hrp* box found	Beginning disrupted by mobile element protein
hopAR1	5244	267	GGAACCGGGAGGGTAATTAGGACACTCA	Transfer *Psm* R2 and *Pss* phylogroup 2
hopAR1	5255	267	GGAACCGAATGGCTCATCTGGACACTCA	
hopAR1	5260	267	GGAACCGAATGGCTCATCTGGACACTCA	
hopAR1	5300	267	GGAACCGGGAGGGTAATTAGGACACTCA	
hopAR1	9097	267	GGAACCGAATGGCTCATCTGGACACTCA	
hopAR1	9326	267	GGAACCGGGAGGGTAATTAGGACACTCA	
hopAR1	9629	267	GGAACCGGGAGGGTAATTAGGACACTCA	
hopAR1	9630	267	GGAACCGAATGGCTCATCTGGACACTCA	
hopAR1	9644	267	GGAACCGAATGGCTCATCTGGACACTCA	
hopAR1	9646	267	GGAACCGGGAGGGTAATTAGGACACTCA	
hopAR1	9654	267	GGAACCGAATGGCTCATCTGGACACTCA	
hopAR1	9657	267	GGAACCGGGAGGGTAATTAGGACACTCA	
hopAR1	9659	267	GGAACCGAATGGCTCATCTGGACACTCA	

Effector	Strain	Protein length	*hrp* box	Comments
hopAR1	leaf	267	GGAACCGAATGGCTCATCTGGACACTCA	
hopAR1	sc214	267	GGAACCGAATGGCTCATCTGGACACTCA	
hopAS1	5244	1361	GGAACCGAATCCACATATCGACCACCCA	
hopAS1	5255	1358	GGAACCGAATCCACATTTCGACCACCCA	
hopAS1	5260	1358	GGAACCGAATCCACATTTCGACCACCCA	
hopAS1	5300	1361	GGAACCGAATCCACATATCGACCACCCA	
hopAS1	9326	1361	GGAACCGAATCCACATATCGACCACCCA	
hopAS1	9629	1361	GGAACCGAATCCACATATCGACCACCCA	
hopAS1	9643	1362	GGAACCGAATCCATATTTCGACCACCCA	
hopAS1	9646	1361	GGAACCGAATCCACATATCGACCACCCA	
hopAS1	9657	1361	GGAACCGAATCCACATATCGACCACCCA	
hopAS1	RMA1	1364	GGAACCGAATCCACATTTCGACCACCCA	
hopAS1	leaf	1358	GGAACCGAATCCACATTTCGACCACCCA	
hopAS1	sc214	1358	GGAACCGAATCCACATTTCGACCACCCA	
hopAT1	5244	82	GGAACCAAATCGCTGGCTACGCTAACCAA	Possible transfer *Psm* R1 and R2
hopAT1	5255	82	GGAACCAAATCGCTGGCTACGCTAACCA	
hopAT1	5260	82	GGAACCAAATCGCTGGCTACGCTAACCA	
hopAT1	9646	82	GGAACCAAATCGCTGGCTACGCTAACCAA	
hopAT1	leaf	82	GGAACCAAATCGCTGGCTACGCTAACCA	
hopAT1	sc214	82	GGAACCAAATCGCTGGCTACGCTAACCA	
hopAU1	5244	731	GGAACCCTCCTGTGATTTCGAACACTCA	
hopAU1	5255	731	GGAACCCTCCTGTGATTTTCGAACACTCA	
hopAU1	5260	731	GGAACCCTCCTGTGATTTTCGAACACTCA	
hopAU1	5300	731	GGAACCCTCCTGTGATTTCGAACACTCA	
hopAU1	9326	731	GGAACCCTCCTGTGATTTCGAACACTCA	
hopAU1	9629	731	GGAACCCTCCTGTGATTTCGAACACTCA	
hopAU1	9646	731	GGAACCCTCCTGTGATTTCGAACACTCA	

Effector	Strain	Protein length	*hrp* box	Comments
hopAU1	9657	731	GGAACCCTCCTGTGATTTCGAACACTCA	
hopAU1	leaf	731	GGAACCCTCCTGTGATTTTCGAACACTCA	
hopAU1	sc214	731	GGAACCCTCCTGTGATTTTCGAACACTCA	
hopAV1	5244	806	GGAACCCTCCTGTGCTTTTTGAACACTCA	
hopAV1	5300	473	No *hrp* box found	Truncated at beginning
hopAV1	9326	473	No *hrp* box found	Truncated at beginning
hopAV1	9629	473	No *hrp* box found	Truncated at beginning
hopAV1	9646	806	GGAACCCTCCTGTGCTTTTTGAACACTCA	
hopAV1	9657	473	No *hrp* box found	Truncated at beginning
hopAW1	9097	220	GGAACTGATAGAAGGTCTCGACCACTCA	Stop codon at beginning
hopAW1	9630	220	GGAACTGATAGAAGGTCTCGACCACTCA	Stop codon at beginning
hopAW1	9644	195	GGAACTGATAGAAGGTCTCGACCACTCA	Insertion at start
hopAW1	9654	195	GGAACTGATAGAAGGTCTCGACCACTCA	Insertion at start
hopAW1	9659	195	GGAACTGATAGAAGGTCTCGACCACTCA	Insertion at start
hopAY1	5244	323	GGAACTTTTTCTTGCTCGTTGCCACACA	Possible transfer *Psm* R1 and R2
hopAY1	5255	323	GGAACTTTTTCTTGCTCGTTGCCACACA	
hopAY1	5260	323	GGAACTTTTTCTTGCTCGTTGCCACACA	
hopAY1	5300	207	GGAACTTTTTCTTGCTCGTTGCCACACA	Frameshift, pseudogenisation
hopAY1	9326	323	GGAACTTTTTCTTGCTCGTTGCCACACA	
hopAY1	9629	207	GGAACTTTTTCTTGCTCGTTGCCACACA	Frameshift, pseudogenisation
hopAY1	9646	323	GGAACTTTTTCTTGCTCGTTGCCACACA	
hopAY1	9657	323	GGAACTTTTTCTTGCTCGTTGCCACACA	
hopAY1	leaf	323	GGAACTTTTTCTTGCTCGTTGCCACACA	
hopAY1	sc214	323	GGAACTTTTTCTTGCTCGTTGCCACACA	
hopAZ1	5244	229	GGAACCTCTCCTCAATGAGTTGCCACTCA	
hopAZ1	5255	217	AGAACCGCCGAATCGGCACCCGCTACGA	
hopAZ1	5260	217	AGAACCGCCGAATCGGCACCCGCTACGA	

Effector	Strain	Protein length	*hrp* box	Comments
hopAZ1	5300	229	GGAACCTCTCCTCAATGAGTTGCCACTCA	
hopAZ1	9293	217	GGAACTGCTCCTCAATGAGTTGCCACTCA	
hopAZ1	9326	229	GGAACCTCTCCTCAATGAGTTGCCACTCA	
hopAZ1	9629	229	GGAACCTCTCCTCAATGAGTTGCCACTCA	
hopAZ1	9644	219	GGAACCATTCTTCTACTCATAGCCACTCA	
hopAZ1	9646	229	GGAACCTCTCCTCAATGAGTTGCCACTCA	
hopAZ1	9654	219	GGAACCATTCTTCTACTCATAGCCACTCA	
hopAZ1	9656	217	GGAACTGCTCCTCAATGAGTTGCCACTCA	
hopAZ1	9657	229	GGAACCTCTCCTCAATGAGTTGCCACTCA	
hopAZ1	9659	219	AGAACCGCCGAATCGGCACCCGCTACGA	
hopAZ1	leaf	217	AGAACCGCCGAATCGGCACCCGCTACGA	
hopB1	9643	465	GAAACGTCGGTGCCCGGGCAGCTTCACG	
hopBB1	5255	280	GGAACTTCAATAGGGTGTCGTACCACTCA	Horizontal transfer *Psm* R1 and R2
hopBB1	5260	280	GGAACTTCAATAGGGTGTCGTACCACTCA	
hopBB1	5300	280	GGAACTTCAATAGGGTGTCGTACCACTCA	
hopBB1	9326	280	GGAACTTCAATAGGGTGTCGTACCACTCA	
hopBB1	9629	280	GGAACTTCAATAGGGTGTCGTACCACTCA	
hopBB1	9646	280	GGAACTTCAATAGGGTGTCGTACCACTCA	
hopBB1	9657	280	GGAACTTCAATAGGGTGTCGTACCACTCA	
hopBB1	leaf	280	GGAACTTCAATAGGGTGTCGTACCACTCA	
hopBB1	sc214	280	GGAACTTCAATAGGGTGTCGTACCACTCA	
hopBD1	5244	300	GGAACCGGTCGAGGGGTTCTGACCACATA	Horizontal transfer *Psm* R1 and R2
hopBD1	5255	300	GGAACCGGTCGAGGGGTTCTGACCACATA	
hopBD1	5260	300	GGAACCGGTCGAGGGGTTCTGACCACATA	
hopBD1	9646	300	GGAACCGGTCGAGGGGTTCTGACCACATA	
hopBD1	leaf	300	GGAACCGGTCGAGGGGTTCTGACCACATA	
hopBD1	sc214	300	GGAACCGGTCGAGGGGTTCTGACCACATA	

Effector	Strain	Protein length	*hrp* box	Comments
hopBE1	9097	588	GGAACCCGATCCATCCGCCGAGCCACTCA	
hopBE1	9630	588	GGAACCCGATCCATCCGCCGAGCCACTCA	
hopBE1	9644	588	GGAACCCGATCCATCCGCCGAGCCACTCA	
hopBF1	5244	192	No *hrp* box found	Possible transfer R1 and R2 via mobile elements
hopBF1	9293	201	GGAACCCAACTCACTCAATTCATCACTCA	
hopBF1	9646	192	No *hrp* box found	
hopBF1	9656	201	GGAACCCAACTCACTCAATTCATCACTCA	
hopBF1	leaf	192	No *hrp* box found	
hopBH1	RMA1	421	GGAACCGATCCTGGCTGCGTGATCACCGA	
hopBK1	5244	55	GCCACATGTTCCTTTTGTTCTACTACTGA	Missing beginning of gene
hopBK1	5300	55	GCCACATGTTCCTTTTGTTCTACTACTGA	Missing beginning of gene
hopBK1	9326	55	GCCACATGTTCCTTTTGTTCTACTACTGA	Missing beginning of gene
hopBK1	9629	55	GCCACATGTTCCTTTTGTTCTACTACTGA	Missing beginning of gene
hopBK1	9646	55	GCCACATGTTCCTTTTGTTCTACTACTGA	Missing beginning of gene
hopBK1	9657	55	GCCACATGTTCCTTTTGTTCTACTACTGA	Missing beginning of gene
hopBL2	5244	752	GGAATTTAAGCTCGATAGTTGCCACAGT	
hopBL2	5300	538	GGAATTTAAGCTCGATAGTTGCCACAGT	Interrupted by mobile element
hopBL2	9326	752	GGAATTTAAGCTCGATAGTTGCCACAGT	
hopBL2	9629	752	GGAATTTAAGCTCGATAGTTGCCACAGT	
hopBL2	9646	752	GGAATTTAAGCTCGATAGTTGCCACAGT	
hopBL2	9657	752	GGAATTTAAGCTCGATAGTTGCCACAGT	
hopC1	RMA1	269	GGAACCAATTCGCTTCTAAAACCACTCA	
hopD1	5244	705	GGAACCCAAGAGCTCTTGCGACCACACA	
hopD1	5255	705	GGAACCCAAGAGCCCTTGTGACCACACA	Stop codon in middle
hopD1	5260	705	GGAACCCAAGAGCCCTTGTGACCACACA	Stop codon in middle
hopD1	5300	705	GGAACCCAAGAGCTCTTGCGACCACACA	

Effector	Strain	Protein length	*hrp* box	Comments
hopD1	9326	705	GGAACCCAAGAGCTCTTGCGACCACACA	Contig break at beginning
hopD1	9629	705	GGAACCCAAGAGCTCTTGCGACCACACA	Contig break at beginning
hopD1	9646	705	GGAACCCAAGAGCTCTTGCGACCACACA	
hopD1	9657	705	GGAACCCAAGAGCTCTTGCGACCACACA	Contig break at beginning
hopD1	leaf	705	GGAACCCAAGAGCCCTTGTGACCACACA	Stop codon in middle
hopD1	sc214	705	GGAACCCAAGAGCCCTTGTGACCACACA	
hopE1	5255	211	GGAACCGAATCCACCTCAAAGTCCACCCA	
hopE1	5260	211	GGAACCGAATCCACCTCAAAGTCCACCCA	
hopE1	leaf	211	GGAACCGAATCCACCTCAAAGTCCACCCA	
hopE1	sc214	211	GGAACCGAATCCACCTCAAAGTCCACCCA	
hopF2	5255	203	GGAACCTGATACCTCTCGATGACCACCCA	
hopF2	5260	203	GGAACCTGATACCTCTCGATGACCACCCA	
hopF2	9643	203	GGAACCTGATACCTCTCGATGACCACCCA	
hopF2	RMA1	203	GGAACCTGATACCTCTCGATGACCACCCA	
hopF2	leaf	203	GGAACCTGATACCTCTCGATGACCACCCA	
hopF2	sc214	203	GGAACCTGATACCTCTCGATGACCACCCA	
hopF3	5244	195	GGAACTAAACCCATAATATCGACCACACA	
hopF3	5300	198	GGAACTAAACCCATAATATCGACCACACA	
hopF3	9326	198	GGAACTAAACCCATAATATCGACCACACA	
hopF3	9629	198	GGAACTAAACCCATAATATCGACCACACA	
hopF3	9646	195	GGAACTAAACCCATAATATCGACCACACA	
hopF3	9657	198	GGAACTAAACCCATAATATCGACCACACA	
hopF4	5244	203	GGAACCAATTACATAGCTGTGACCACTCA	
hopF4	5255	203	GGAACCAGAAGCCTAGGTATGACCACTGA	
hopF4	5260	203	GGAACCAGAAGCCTAGGTATGACCACTGA	
hopF4	9646	203	GGAACCAATTACATAGCTGTGACCACTCA	

Effector	Strain	Protein length	*hrp* box	Comments
hopF4	leaf	203	GGAACCAGAAGCCTAGGTATGACCACTGA	
hopF4	sc214	203	GGAACCAGAAGCCTAGGTATGACCACTGA	
hopG1	5300	493	GGAACTCTACGCTTGGCGACGACCACGTA	
hopG1	9326	493	GGAACTCTACGCTTGGCGACGACCACGTA	
hopG1	9629	493	GGAACTCTACGCTTGGCGACGACCACGTA	
hopG1	9657	493	GGAACTCTACGCTTGGCGACGACCACGTA	
hopH1	5255	218	GGAACCGAATCCATCTCGAGGGCCACTCA	
hopH1	5260	218	GGAACCGAATCCATCTCGAGGGCCACTCA	
hopH1	9097	218	GGAACCGAACCCATCTCAAGGGCCACTCA	
hopH1	9630	218	GGAACCGAACCCATCTCAAGGGCCACTCA	
hopH1	9644	218	GGAACCGAACCCATCTCAAGGGCCACTCA	
hopH1	9654	218	GGAACCGAACCCATCTCAAGGGCCACTCA	
hopH1	9659	218	GGAACCGAACCCATCTCAAGGGCCACTCA	
hopH1	leaf	218	GGAACCGAATCCATCTCGAGGGCCACTCA	
hopI1	5244	336	GGAACCAGATCCTGTTGCTTGCCACCAA	
hopI1	5255	413	GGAACCAGATCTCGTTGCTTGCCACAAA	
hopI1	5260	413	GGAACCAGATCTCGTTGCTTGCCACAAA	
hopI1	5300	336	GGAACCAGATCCTGTTGCTTGCCACCAA	
hopI1	9097	336	GGAACCGGATCTCGTTGCTTGCCACCAA	
hopI1	9293	336	GGAACCGGATCTCGTTGCTTGCCACCAA	
hopI1	9326	336	GGAACCAGATCCTGTTGCTTGCCACCAA	
hopI1	9629	336	GGAACCAGATCCTGTTGCTTGCCACCAA	
hopI1	9630	336	GGAACCGGATCTCGTTGCTTGCCACCAA	
hopI1	9643	410	GGAACCAGATCTCGTTGCTTGCCACCAA	
hopI1	9644	336	GGAACCGGATCTCGTTGCTTGCCACCAA	
hopI1	9646	336	GGAACCAGATCCTGTTGCTTGCCACCAA	
hopI1	9654	336	GGAACCGGATCTCGTTGCTTGCCACCAA	

Effector	Strain	Protein length	*hrp* box	Comments
hopI1	9656	336	GGAACCGGATCTCGTTGCTTGCCACCAA	
hopI1	9657	336	GGAACCAGATCCTGTTGCTTGCCACCAA	
hopI1	9659	336	GGAACCGGATCTCGTTGCTTGCCACCAA	
hopI1	RMA1	286	GGAACCAGATCTCGTTGCTTGCCACCAA	
hopI1	leaf	413	GGAACCAGATCTCGTTGCTTGCCACAAA	
hopM1	5244	528	No *hrp* box found	Missing beginning of gene, due to inversion
hopM1	5255	712	GGAACTGAAATGCCTATGCCTGCGACTCA	
hopM1	5260	712	GGAACTGAAATGCCTATGCCTGCGACTCA	
hopM1	5300	633	GGAACCAAAACGCCTTTACATTCGACTCA	Missing beginning of gene due to stop codon
hopM1	9097	718	GGAACCAAAACCCCTTGCCCTCCGACTCA	
hopM1	9293	715	GGAACCGACAACGCCAGGCAGCCACTCA	
hopM1	9326	633	GGAACCAAAACGCCTTTACATTCGACTCA	Missing beginning of gene due to stop codon
hopM1	9629	633	GGAACCAAAACGCCTTTACATTCGACTCA	Missing beginning of gene due to stop codon
hopM1	9630	718	GGAACCAAAACCCCTTGCCCTCCGACTCA	
hopM1	9643	712	GGAACTGAAATGCCTATGCCTGCGACTCA	
hopM1	9644	718	GGAACCAAAACCCCTTGCCCTCCGACTCA	
hopM1	9646	528	No *hrp* box found	Missing beginning of gene, due to inversion
hopM1	9654	718	GGAACCAAAACCCCTTGCCCTCCGACTCA	
hopM1	9656	715	GGAACCGACAACGCCAGGCAGCCACTCA	
hopM1	9657	633	GGAACCAAAACGCCTTTACATTCGACTCA	Missing beginning of gene due to stop codon
hopM1	9659	718	GGAACCAAAACCCCTTGCCCTCCGACTCA	
hopM1	RMA1	712	GGAACTGAAATGCCTATGCCTGCGACTCA	
hopM1	leaf	712	GGAACTGAAATGCCTATGCCTGCGACTCA	
hopM1	sc214	712	GGAACTGAAATGCCTATGCCTGCGACTCA	
hopN1	5255	350	GGAACCGCATCACGTCTTGAACCACAGA	
hopN1	5260	350	GGAACCGCATCACGTCTTGAACCACAGA	
hopN1	9643	350	GGAACCGCATCACGTCTTGAACCACAGA	

Effector	Strain	Protein length	*hrp* box	Comments
hopN1	RMA1	350	GGAACCGCATCACGTCTTGAACCACAGA	
hopN1	leaf	350	GGAACCGCATCACGTCTTGAACCACAGA	
hopN1	sc214	350	GGAACCGCATCACGTCTTGAACCACAGA	
hopO1	5255	283	GGAACCTTACGGAGGTTCACGCCACCAA	
hopO1	5260	283	GGAACCTTACGGAGGTTCACGCCACCAA	
hopO1	9643	298	GGAACCGGACGAGGCTTTCGGCCACTCA	
hopQ1	5244	447	GAAACCGAAACGGCGTTGCTTGCCACACA	
hopQ1	5300	447	GAAACCGAAACGGCGTTGCTTGCCACACA	
hopQ1	9326	447	GAAACCGAAACGGCGTTGCTTGCCACACA	Contig break, whole gene confirmed by PCR
hopQ1	9629	447	GAAACCGAAACGGCGTTGCTTGCCACACA	Contig break, whole gene confirmed by PCR
hopQ1	9646	447	GAAACCGAAACGGCGTTGCTTGCCACACA	
hopQ1	9657	447	GAAACCGAAACGGCGTTGCTTGCCACACA	Contig break, whole gene confirmed by PCR
hopR1	5244	1959	GGAACCGATCTGGCTGCACAGCCACTCA	
hopR1	5255	1957	GGAACCGATCTGGCCGCCTGACCACTCA	
hopR1	5260	1957	GGAACCGATCTGGCCGCCTGACCACTCA	
hopR1	5300	1959	GGAACCGATCTGGCTGCACAGCCACTCA	
hopR1	9326	1959	GGAACCGATCTGGCTGCACAGCCACTCA	
hopR1	9629	1959	GGAACCGATCTGGCTGCACAGCCACTCA	
hopR1	9643	1957	GGAACCGATCCGGTTGCTTGGCCACTCA	
hopR1	9646	1959	GGAACCGATCTGGCTGCACAGCCACTCA	
hopR1	9657	1959	GGAACCGATCTGGCTGCACAGCCACTCA	
hopR1	RMA1	1957	GGAACCGATCCGGCTGTCTGACCACTCA	
hopR1	leaf	1957	GGAACCGATCTGGCCGCCTGACCACTCA	
hopR1	sc214	1957	GGAACCGATCTGGCCGCCTGACCACTCA	
hopS1	9643	118	GGAACCGGACGAGGCTTTCGGCCACTCA	
hopS2	5255	177	GGAACCTCGGCGCGTAGTGCGCCACTCA	

Effector	Strain	Protein length	*hrp* box	Comments
hopS2	5260	177	GGAACCTCGGCGCGTAGTGCGCCACTCA	
hopS2	9643	177	GGAACCCCTGCGCGTCCAGCGCCACTCA	
hopS2	RMA1	177	GGAACCGCGGCGCGTCCTTTGCCACTCA	
hopS2	leaf	177	GGAACCTCGGCGCGTAGTGCGCCACTCA	
hopS2	sc214	177	GGAACCTCGGCGCGTAGTGCGCCACTCA	
hopT1	5255	379	TGAACGAGATCGCAGAGAGCTCTGACGA	
hopT1	5260	379	TGAACGAGATCGCAGAGAGCTCTGACGA	
hopT1	9643	379	TGAACGAGATAGCAGATAATTCAGACGA	
hopV1	5244	324	GGAACCACAGCGTGCGTTCCAGCCACTCA	
hopV1	5300	324	GGAACCACAGCGTGCGTTCCAGCCACTCA	
hopV1	9326	324	GGAACCACAGCGTGCGTTCCAGCCACTCA	
hopV1	9629	324	GGAACCACAGCGTGCGTTCCAGCCACTCA	
hopV1	9646	324	GGAACCACAGCGTGCGTTCCAGCCACTCA	
hopV1	9657	324	GGAACCACAGCGTGCGTTCCAGCCACTCA	
hopW1	5255	768	GGAACCGATCTGGGCGGAGCGCCACTCA	
hopW1	5260	768	GGAACCGATCTGGGCGGAGCGCCACTCA	
hopW1	5300	776	GGAACCCACGGGCCCTTGTGGTCACACA	
hopW1	9326	772	GGAATCCACGGGCCCTTGTGGTCACACA	
hopW1	9629	772	GGAATCCACGGGCCCTTGTGGTCACACA	
hopW1	9657	772	GGAATCCACGGGCCCTTGTGGTCACACA	
hopW1	RMA1	774	GGAACCCACGGGCCCTTGTGGCCACACA	
hopW1	leaf	768	GGAACCGATCTGGGCGGAGCGCCACTCA	
hopX1	5300	384	GGAACTGACCCGGCCCAAATGACGACATA	
hopX1	9293	380	No *hrp* box found	
hopX1	9326	379	GGAACTGACCCGGCCCAAATGACGACATA	
hopX1	9629	379	GGAACTGACCCGGCCCAAATGACGACATA	
hopX1	9656	380	No *hrp* box found	

Effector	Strain	Protein length	*hrp* box	Comments
hopX1	9657	379	GGAACTGACCCGGCCCAAATGACGACATA	
hopX1	leaf	379	GCGAGCCAGCACATCTATACCCCACACA	Putative *hrp* box
hopX1	sc214	379	GCGAGCCAGCACATCTATACCCCACACA	Putative *hrp*
hopX2	9643	352	GCGAGCCAGCACATCTATACCCCACACA	
hopY1	5255	287	GGAACTCATTCCCGCGAATCGCCACTCA	
hopY1	5260	287	GGAACTCATTCCCGCGAATCGCCACTCA	
hopY1	9643	286	GGAACTCATTACCGCGAATCGCCACTCA	
hopY1	RMA1	286	GGAACTCATTACCGCGAATCGCCACTCA	
hopY1	leaf	286	GGAACTCATTCCCGCGAATCGCCACTCA	
hopY1	sc214	286	GGAACTCATTCCCGCGAATCGCCACTCA	
hopZ3	9643	206	GTTACTGCGTGGCCATTCTGCTCCACTCA	
hopZ4	5244	335	GGAACCTTTTTTTGGTGAAAACCACTGA	
hopZ4	9646	335	GGAACCTTTTTTTGGTGAAAACCACTGA	

Table S4.1: Effectors identified in the genome assemblies of* Prunus*-isolates and RMA1 sequenced in this study. *Effector name, length of the putatively produced protein and identified upstream* hrp*-box promoters are listed. Comments on pseudogenisation, horizontal gene transfers and contig breaks in effector genes are also listed.*

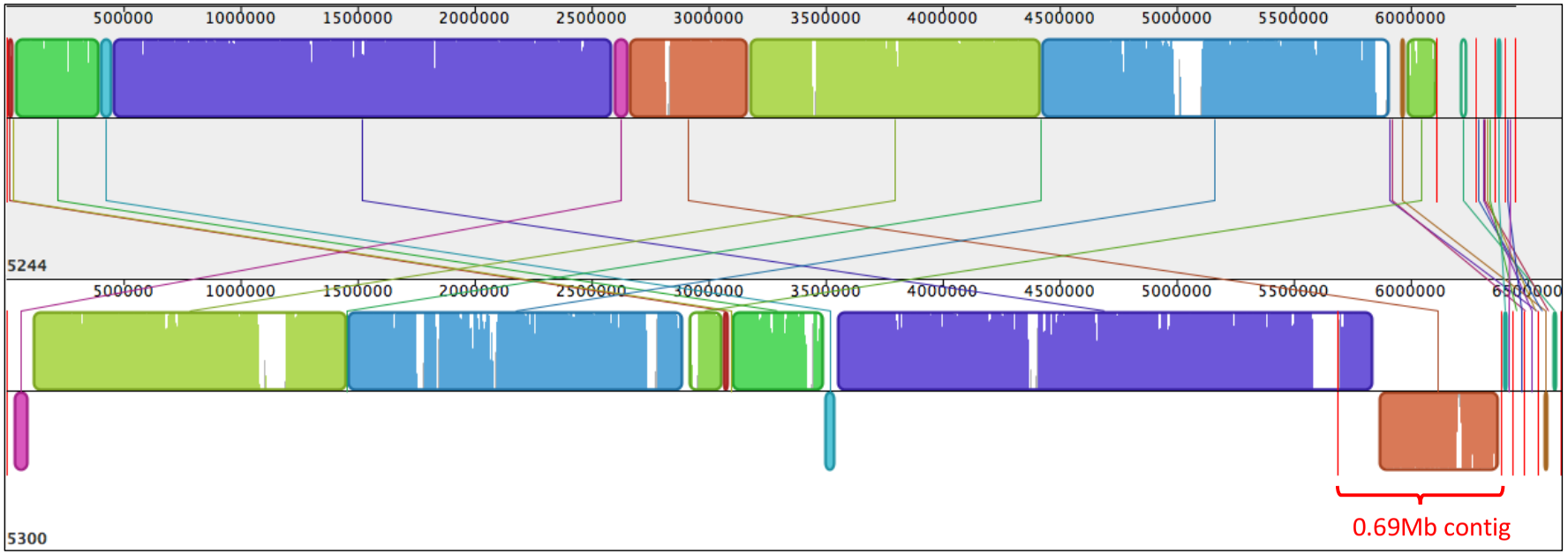


Figure S4.1: Whole-genome alignment of R1-5244 and R1-5300 using ProgressiveMauve. *Different coloured blocks indicate homologous sequences between the two genomes (Locally Co-linear Blocks, LCBs). White shading indicates where this homology is lost. LCBs are connected with lines between the two genomes. Where the LCB appears underneath the horizontal line it is inverted in comparison to the other genome. Contig breaks are indicated by vertical red lines. The 0.69Mb contig of R1-5300 is homologous to two LCBs found in the chromosomal contig of the R1-5244 assembly.*

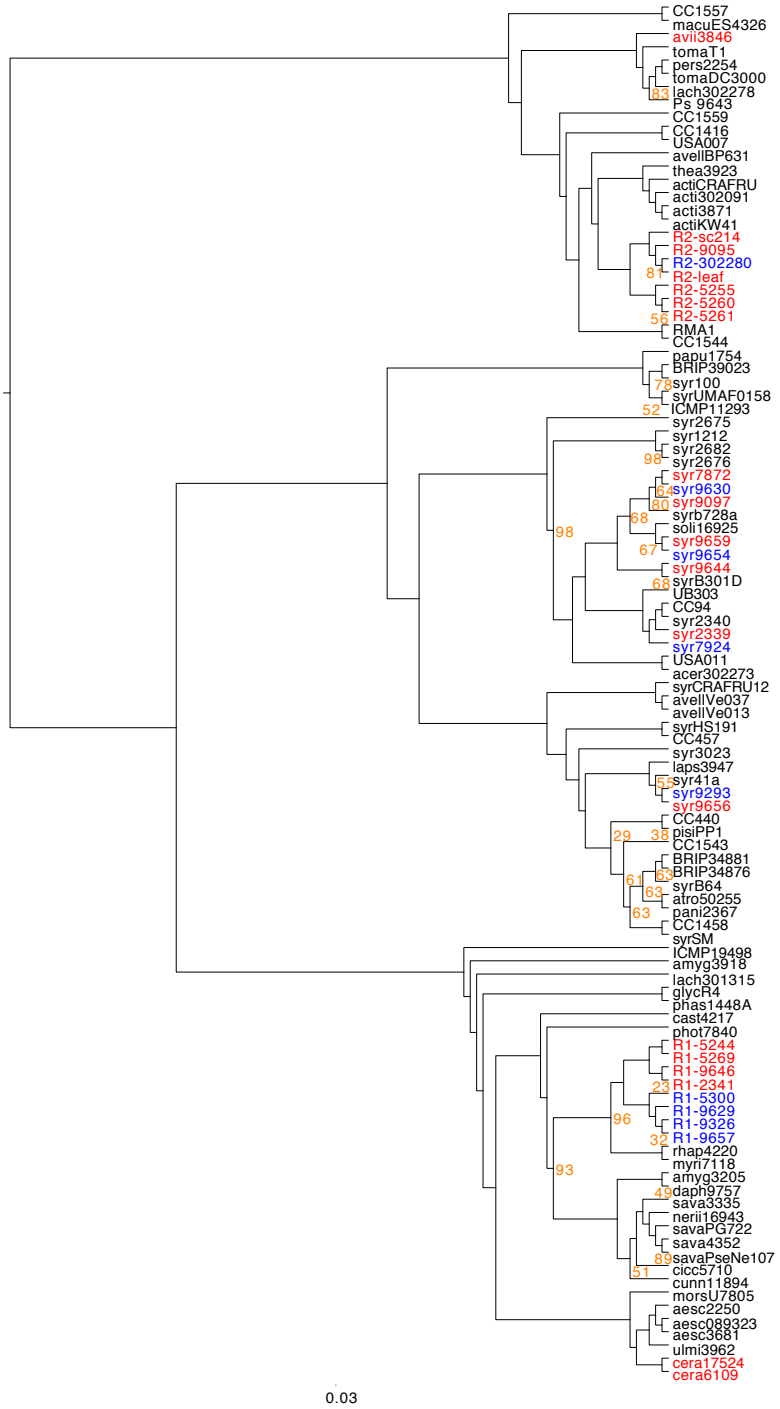


***Figure S4.2: ML phylogeny based on 846 genes from 104 strains, which represent the core genome of* P. syringae.** *The tree was built using RaxML. Isolates from* Prunus *are highlighted dependent on host of isolation with cherry isolates in red and plum isolates in blue. R1-9657 was originally isolated from cherry but due to pathogenicity tests showing it to be non-pathogenic it is coloured in blue as a non-cherry pathogen within* Psm *R1. Support values below 99% are shown for each node.*

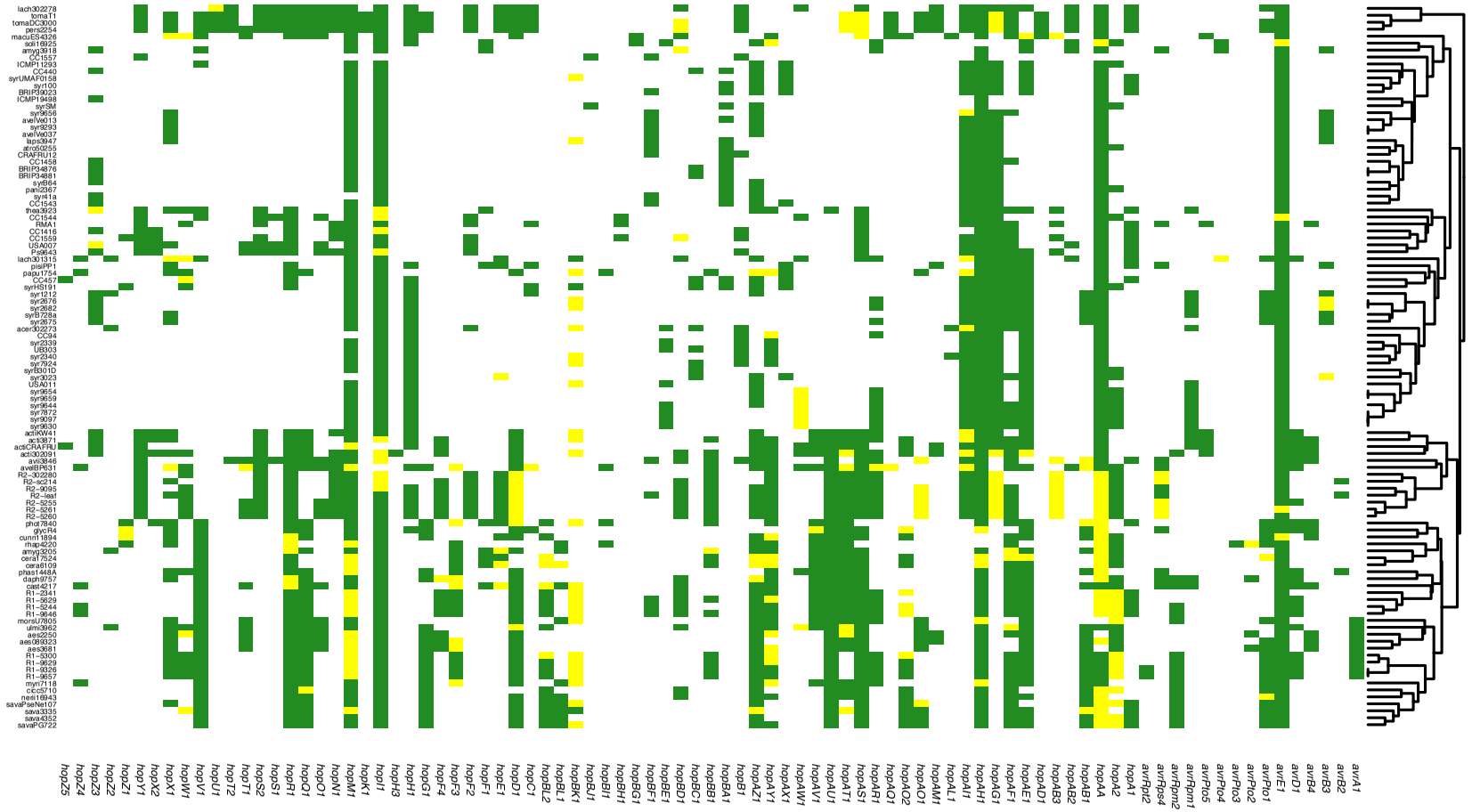


Figure S4.3: Effector presence and absence heatmaps for all analysed* P. syringae *strains. *The heatmap was generated using the R gplots library. The green squares indicate presence of a full-length homolog of the effector gene whereas yellow squares indicate that the gene is disrupted or truncated in some way. The dendrogram clustered strains by similarity in effector gene presence and absence.*

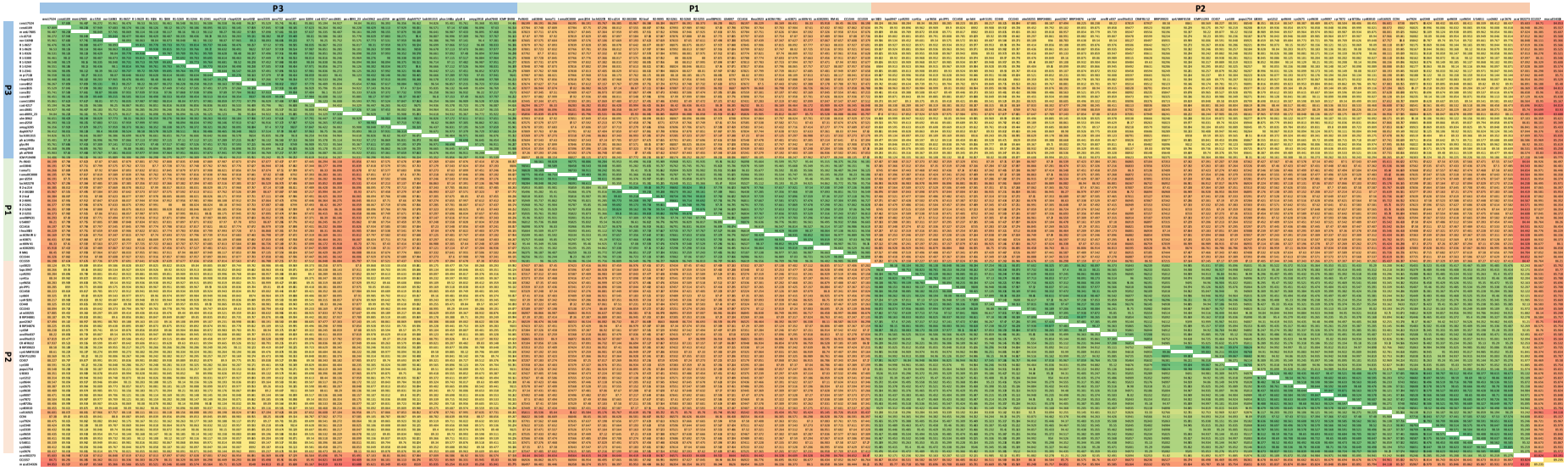


Figure S4.4: Heatmap of average nucleotide identity (ANI) in core genome alignment of 104* P. syringae *strains. *The heatmap was generated using the Geneious and is based on ANI for all pairwise combinations. The different phylogroups 1-3 are labelled. Green squared indicate over 95% ANI. From yellow to red ANI decreases. White squared appear when comparing a strain with itself. The phylogroups represent separate species based on an ANI of >95%, whilst strains in different phylogroups were distantly related. At the end of the table, the divergent strains macuES4326 and CC1557 are included for comparison.*

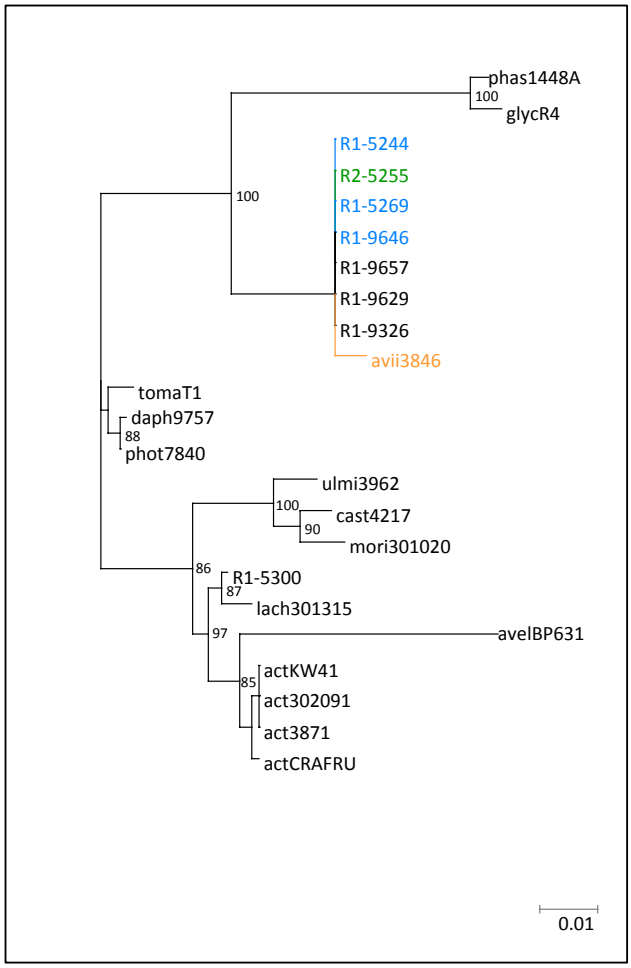


Figure S4.5: ML phylogenetic tree based on the* avrD1 *gene generated using FastTree. *Cherry and plum isolates are highlighted and coloured dependent on clade:* Psm *R1 pathogens: blue,* Psm *R2: green and other cherry pathogens (*P.s *pv.* avii *and* P.s *pv.* cerasicola*) in orange. Support values below 99% are shown for each node. The scale bar shows substitutions per site.*

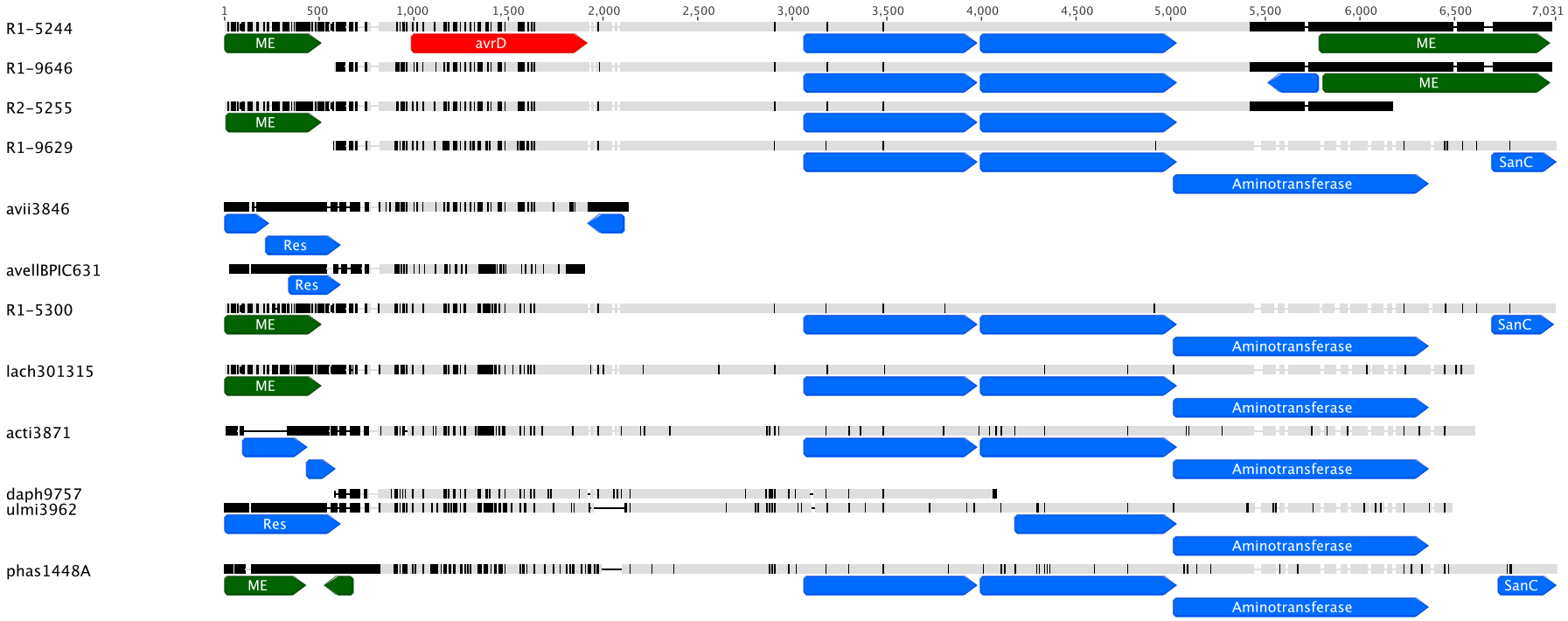


Figure S4.6: Alignment of the DNA region surrounding the avrD1 gene in* Prunus *strains and out-groups. *Grey indicates sequence identity whereas black indicates divergence. Homology between different* Prunus*-infecting clades indicates a horizontal gene transfer may have occurred. The effector gene is coloured in red, whereas other CDS are in blue, mobile element genes are in green. Where the gene annotation was too long to be clearly shown for other CDS genes, it is not presented.*

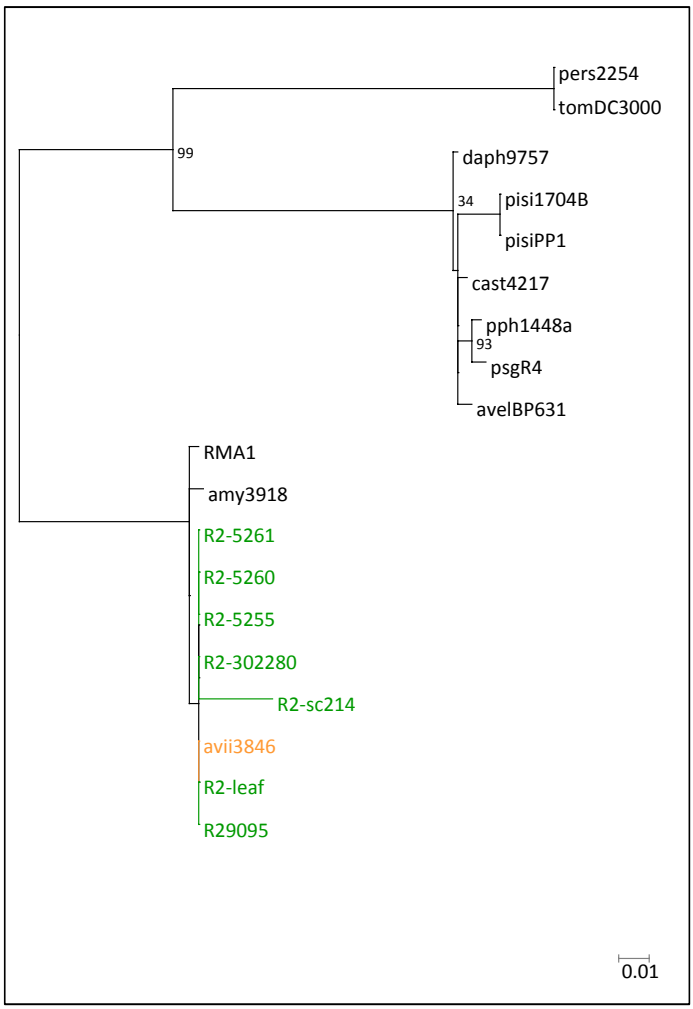


Figure S4.7: ML phylogenetic tree based on the* avrRps4 *gene generated using FastTree. *Cherry and plum isolates are highlighted and coloured dependent on clade:* Psm *R2: green and the other cherry pathogens (*P.s *pv.* avii*) in orange. Support values below 99% are shown for each node. The scale bar shows substitutions per site.*

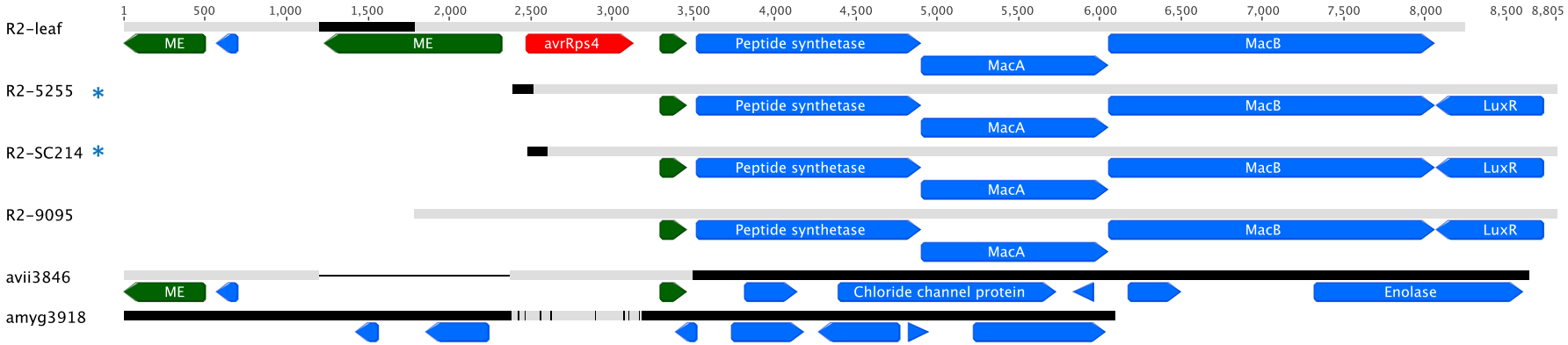


Figure S4.8: Alignment of the DNA region surrounding the* avrRps4 *gene in* Prunus *strains and out-groups. *Grey indicates sequence identity whereas black indicates divergence. Homology between different* Prunus*-infecting clades indicates a horizontal gene transfer may have occurred. *: The gene is disrupted in this strain. The effector gene is coloured in red, whereas other CDS are in blue, mobile element genes are in green. Where the gene annotation was too long to be clearly shown for other CDS genes, it is not presented.*

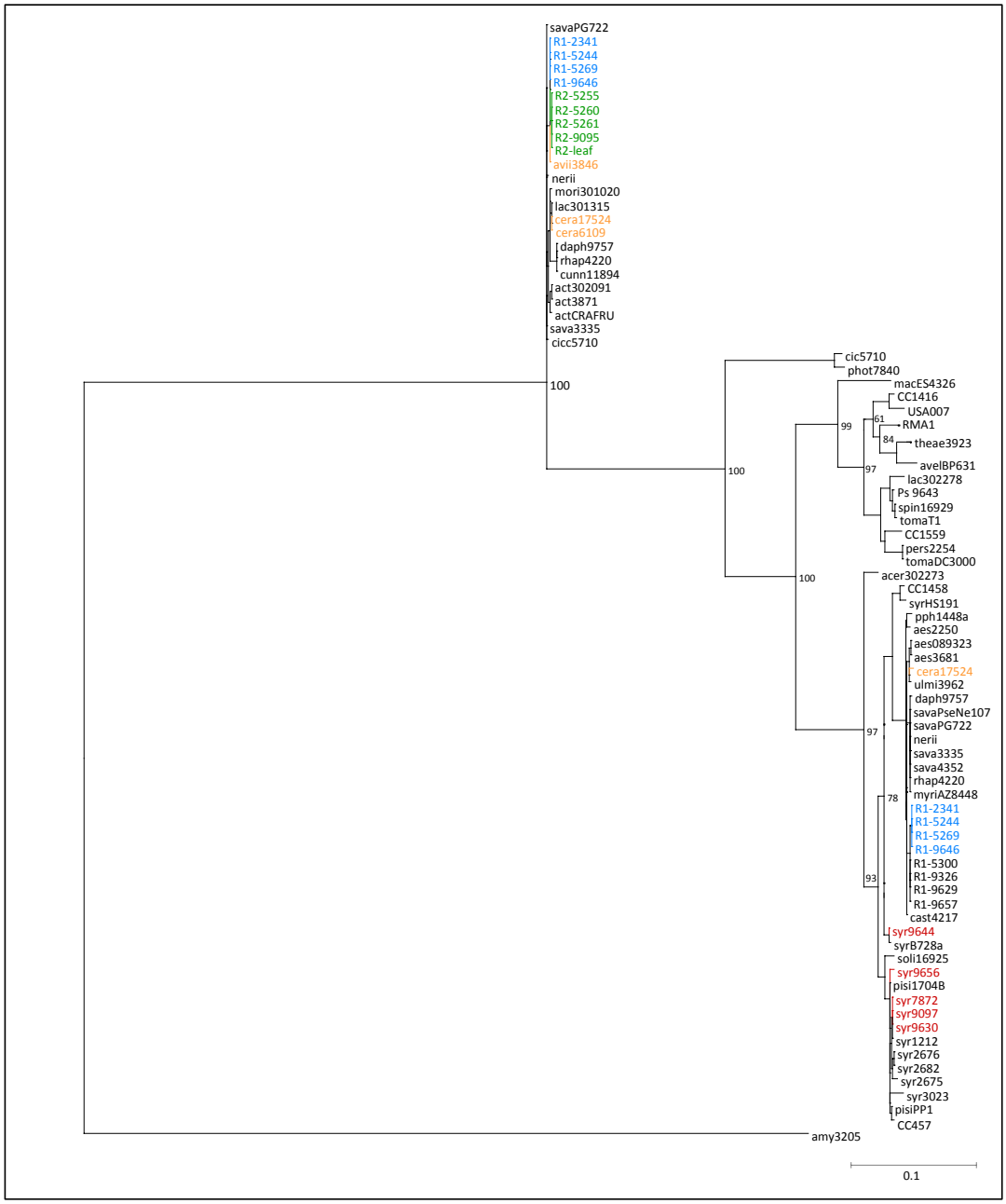


Figure S4.9: ML phylogenetic tree based on the* hopAF1 *gene generated using FastTree. *Cherry and plum isolates are highlighted and coloured dependent on clade:* Psm *R1 pathogens: blue,* Psm *R2: green,* Pss*: red and the other cherry pathogens (*P.s *pv.* avii *and* P.s *pv.* cerasicola*) in orange. Support values below 99% are shown for each node. The scale bar shows substitutions per site.*

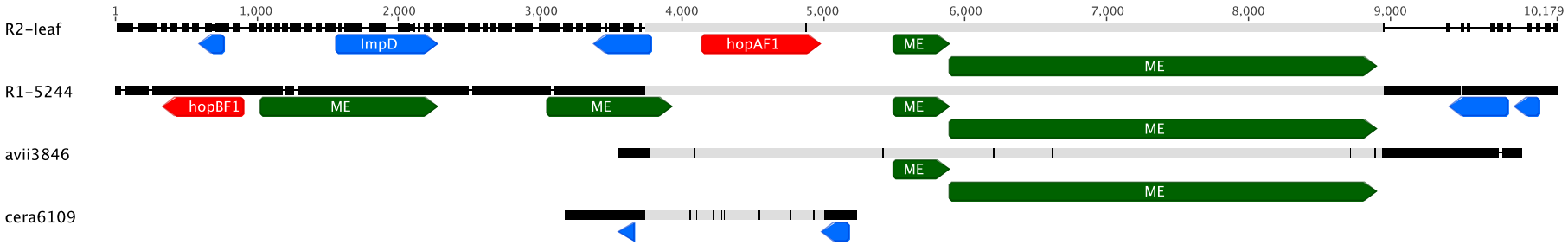


Figure S4.10: Alignment of the DNA region surrounding the* hopAF1 *gene in* Prunus *strains. *Grey indicates sequence identity whereas black indicates divergence. Homology between different* Prunus*-infecting clades indicates a horizontal gene transfer may have occurred. The effector genes are coloured in red, whereas other CDS are in blue, mobile element genes are in green. Where the gene annotation was too long to be clearly shown for other CDS genes, it is not presented.*

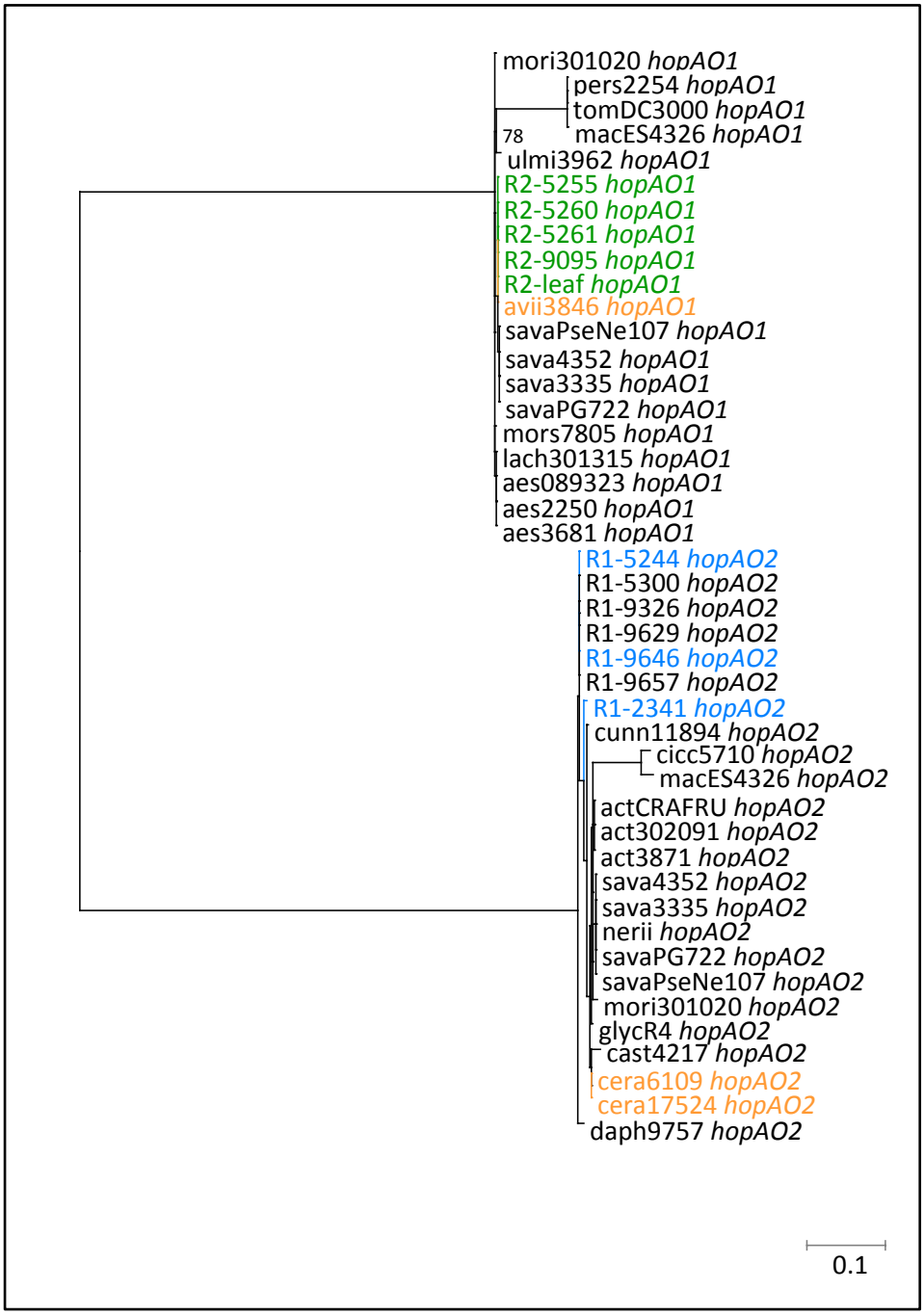


Figure S4.11: ML phylogenetic tree based on the* hopAO1 *gene family generated using FastTree.

Cherry and plum isolates are highlighted and coloured dependent on clade: Psm *R1 pathogens: blue,* Psm *R2: green and the other cherry pathogens (*P.s *pv.* avii *and* P.s *pv.* cerasicola*) in orange. Support values below 99% are shown for each node. The scale bar shows substitutions per site.*

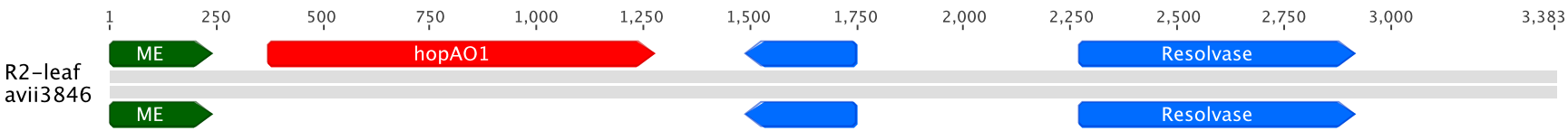


Figure S4.12: Alignment of the DNA region surrounding the* hopAO1 *gene in two* Prunus *strains. *Grey indicates sequence identity whereas black indicates divergence. Homology between the two* Prunus*-infecting strains indicates a horizontal gene transfer may have occurred. The effector gene is coloured in red, whereas other CDS are in blue, mobile element genes are in green. Where the gene annotation was too long to be clearly shown for other CDS genes, it is not presented.*

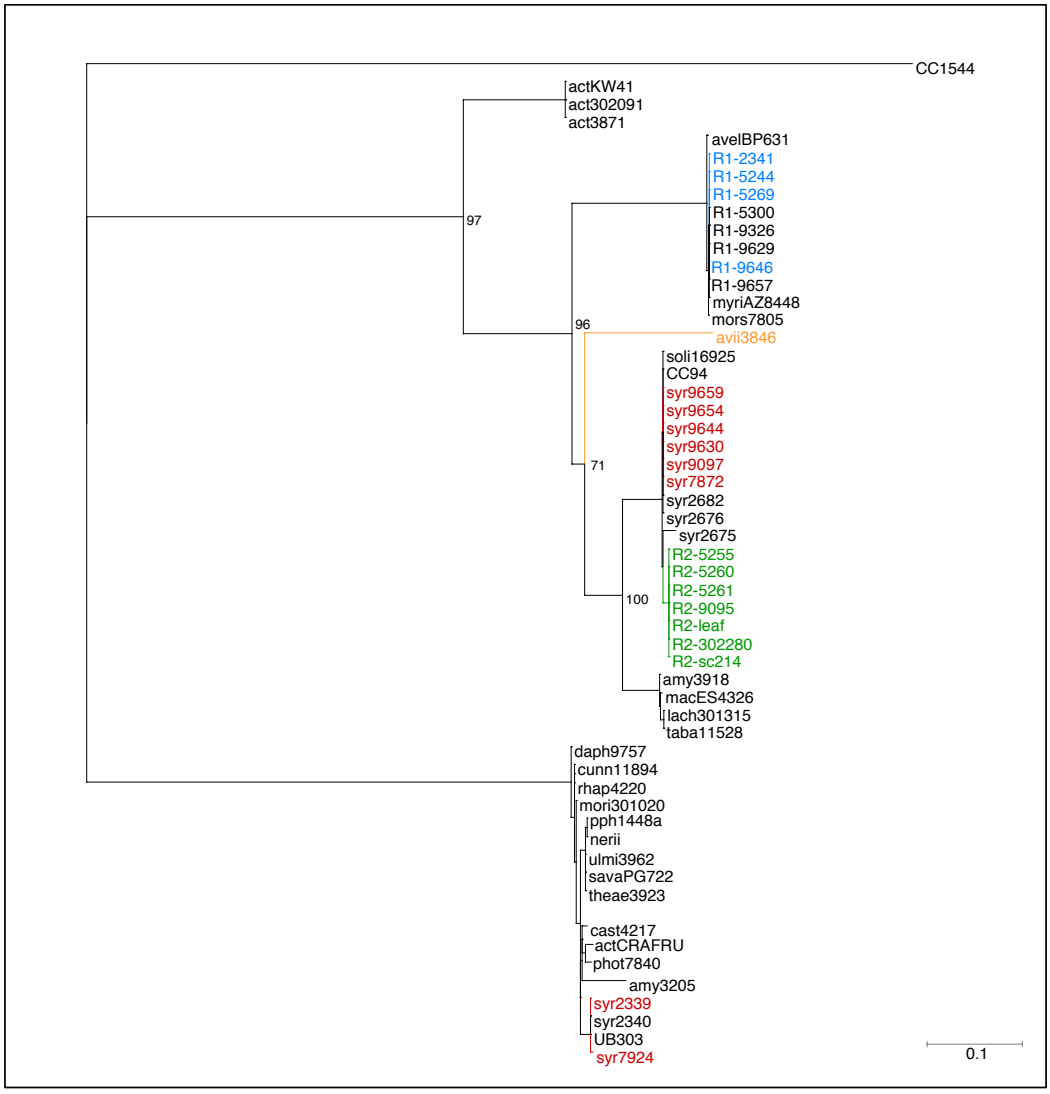


Figure S4.13: ML phylogenetic tree based on the* hopAR1 *gene generated using FastTree. *Cherry and plum isolates are highlighted and coloured dependent on clade:* Psm *R1 pathogens: blue,* Psm *R2: green, Pss: red and the other cherry pathogen (*P.s *pv.* avii*) in orange. Support values below 99% are shown for each node. The scale bar shows substitutions per site.*

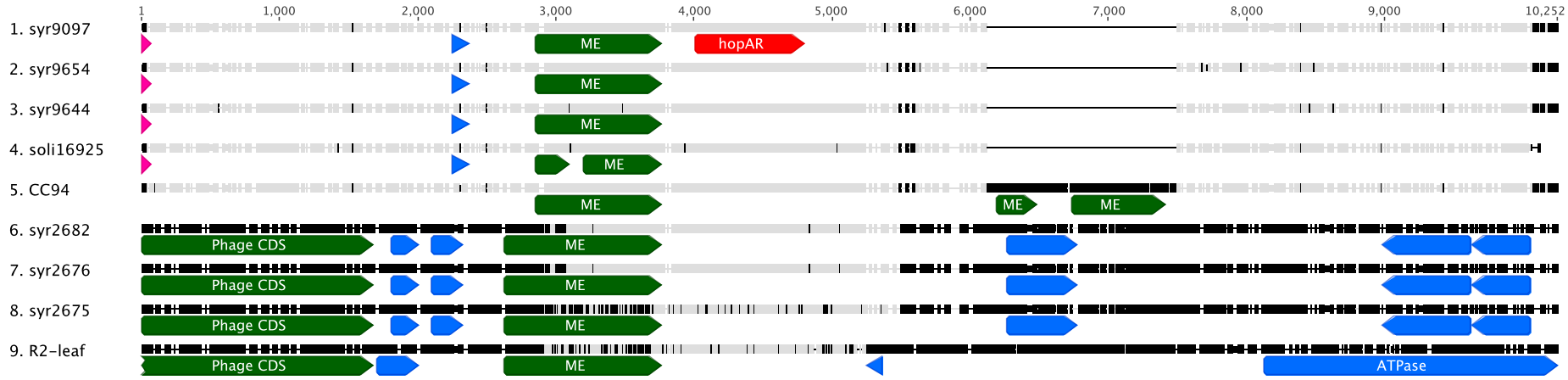


Figure S4.14: Alignment of the DNA region surrounding the* hopAR1 *gene in* Prunus *strains and out-groups. *Grey indicates sequence identity whereas black indicates divergence. Homology between* Psm *R2-leaf and phylogroup 2 strains that infect bean (e.g. syr2675) indicates a horizontal gene transfer may have occurred. The effector gene is coloured in red, whereas other CDS are in blue, mobile element genes are in green. Where the gene annotation was too long to be clearly shown for other CDS genes, it is not presented.*

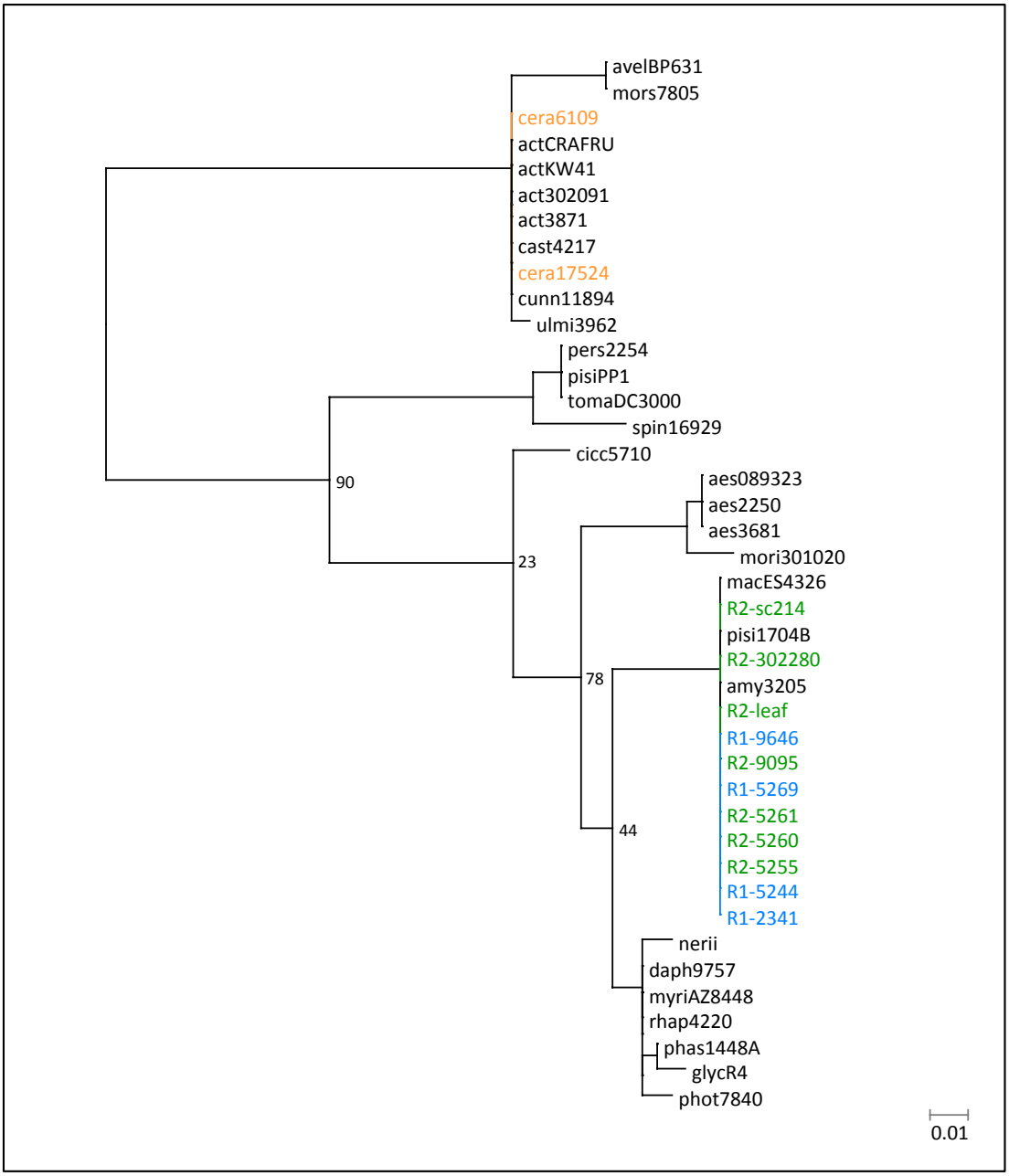


Figure S4.15: ML phylogenetic tree based on the* hopAT1 *gene generated using FastTree. *Cherry and plum isolates are highlighted and coloured dependent on clade:* Psm *R1 pathogens: blue and* Psm *R2: green. Support values below 99% are shown for each node. The scale bar shows substitutions per site.*

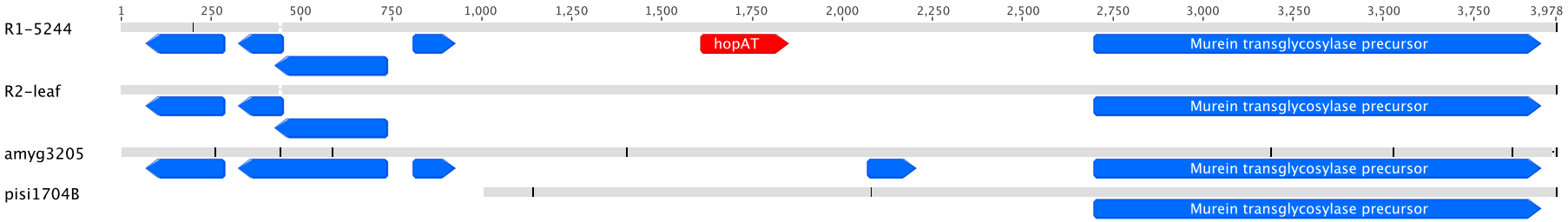


Figure S4.16: Alignment of the DNA region surrounding the* hopAT1 *gene in* Prunus *strains and out-groups. *Grey indicates sequence identity whereas black indicates divergence. Homology between* Prunus*-infecting strains indicates a horizontal gene transfer may have occurred. The effector gene is coloured in red, whereas other CDS are in blue, mobile element genes are in green. Where the gene annotation was too long to be clearly shown for other CDS genes, it is not presented.*

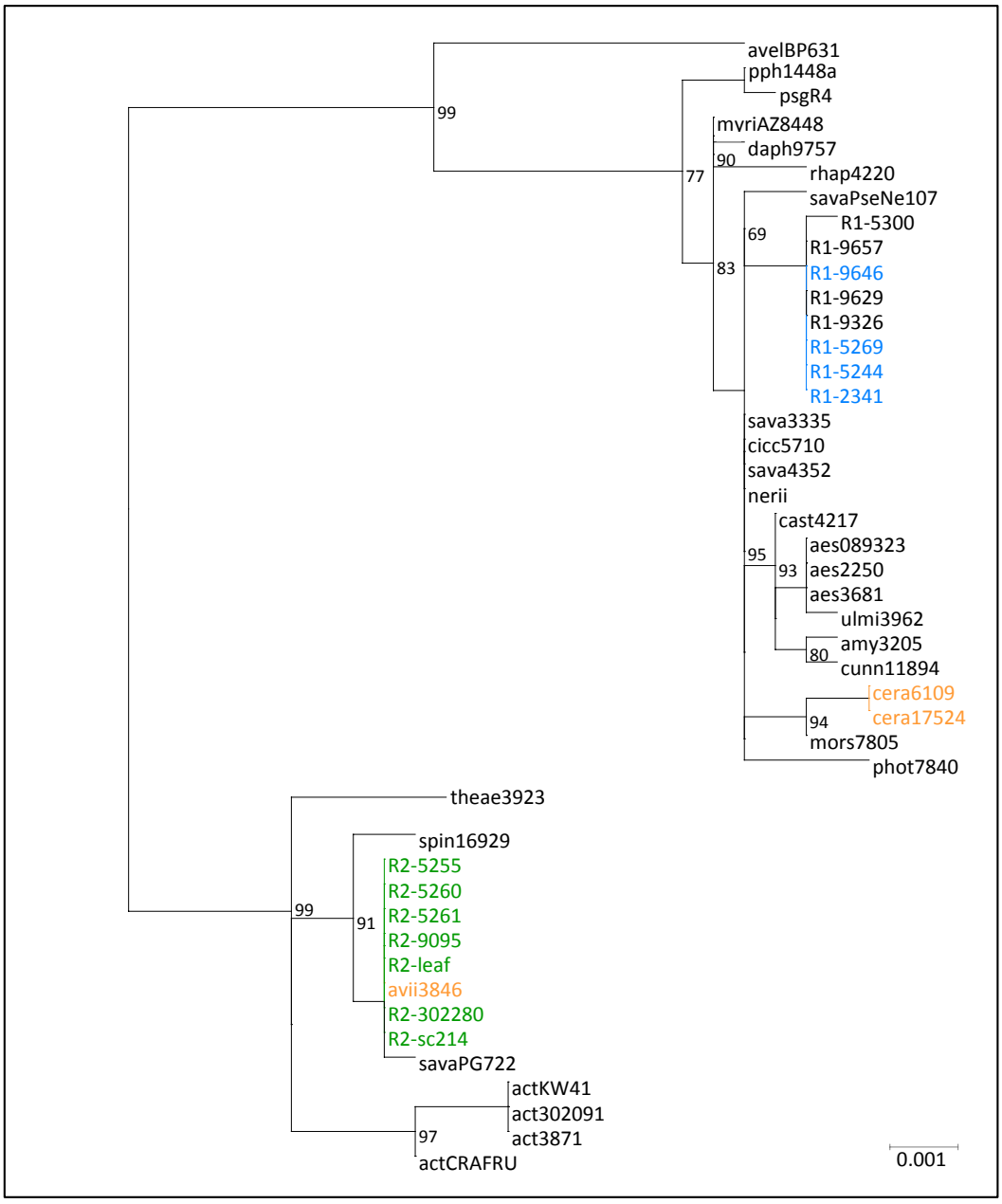


Figure S4.17: ML phylogenetic tree based on the* hopAU1 *gene generated using FastTree. *Cherry and plum isolates are highlighted and coloured dependent on clade:* Psm *R1 pathogens: blue,* Psm *R2: green and the other cherry pathogens (*P.s *pv.* avii *and* P.s *pv.* cerasicola*) in orange. Support values below 99% are shown for each node. The scale bar shows substitutions per site.*

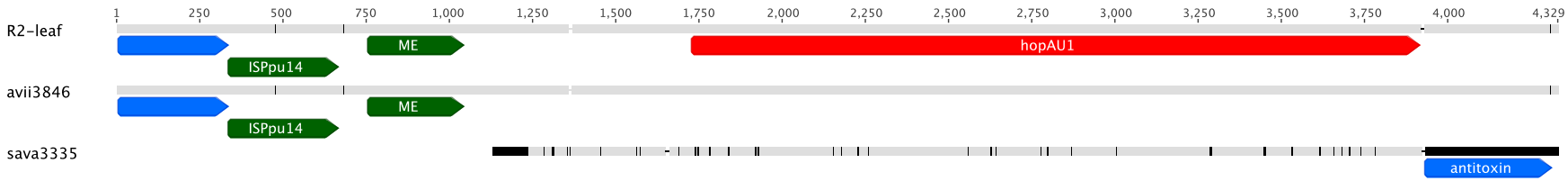


Figure S4.18: Alignment of the DNA region surrounding the* hopAU1 *gene in* Prunus *strains and an out-group. *Grey indicates sequence identity whereas black indicates divergence. Homology between* Prunus*-infecting strains indicates a horizontal gene transfer may have occurred. The effector gene is coloured in red, whereas other CDS are in blue, mobile element genes are in green. Where the gene annotation was too long to be clearly shown for other CDS genes, it is not presented.*

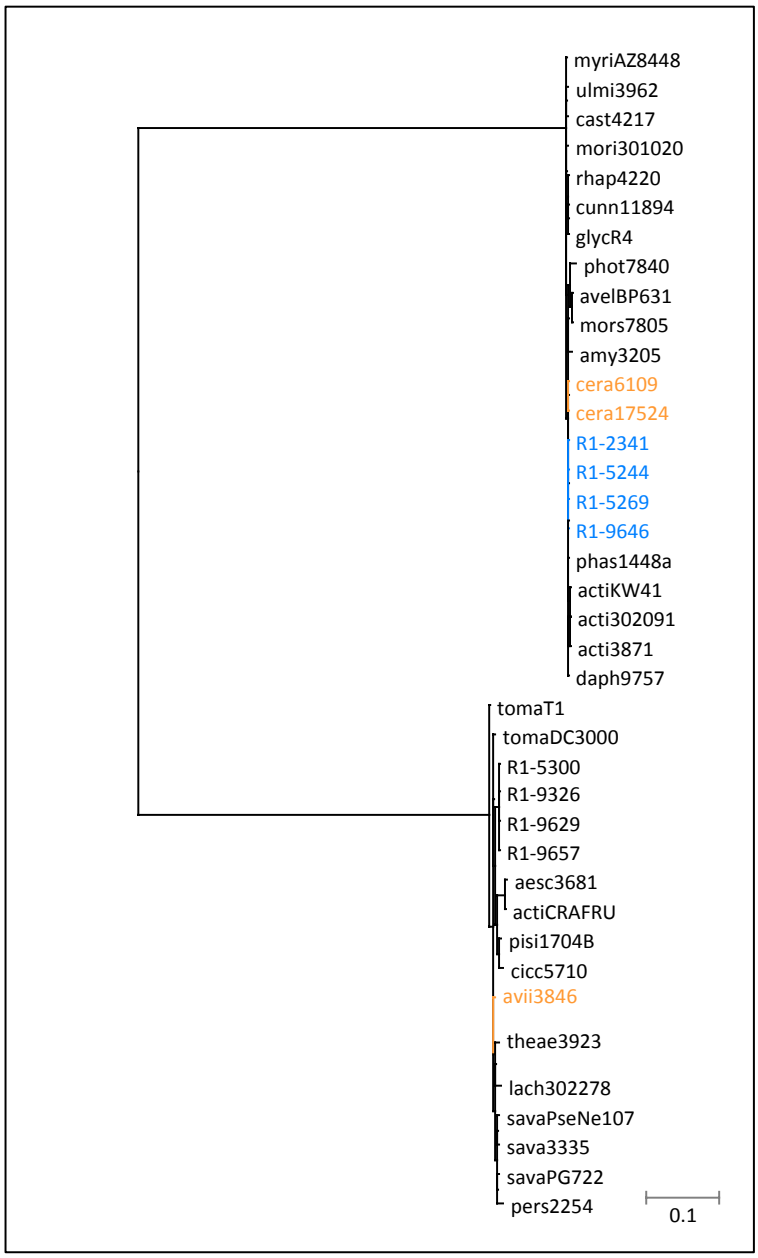


Figure S4.19: ML phylogenetic tree based on the* hopAV1 *gene generated using FastTree. *Cherry and plum isolates are highlighted and coloured dependent on clade:* Psm *R1 pathogens: blue and the other cherry pathogens (*P.s *pv.* avii *and* P.s *pv.* cerasicola*) in orange. Support values below 99% are shown for each node. The scale bar shows substitutions per site.*

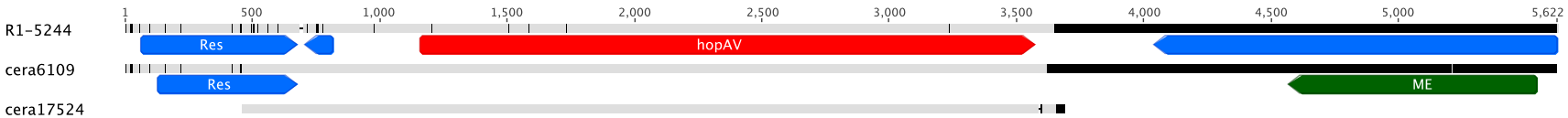


Figure S4.20: Alignment of the DNA region surrounding the* hopAV1 *gene in* Prunus *strains. *Grey indicates sequence identity whereas black indicates divergence. Homology between* Prunus*-infecting strains indicates a horizontal gene transfer may have occurred. The effector gene is coloured in red, whereas other CDS are in blue, mobile element genes are in green. Where the gene annotation was too long to be clearly shown for other CDS genes, it is not presented.*

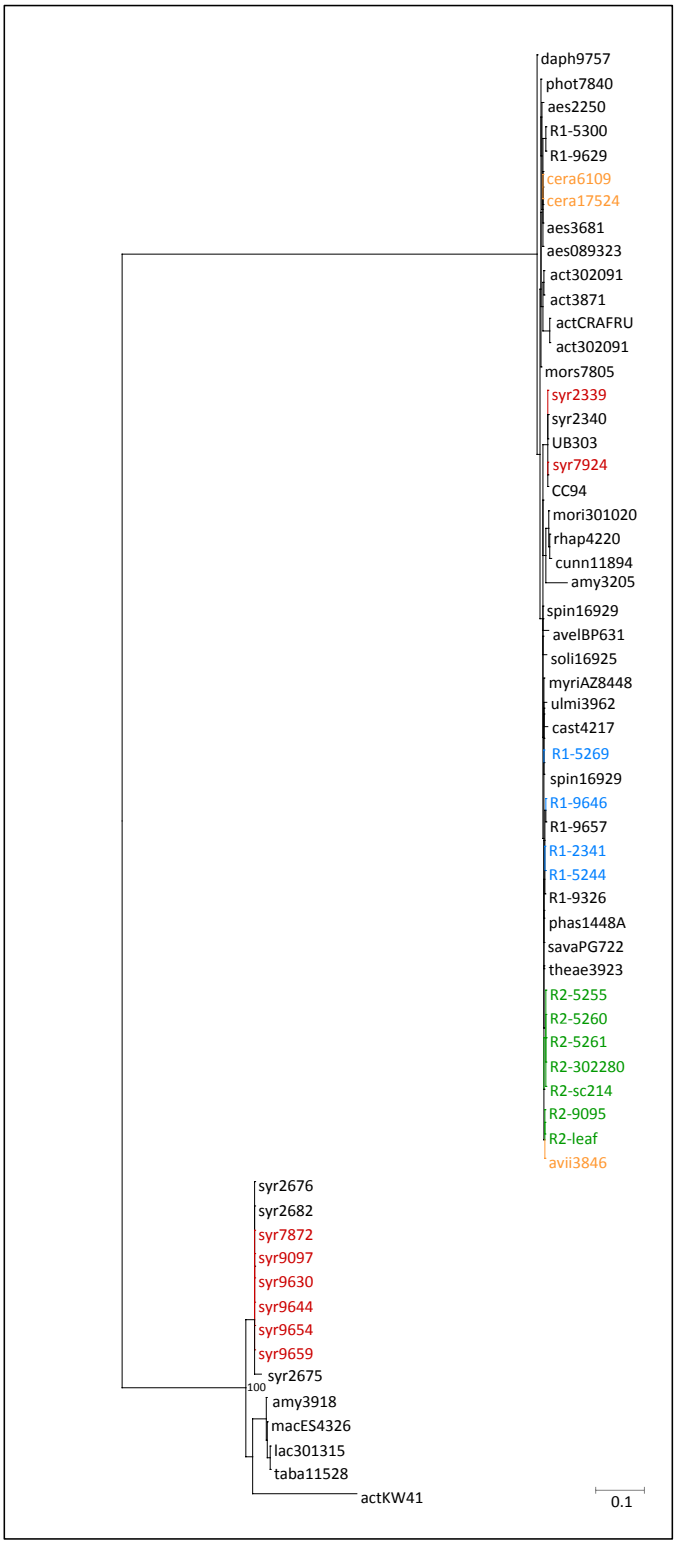


Figure S4.21: ML phylogenetic tree based on the* hopAY1 *gene generated using FastTree. *Cherry and plum isolates are highlighted and coloured dependent on clade:* Psm *R1 pathogens: blue,* Psm *R2: green,* Pss*: red and the other cherry pathogens (*P.s *pv.* avii *and* P.s *pv.* cerasicola*) in orange. Support values below 99% are shown for each node. The scale bar shows substitutions per site.*

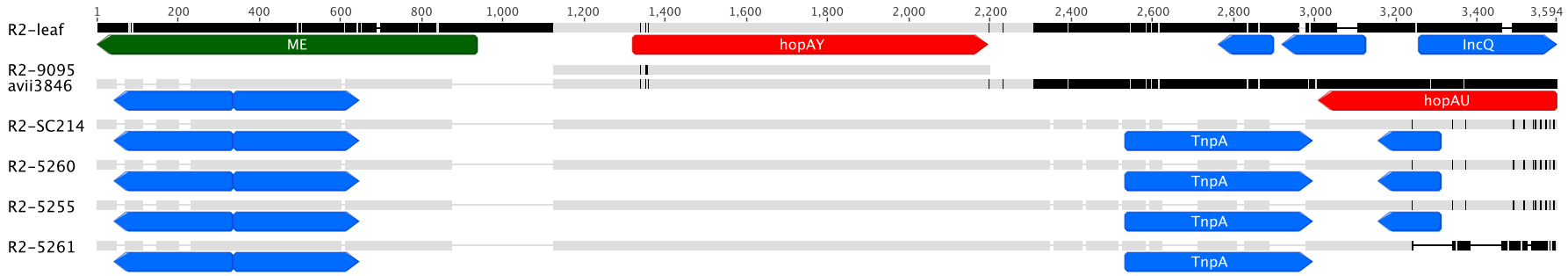


Figure S4.22: Alignment of the DNA region surrounding the* hopAY1 *gene in* Prunus *strains. *Grey indicates sequence identity whereas black indicates divergence. Homology between* Prunus*-infecting strains indicates a horizontal gene transfer may have occurred. The effector gene is coloured in red, whereas other CDS are in blue, mobile element genes are in green. Where the gene annotation was too long to be clearly shown for other CDS genes, it is not presented.*

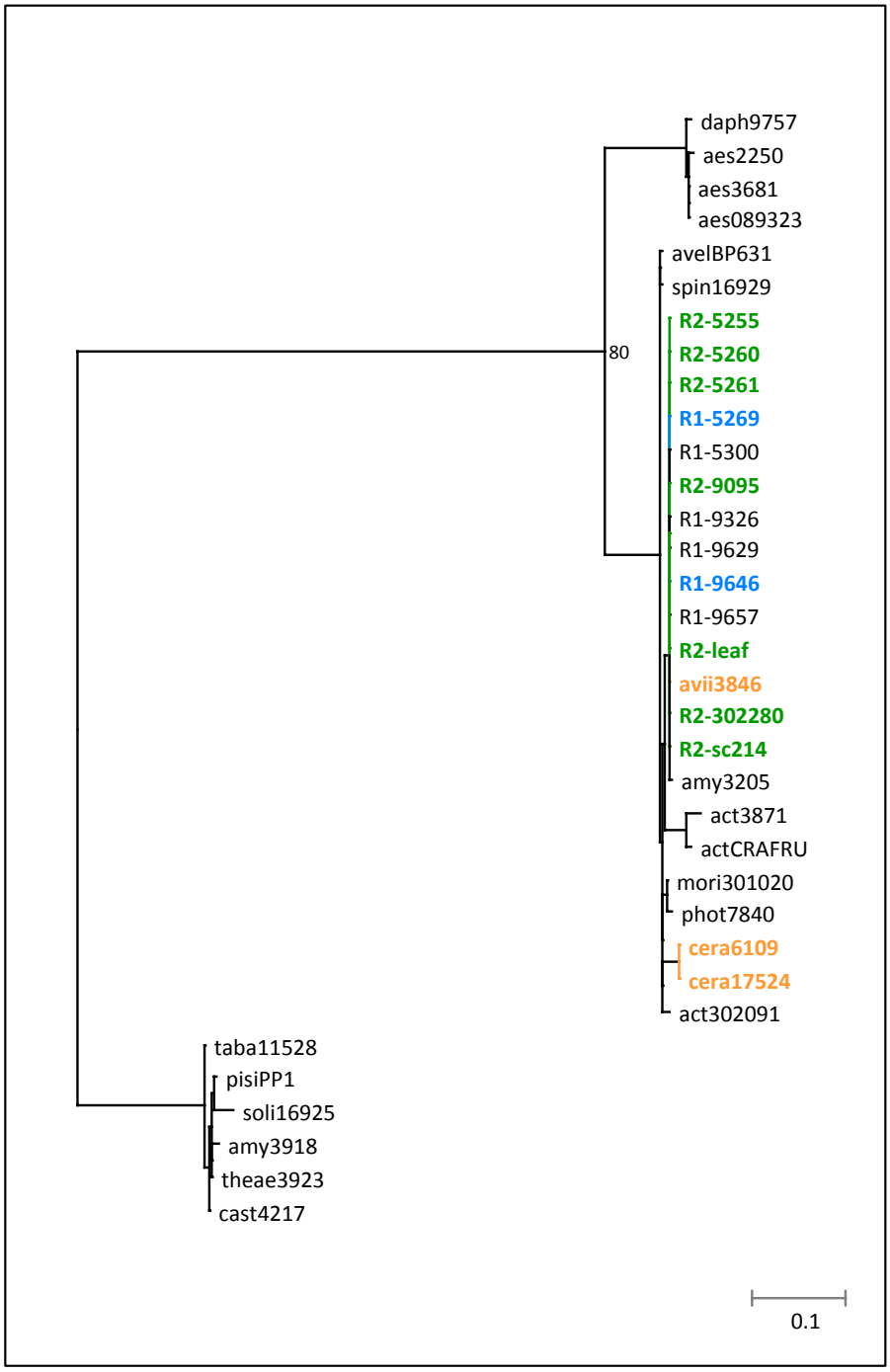


Figure S4.23: ML phylogenetic tree based on the* hopBB1 *gene generated using FastTree. *Cherry and plum isolates are highlighted and coloured dependent on clade:* Psm *R1 pathogens: blue,* Psm *R2: green and the other cherry pathogens (*P.s *pv.* avii *and* P.s *pv.* cerasicola*) in orange. Support values below 99% are shown for each node. The scale bar shows substitutions per site.*

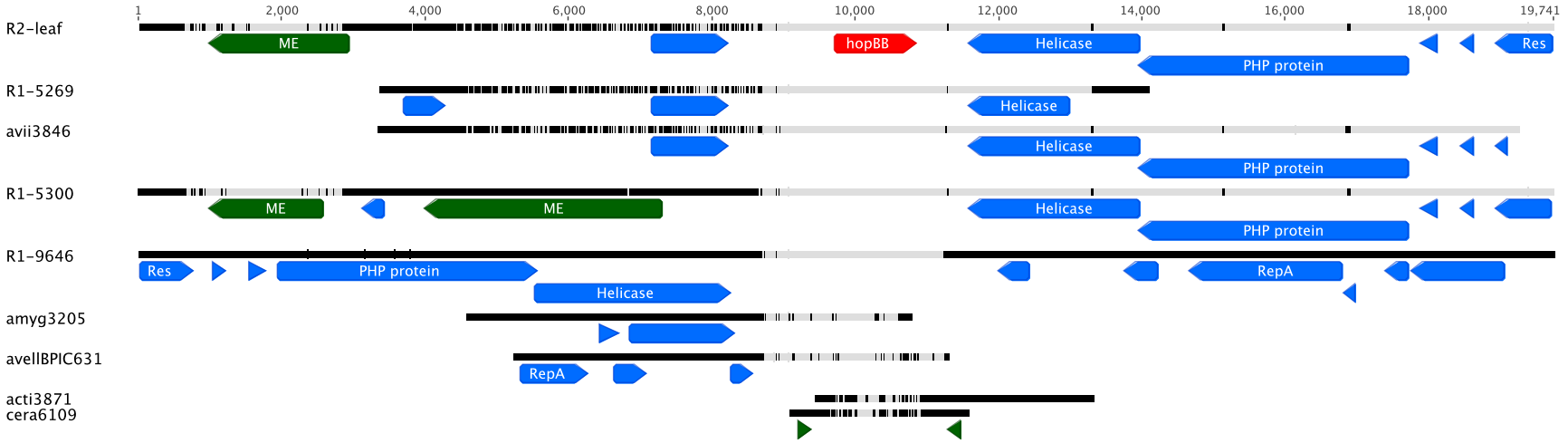


Figure S4.24: Alignment of the DNA region surrounding the* hopBB1 *gene in* Prunus *strains. *Grey indicates sequence identity whereas black indicates divergence. Homology between* Prunus*-infecting strains indicates a horizontal gene transfer may have occurred. The effector gene is coloured in red, whereas other CDS are in blue, mobile element genes are in green. Where the gene annotation was too long to be clearly shown for other CDS genes, it is not presented.*

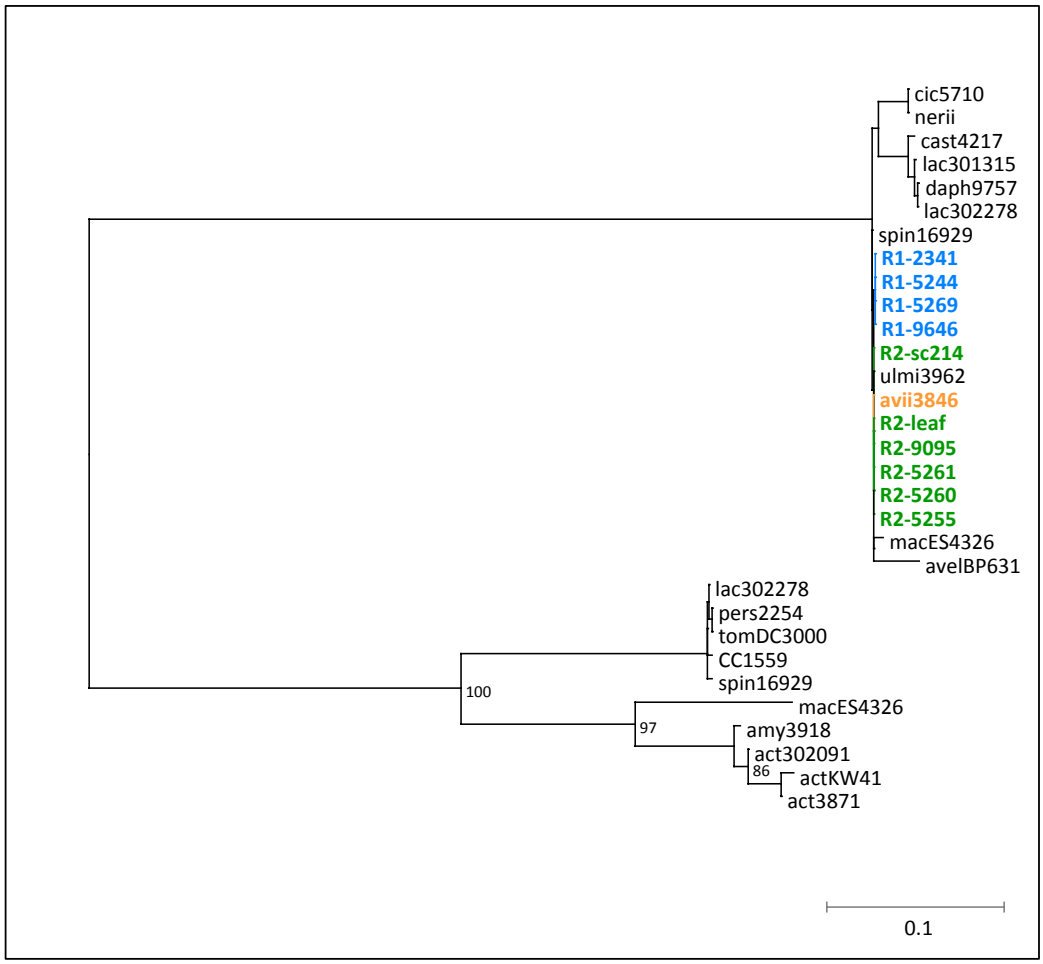


Figure S4.25: ML phylogenetic tree based on the* hopBD1 *gene generated using FastTree. *Cherry and plum isolates are highlighted and coloured dependent on clade:* Psm *R1 pathogens: blue,* Psm *R2: green and the other cherry pathogen (*P.s *pv.* avii*) in orange. Support values below 99% are shown for each node. The scale bar shows substitutions per site.*

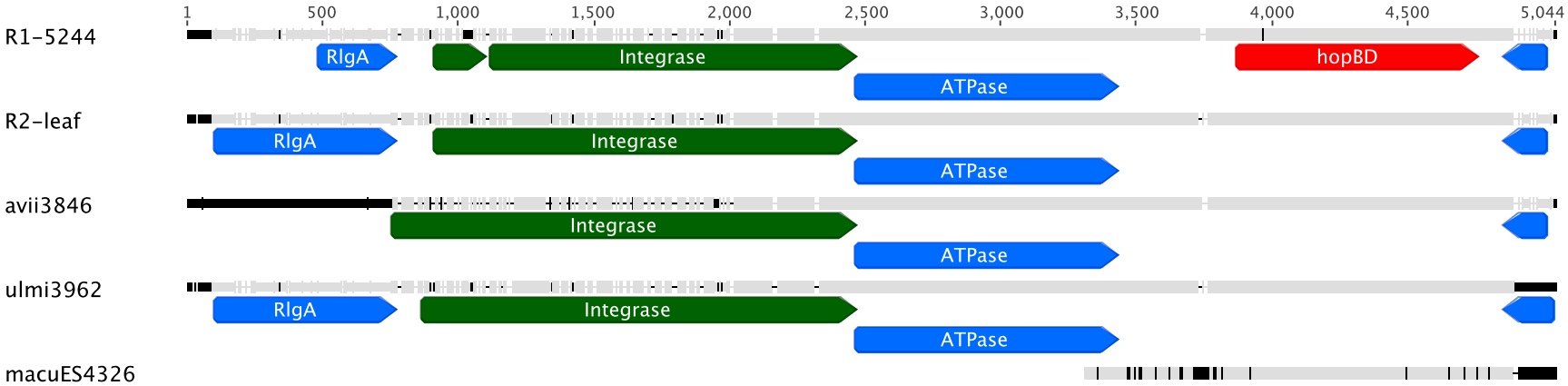


Figure S4.26: Alignment of the DNA region surrounding the* hopBD1 *gene in* Prunus *strains and some out-groups. *Grey indicates sequence identity whereas black indicates divergence. Homology between* Prunus*-infecting strains indicates a horizontal gene transfer may have occurred. The effector gene is coloured in red, whereas other CDS are in blue, mobile element genes are in green. Where the gene annotation was too long to be clearly shown for other CDS genes, it is not presented.*

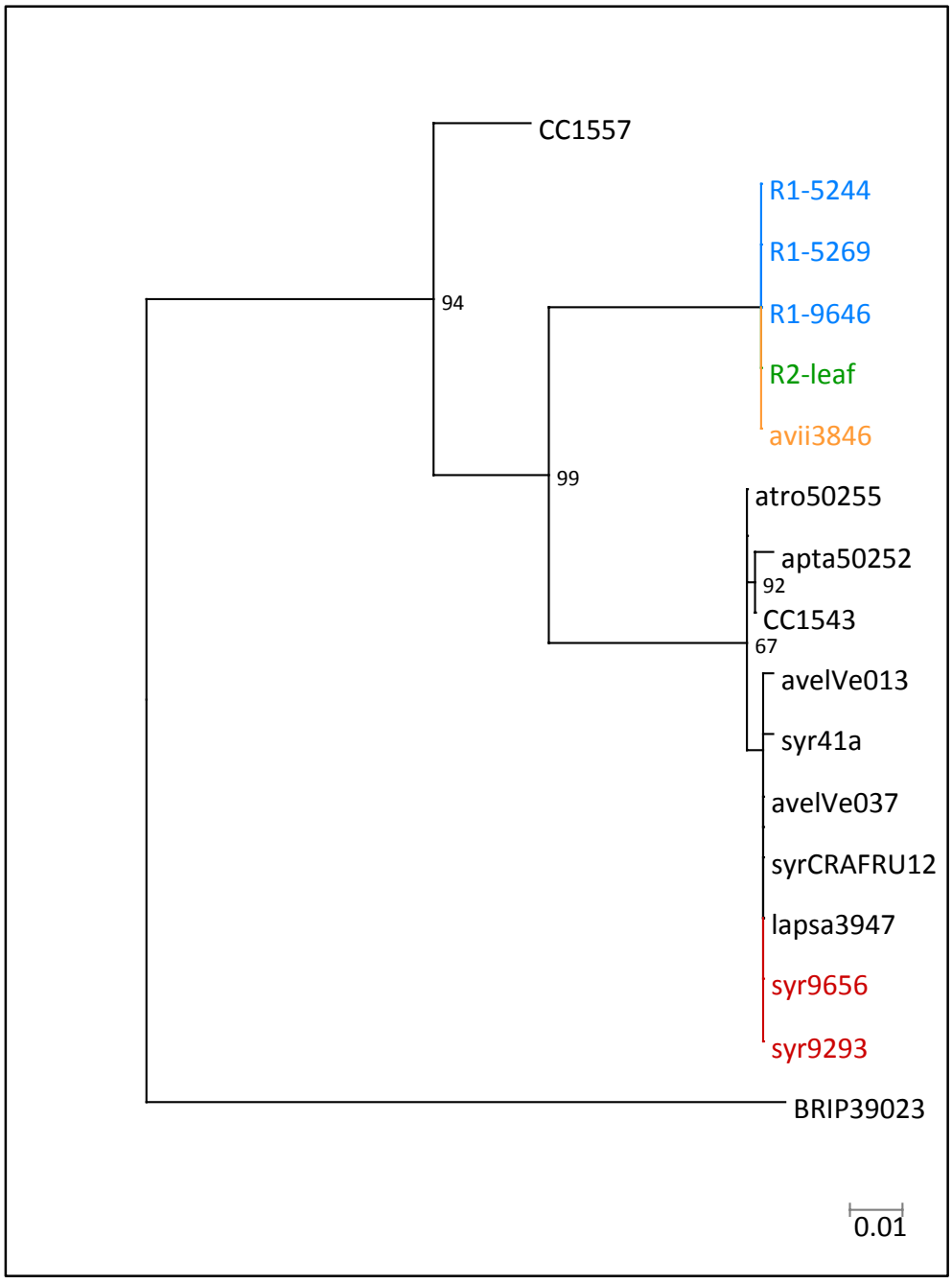


Figure S4.27: ML phylogenetic tree based on the* hopBF1 *gene generated using FastTree. *Cherry and plum isolates are highlighted and coloured dependent on clade:* Psm *R1 pathogens: blue,* Psm *R2: green,* Pss*: red and the other cherry pathogen (*P.s *pv.* avii*) in orange. Support values below 99% are shown for each node. The scale bar shows substitutions per site.*

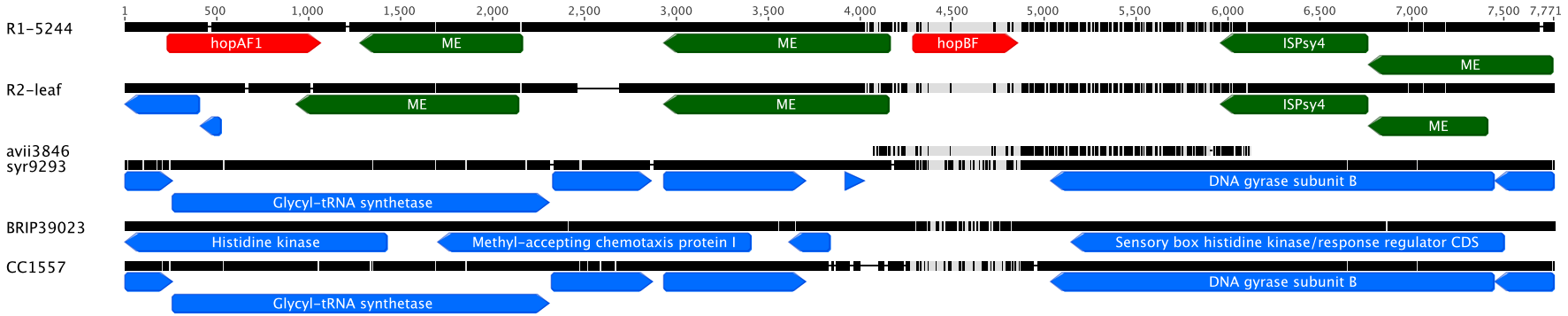


Figure S4.28: Alignment of the DNA region surrounding the* hopBF1 *gene in* Prunus *strains and some out-groups. *Grey indicates sequence identity whereas black indicates divergence. Homology between* Prunus*-infecting strains indicates a horizontal gene transfer may have occurred. The effector gene is coloured in red, whereas other CDS are in blue, mobile element genes are in green. Where the gene annotation was too long to be clearly shown for other CDS genes, it is not presented.*

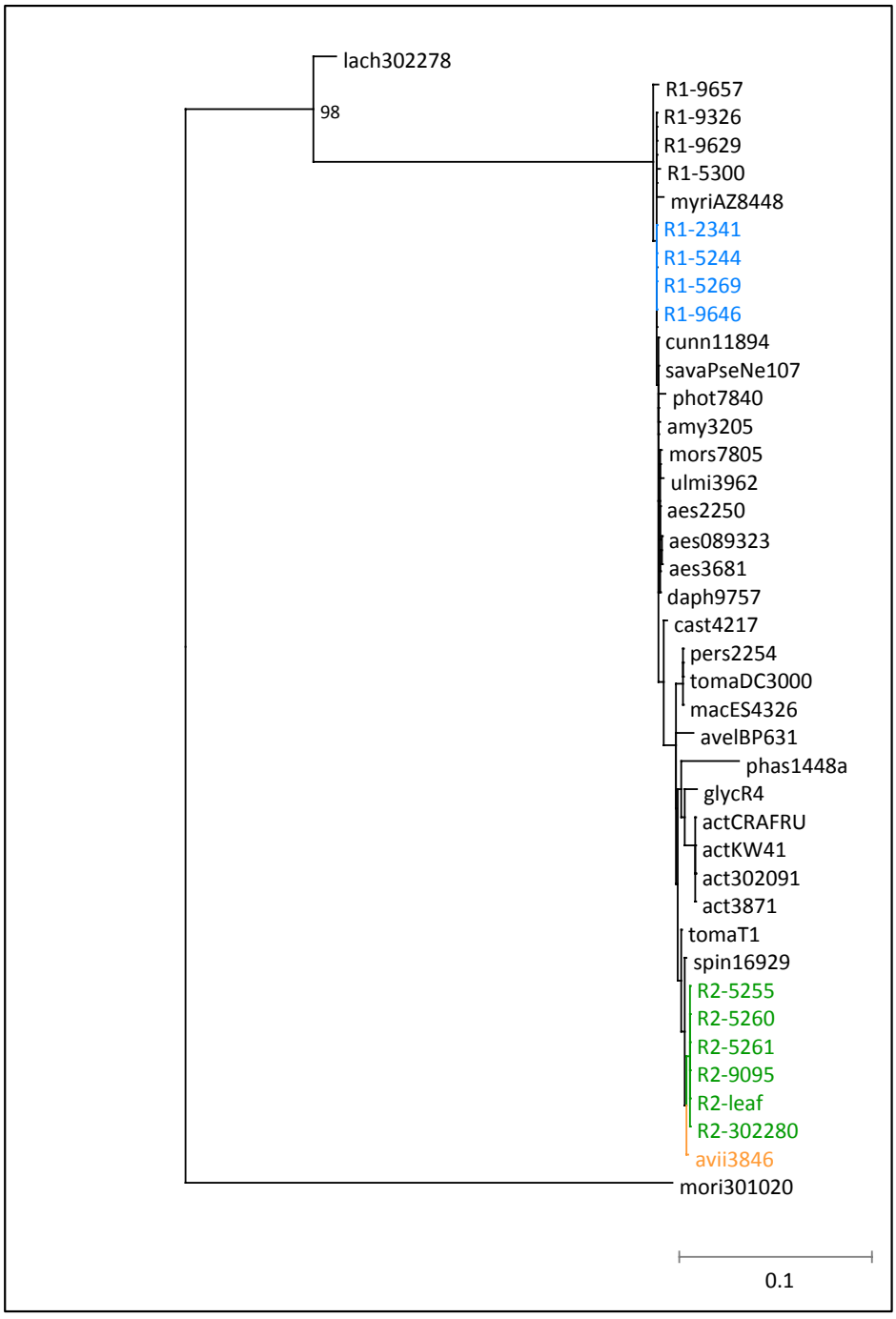


Figure S4.29: ML phylogenetic tree based on the* hopD1 *gene generated using FastTree. *Cherry and plum isolates are highlighted and coloured dependent on clade:* Psm *R1 pathogens: blue,* Psm *R2: green and the other cherry pathogen (*P.s *pv.* avii*) in orange. Support values below 99% are shown for each node. The scale bar shows substitutions per site.*

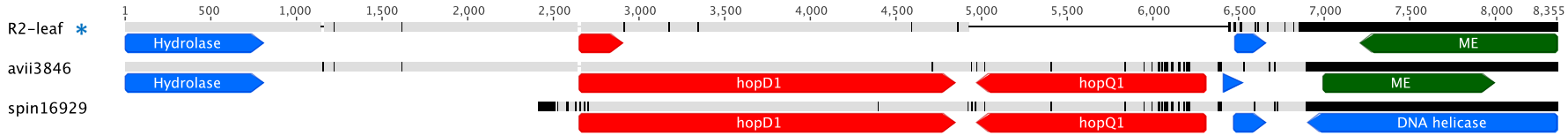


Figure S4.30: Alignment of the DNA region surrounding the* hopD1 *gene in* Prunus *strains and an out-group. *Grey indicates sequence identity whereas black indicates divergence. Homology between* Prunus*-infecting strains indicates a horizontal gene transfer may have occurred. The effector gene is coloured in red, whereas other CDS are in blue, mobile element genes are in green. * The gene is truncated in R2-leaf due to a premature stop codon. Where the gene annotation was too long to be clearly shown for other CDS genes, it is not presented.*

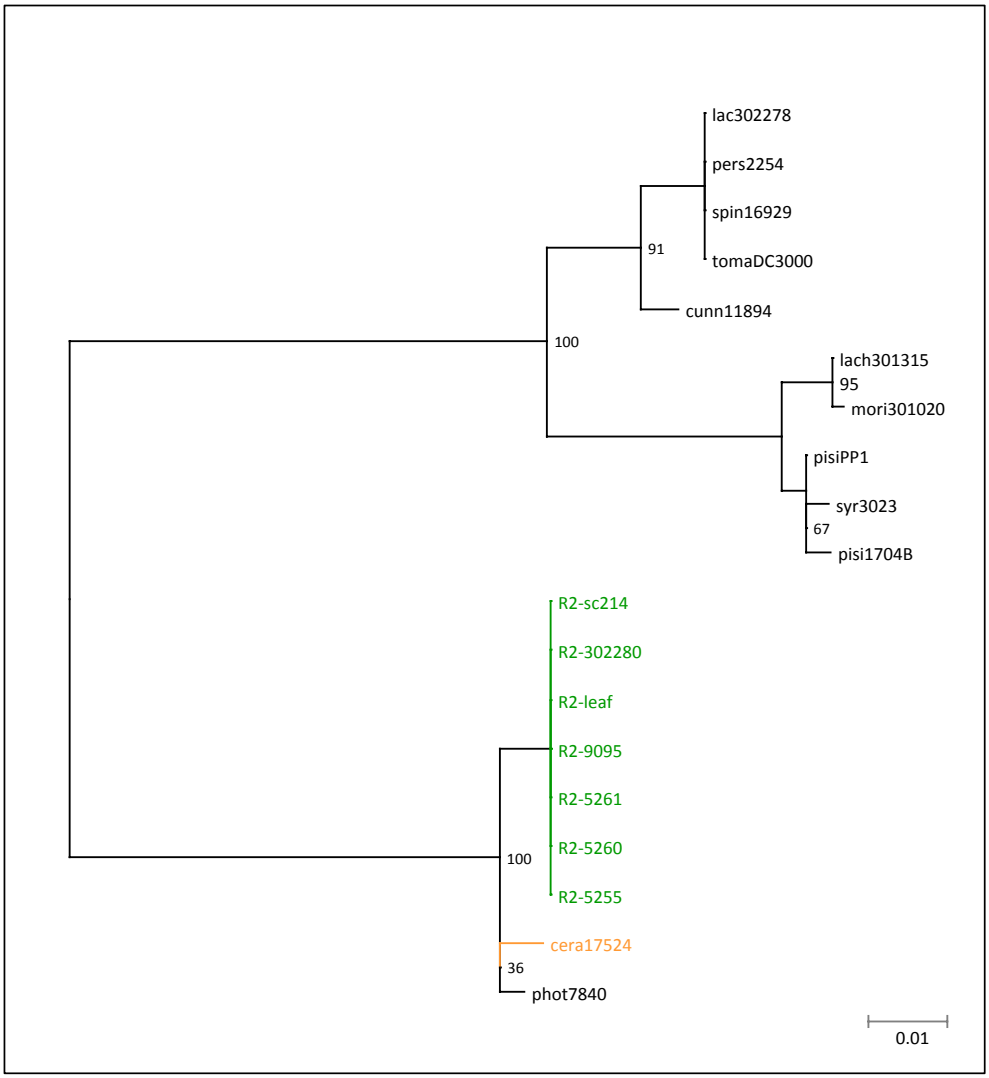


Figure S4.31: ML phylogenetic tree based on the* hopE1 *gene generated using FastTree. *Cherry and plum isolates are highlighted and coloured dependent on clade:* Psm *R2: green and the other cherry pathogen (*P.s *pv.* cerasicola*) in orange. Support values below 99% are shown for each node. The scale bar shows substitutions per site.*

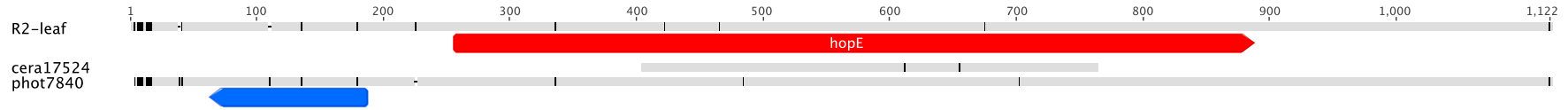


Figure S4.32: Alignment of the DNA region surrounding the* hopE1 *gene in* Prunus *strains and an out-group. *Grey indicates sequence identity whereas black indicates divergence. Homology between* Prunus*-infecting strains indicates a horizontal gene transfer may have occurred. The effector gene is coloured in red, whereas other CDS are in blue, mobile element genes are in green. Where the gene annotation was too long to be clearly shown for other CDS genes, it is not presented.*

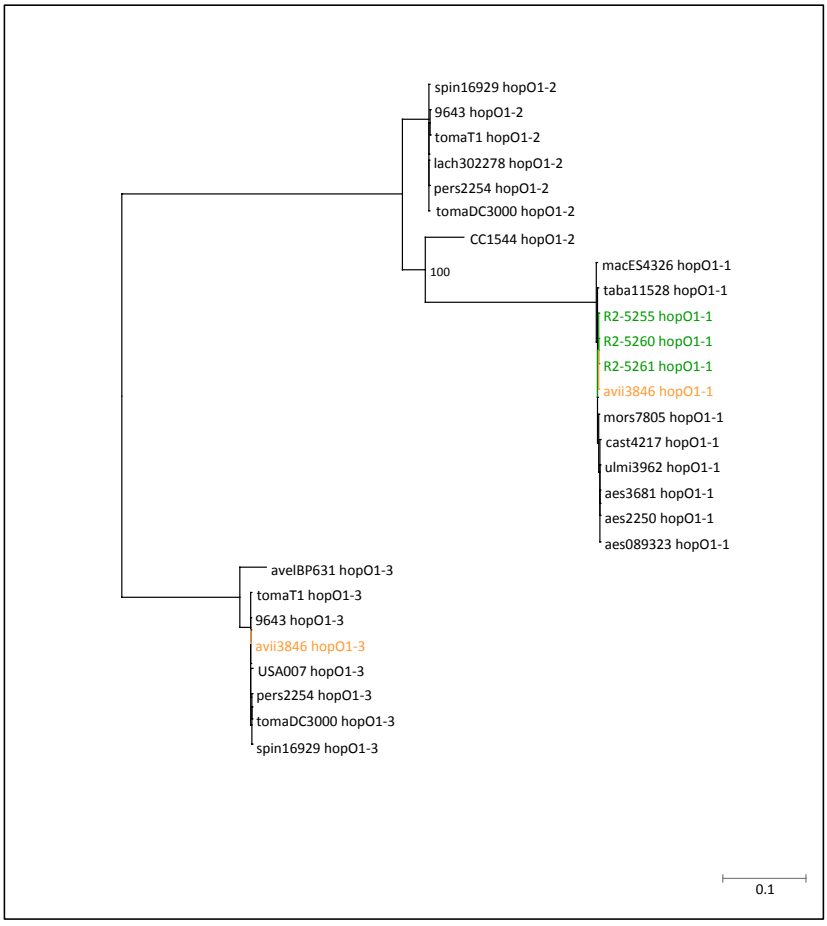


Figure S4.33: ML phylogenetic tree based on the* hopO *gene family generated using FastTree.

Cherry and plum isolates are highlighted and coloured dependent on clade: Psm *R2: green and the other cherry pathogen (*P.s *pv.* avii*) in orange. Support values below 99% are shown for each node. The scale bar shows substitutions per site.*

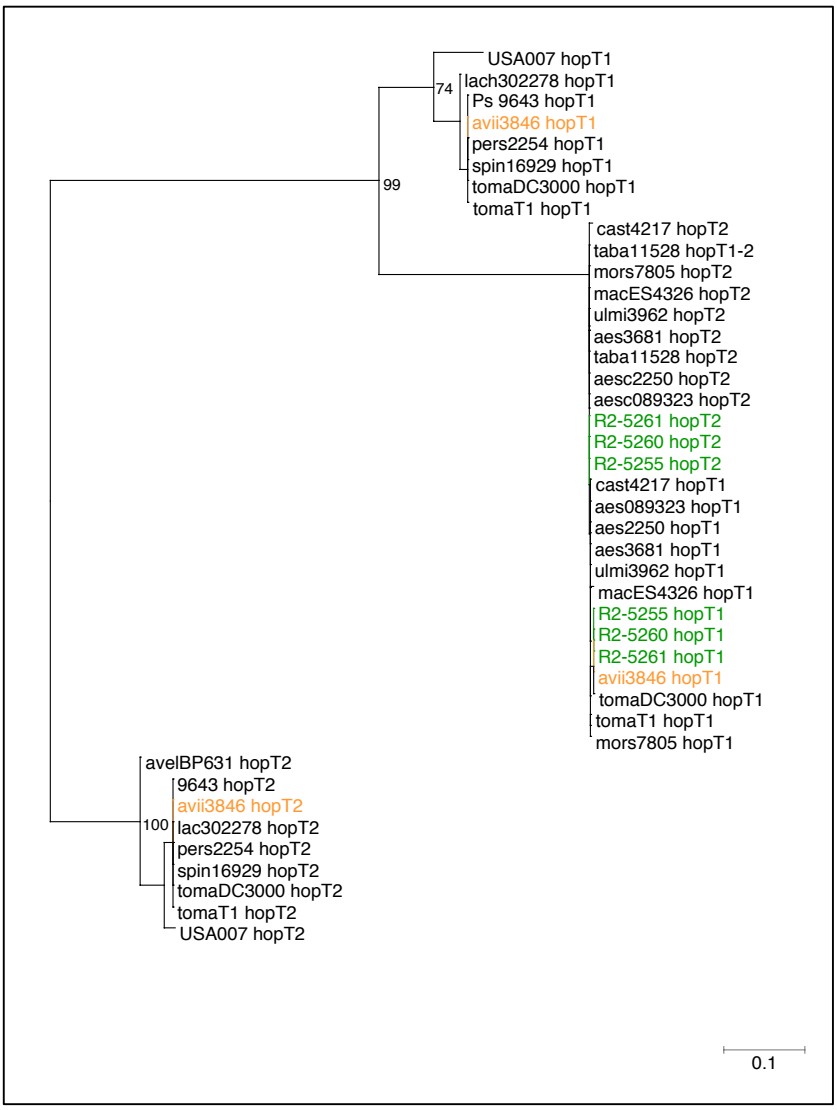


Figure S4.34: ML phylogenetic tree based on the* hopT *gene family generated using FastTree. *Cherry and plum isolates are highlighted and coloured dependent on clade:* Psm *R2: green and the other cherry pathogen (*P.s *pv.* avii*) in orange. Support values below 99% are shown for each node. The scale bar shows substitutions per site.*

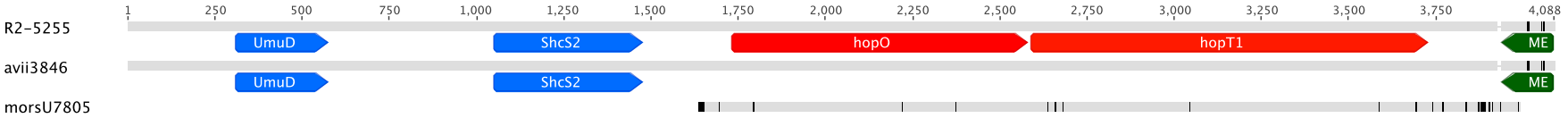


Figure S4.35: Alignment of the DNA region surrounding the* hopO1 and hopT1 *gene in* Prunus *strains isolated from cherry and apricot. *Grey indicates sequence identity whereas black indicates divergence. Homology between cherry-infecting strains indicates a horizontal gene transfer may have occurred. The effector gene is coloured in red, whereas other CDS are in blue, mobile element genes are in green. Where the gene annotation was too long to be clearly shown for other CDS genes, it is not presented.*

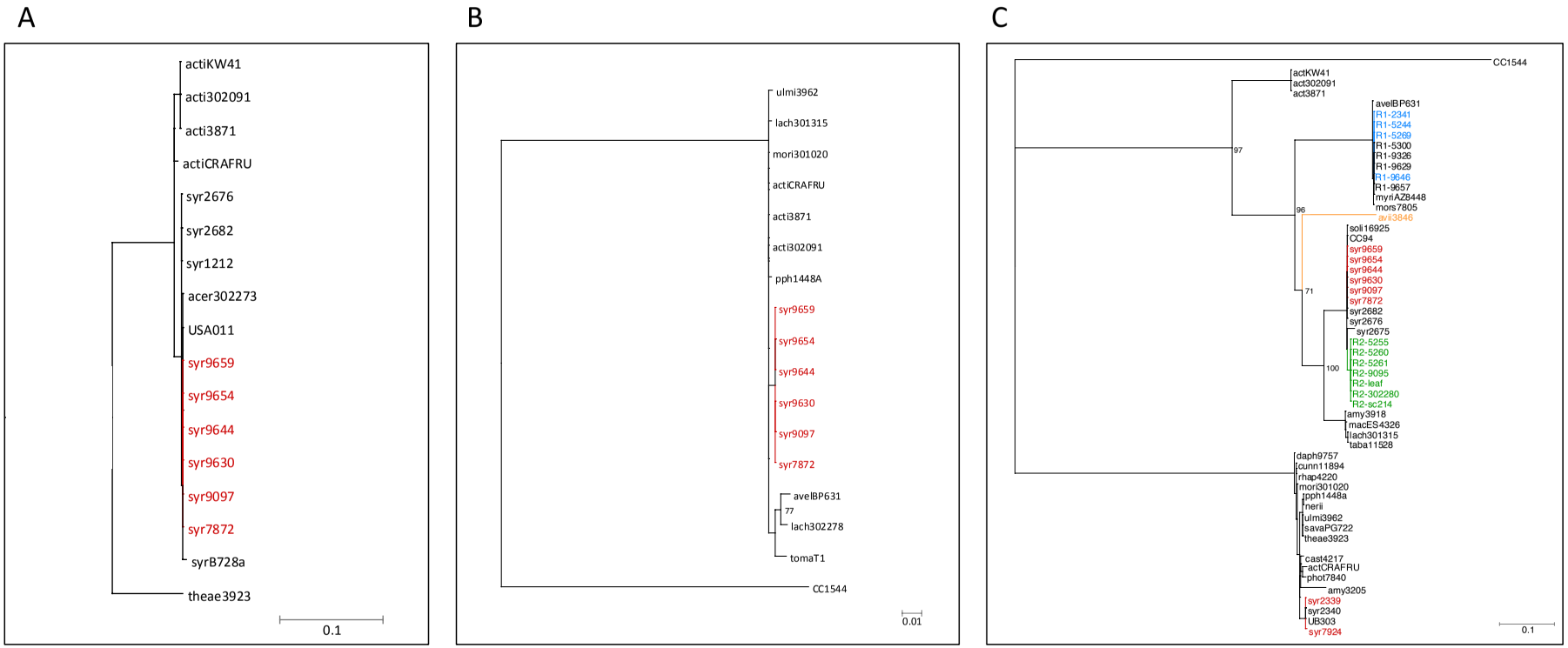


Figure S4.36: ML phylogenetic trees of convergently acquired Pss effector genes. A: avrRpm1, B: hopAW1, C: hopAR1. *Trees were generated using FastTree. Cherry and plum isolates are highlighted and coloured dependent on clade:* Psm *R1: blue,* Psm *R2: green,* Pss*: red and the other cherry pathogen (*P.s *pv.* avii*) in orange. Support values below 99% are shown for each node. The scale bar shows substitutions per site.*

Chapter 5: Cloning of candidate virulence and avirulence effector genes involved in host specificity of *P. syringae* on cherry

5.1 Abstract

Pseudomonas syringae is a highly host-specific bacterial plant pathogen. Differences in host range are thought to be strongly linked to the repertoire of type III effectors possessed by individual strains. The topic of host specificity was explored using strains that cause bacterial canker of cherry. Several distinct *P. syringae* clades can infect cherry. The three major bacterial canker pathogens are *P.s.* pv. *morsprunorum* race 1, *P.s.* pv. *morsprunorum* race 2 and *P.s.* pv. *syringae*. Nine effector genes were initially cloned from closely related non-pathogenic strains isolated from other plant species and expressed in pathogenic strains. Two effectors, HopAB1 and HopC1, were found to significantly reduce the population growth of pathogenic strains and putatively induce a hypersensitive response (HR) in cherry leaves. This suggests that these two effectors are being detected by the cherry immune system. Cloning of additional members of HopAB (HopAB2 and HopAB3) confirmed these effectors also reduced pathogen population growth, however this reduction was smaller in magnitude compared to HopAB1 and the HR induced by these effectors appeared later. Deletion of these HopAB alleles in non-pathogenic strains did not cause a statistically significant increase their growth in cherry leaves, however an increase in multiplication of *P.s.* pv. *avellanae* Δ*hopAB2* mutants was recorded. The lack of clear population increase after the deletion of avirulence factors is likely due to the presence of other effectors/PAMPs that still trigger plant immune responses or because the non-pathogens lack the virulence factors required to reach higher population levels. The convergent loss/pseudogenisation of *hopAB* alleles in cherry pathogenic strains is therefore likely to be the result of selective pressures exerted by the host immune response towards this effector. Finally, to study the role of virulence factors important for disease on cherry, three effectors significantly associated with disease on cherry and found in multiple cherry pathogenic clades, were deleted from *Psm* R2. Single,

double and triple mutants were not significantly reduced in population growth, suggesting that they are not crucial for virulence or that functional redundancy with other effectors is masking the phenotypic effect of their deletion.

5.2 Introduction

This chapter used cloning to determine the role of particular effectors in host specificity of *Pseudomonas syringae*. Understanding the evolutionary processes that restrict host range in *P. syringae* is important for disease prevention, management and the successful prediction of the epidemics (Monteil *et al.* 2016). Host specificity is dictated by genetic variation in both pathogen virulence and host immunity (Reignault & Sancholle 2005; Lindeberg *et al.* 2009). The molecular interplay between effectors and the plant immune system (discussed in detail in Chapter 1) determines if a bacteria-plant interaction leads to disease (compatible) or resistance (incompatible). However, many interactions fall somewhere between these two extremes due to the quantitative nature of disease and plant resistance (Iakovidis *et al.* 2016).

The ability of host resistance mechanisms to shape pathogen host range has been widely explored in *P. syringae*. Before the advent of genomics, early molecular studies utilised genomic libraries to identify factors involved in host specificity. These approaches involved the creation of a library containing fragments of a non-pathogenic strain's genome inserted into cosmid vectors. The library was then transferred into a pathogenic strain, and screened for the ability to change the pathogen's virulence *in planta*. A reduction in virulence or induction of the hypersensitive response, due to gene products expressed by particular clones indicated the presence of an avirulence factor. This allowed the identification of several avirulence effector proteins that trigger plant immune responses and the realisation that the same effectors can invoke both host and non-host resistance responses towards *P. syringae* (Staskawicz *et al.* 1987; Ronald *et al.* 1992; Jackson *et al.* 1999; Fillingham *et al.* 1992; Kobayashi *et al.* 1989; Whalen *et al.* 1991; Arnold *et al.* 2001a). For example, Staskawicz *et al.* (1987), identified that host resistance of soybean cultivars to races 1 and 0 of the soybean pathogen *P.s* pv. *glycinea*, involved

the elicitation of a HR by the effector AvrB. Subsequent studies determined that this is due to indirect detection by the Nucleotide Binding Leucine-Rich Repeat (NB-LRR) R protein Rpg1b, which detects effector modification of RIN4-like proteins (Ashfield *et al.* 2003; Selote & Kachroo 2010). Studies of AvrB in *Arabidopsis* have found that *Arabidopsis* has convergently evolved to detect the presence of AvrB via a similar mechanism, whereby the RPM1 R protein detects effector modifications of RIN4 (Mackey *et al.* 2002; Ashfield *et al.* 2014; Whitham *et al.* 2016). The use of genomic libraries also led to the identification of AvrPto, an avirulence effector which triggers host resistance in Pto-expressing tomato lines (Ronald *et al.* 1992). Subsequent studies revealed that diverse homologues of this effector and a related protein AvrPtoB (now called HopAB2) from various *P. syringae* pathovars also elicit Pto triggered immunity in tomato, indicating that non-host resistance mechanisms may also involve detection of these effectors (Lin & Martin 2007). They found that deletion of these effectors allowed non-pathogens to grow to higher levels *in planta*, indicating that the loss of particular effectors may drive changes in host range.

These studies laid the ground-work for current hypotheses linking effector repertoires to host range in *P. syringae* and other plant pathogens (Lindeberg *et al.* 2012; O'Brien *et al.* 2011). Genome sequencing has allowed the comparison of effector repertoires, leading to the identification of candidate avirulence factors for cloning. A study comparing closely related bean pathogens showed that subtle differences in effector repertoires underlie host specificity (Baltrus *et al.* 2012). Two effectors (HopC1 and HopM1) were found to be unique to the soybean pathogen *P.s* pv. *glycinea*. When these effectors were expressed in the French bean pathogen *P.s* pv. *phaseolicola*, the growth of this pathogen was reduced on its homogeneous host. This indicated they may be avirulence factors that operate during non-host resistance to *P.s* pv. *glycinea*, preventing it from causing disease on French bean. A study of the model strain, *P.s* pv. *tomato* DC3000, showed that the expression of the HopQ1-1 effector activates ETI in *Nicotiana benthamiana*. Deletion of this effector allowed this strain to grow and cause disease, again showing that host range can be expanded by small changes in effector complement (Wei et al. 2007). HopQ1-1 is absent from many strains of *P.s* pv. *tabaci*. The absence of this effector may therefore reflect this pathovar's adaptation to cause

disease on *Nicotiana benthamiana* (Ferrante *et al.* 2009). More subtle, allelic variation of effectors can also dictate host specificity. Stevens *et al.* (1998), found that single amino acid substitutions in the HopX1 (AvrPphE) protein of different races of *P.s* pv. *phaseolicola* led to evasion of *R2*-mediated host resistance in certain cultivars of bean.

Evolution experiments have confirmed that host selection pressures drive the loss or inactivation of avirulence effector genes in bacterial populations. Strains of *P.s* pv. *phaseolicola* race 4 expressing HopAR1 (AvrPphB) cause the HR on resistant plants that express the R3 R protein (Jackson *et al.* 2000). Host resistance drives the excision of the genomic island, containing the *hopAR1* gene, from the chromosome when these strains are passaged on the resistant host, facilitating the evolution of virulence via natural selection (Pitman *et al.* 2005). Further experimental evolution studies revealed that the genomic island is initially not lost, instead it resides within the cell as a circular molecule separate from the chromosome and gene expression is suppressed (Godfrey *et al.* 2011). Under constant host resistance selection, the island is lost from most members of the population. However, it is still maintained at low frequencies as it confers a fitness benefit on plants lacking the *R3* resistance gene (Neale *et al.* 2016).

Therefore, the perception of effectors has played an integral role in the evolution of host specificity in *P. syringae*. However, the early perception of PAMPs such as flagellin is also an important stage during the *P. syringae*-plant interaction. Takeuchi *et al.* (2003) found that differences in post-translational modification of flagellin influenced host specificity. Flagellins from different pathovars varied in the ability to induce plant cell death due to differences in post-translation glycosylation. Clarke *et al.* (2013) found that allelic variation in flagellin proteins significantly influenced the ability of *Arabidopsis* to detect *P. syringae*. Therefore, although PAMPs are more conserved than effectors, diversification of PAMP sequences or post-translational modification may allow *P. syringae* to evade detection on particular hosts.

During host specialisation towards particular plant species, the acquisition of specific factors that increase bacterial fitness and promote disease is also required. Various studies have implicated certain effectors with specific plant diseases. The effectors

AvrPto1 and HopAB2 have been shown to be important for virulence of *P.s* pv. *tomato* DC3000 on tomato, as their deletion resulted in reduced population growth (Lin & Martin 2005). HopAB1 (a homologue of HopAB2) has also been shown to be required for full virulence of *P.s* pv. *phaseolicola* on bean due to its ability to suppress ETI (Jackson *et al.* 1999). The acquisition of HopAB1 in strains of *P.s* pv. *syringae* that are distantly related to *P.s* pv. *phaseolicola* but have convergently evolved to infect bean, indicates it may be important for host specificity (Vinatzer *et al.* 2006). Bioinformatics and functional analysis of the olive pathogen *P.s* pv. *savastanoi* revealed two previously unknown effectors HopBL1 and HopBL2, which are present in *P.s* pv. *savastanoi* strains and other strains infecting woody host plants (Matas *et al.* 2014). Nowell *et al.* (2016) determined that the effector genes, *hopAY1* and *hopAO1,* are also associated with specialisation for woody host species. The presence of these effectors in multiple wood-infecting pathovars therefore indicates that they may be important for virulence on woody hosts.

However, the identification of crucial virulence-associated effectors has been hindered by the fact that most deletion mutants show little or no attenuation in virulence due to functional redundancy of effector proteins (Wei et al. 2007). To study the individual contributions of effectors to virulence of *P.s* pv. *tomato* DC3000 on *N. benthamiana*, Cunnac *et al.* (2011) used a poly-mutant approach. They deleted the majority of effectors from *P.s* pv. *tomato* DC3000 and then reassembled the effector repertoire to find the minimal number required for virulence. They found that the minimal repertoire showed functional hierarchy, with AvrPto/AvrPtoB acting early to prevent PAMP perception. Six other effectors (HopM1, HopE1, AvrE1, HopAM1, HopG1 and HopAA1 or HopN1) were required to reach wild-type growth and disease symptoms. This study revealed that *P. syringae* only requires a small number of its effectors to cause disease, but carries a large functionally redundant repertoire to allow rapid adaptation to new niches. This functional redundancy helped explain why strains infecting the same plant species often possess divergent effector repertoires (Almeida *et al.* 2009; O'Brien *et al.* 2012). The generation of polymutants, as in Cunnac *et al.* (2011), therefore provided a method to identify effectors vital for pathogenicity on particular plant species.

Type III secreted effectors are not the only factors implicated in *P. syringae* virulence and host specialisation. Many strains of *P. syringae* produce non-enzymatic, low MW phytotoxins. The toxins are generally not associated with host specificity (Bender *et al.* 1999) as the same toxin can be found in strains infecting different plant species, e.g. the phytotoxin coronatine genes are present in the genomes of various pathovars, including *P.s* pv. *tomato*, *P.s* pv. *aesculi* and *P.s* pv. *morsprunorum* (Cuppels & Ainsworth 1995). Phytotoxins may provide similar virulence functions to effectors, such as induction of necrosis and chlorosis, allowing different lineages to specialise on the same host using variable effector repertoires (Dudnik & Dudler 2014). Apart from phytotoxins, different metabolic subsystems may be important for specialisation. Green *et al.* (2010) found that pathogenic *P.s* pv. *aesculi* strains have acquired a cluster of genes whose products are involved in sugar utilisation that may be important for occupying vascular plant tissue. Genomics of wood-infecting pathovars showed that many strains produce proteins involved in the metabolism of aromatic compounds (Rodríguez-Palenzuela *et al.* 2010; Buonaurio *et al.* 2015; Caballo-Ponce *et al.* 2016). These pathways may have allowed specialisation of wood-infecting strains due to the ability to utilise compounds found in wood or detoxify compounds with antimicrobial properties (Bartoli *et al.* 2015).

This study combines the knowledge gained through bioinformatics and pathology to use cloning to explore the host specificity of *P. syringae* strains that cause bacterial canker on cherry. The cherry-infecting clades are reported to show subtle differences in virulence, distribution and lifestyle (Scortichini 2010) and therefore provide an intriguing opportunity to study host specificity.

Focusing on effectors as the main determinant of specificity, two parallel approaches were used. One involved the identification of candidate avirulence effectors that trigger resistance to non-pathogens, whilst the other involved searching for important virulence-associated effectors that may be vital for disease on cherry. Finally, an unbiased approach was taken, by using genomic libraries which could help identify

candidate host specificity factors not identified through genomic comparisons based on bioinformatics analyses.

5.3 Methods

5.3.1 Bacterial strains

Bacteria were grown as in Chapter 2 on Kings B agar or in Luria Broth (LB). As plasmids were introduced, antibiotics were added to the media. Antibiotics were used at the following concentrations (μg/ml): Kanamycin 50, Gentamycin 10, Tetracycline 20, Spectinomycin 100, Nitrofurantoin 100. For blue-white screening of *E. coli*, X-gal was added to the media at a concentration of 80μg/ml. All strains were stored in 40% glycerol at -80°C. Tables 5.1 and 5.2 list the *P. syringae* strains and plasmids used in this study.

Strain	Description	Reference
R1-5244	*P.s* pv. *morsprunorum* R1 5244 isolated from a diseased cherry cv. Napoleon	This study
R1-5300	*P.s* pv. *morsprunorum* R1 5300 isolated from a diseased plum cv. Victoria	This study
R1-9657	*P.s* pv. *morsprunorum* R1 9657 isolated during a leaf wash from cherry cv. Kiku-Shidare	This study
R2-leaf	*P.s* pv. *morsprunorum* R2 leaf isolated from a leaf disease lesion from cherry cv. Napoleon	This study
syr9644	*P.s* pv. *syringae* 9644 isolated during a leaf wash from cherry cv. Napoleon	This study
PsavBPIC631	*P.s* pv. *avellanae* BPIC631 type strain isolated from hazelnut	O'Brien *et al.* 2012
Pph1448A	*P.s* pv. *phaseolicola* race 6	Joardar *et al.* 2005
RMA1	*P.s* strain RMA1 isolated from a diseased *Aquilegia vulgaris* cv. Winky	This study
R1-5244 pBBR-1	*P.s* pv. *morsprunorum* R1 5244 with pBBR-1	This study
R1-5244 pBBR-2	*P.s* pv. *morsprunorum* R1 5244 with pBBR-2	This study
R1-5244 pBBR-3	*P.s* pv. *morsprunorum* R1 5244 with pBBR-3	This study
R1-5244 pBBR-4	*P.s* pv. *morsprunorum* R1 5244 with pBBR-4	This study
R1-5244 pBBR-5	*P.s* pv. *morsprunorum* R1 5244 with pBBR-5	This study
R1-5244 pBBR-6	*P.s* pv. *morsprunorum* R1 5244 with pBBR-6	This study
R1-5244 pBBR-7	*P.s* pv. *morsprunorum* R1 5244 with pBBR-7	This study
R1-5244 pBBR-8	*P.s* pv. *morsprunorum* R1 5244 with pBBR-8	This study
R1-5244 pBBR-9	*P.s* pv. *morsprunorum* R1 5244 with pBBR-9	This study
R1-5244 pBBR-10	*P.s* pv. *morsprunorum* R1 5244 with pBBR-10	This study
R1-5244 pBBR-11	*P.s* pv. *morsprunorum* R1 5244 with pBBR-11	This study
R1-5244 pBBR-12	*P.s* pv. *morsprunorum* R1 5244 with pBBR-12	This study
R2-leaf pBBR-1	*P.s* pv. *morsprunorum* R2 leaf with pBBR-1	This study
R2-leaf pBBR-2	*P.s* pv. *morsprunorum* R2 leaf with pBBR-2	This study
R2-leaf pBBR-3	*P.s* pv. *morsprunorum* R2 leaf with pBBR-3	This study
R2-leaf pBBR-4	*P.s* pv. *morsprunorum* R2 leaf with pBBR-4	This study
R2-leaf pBBR-5	*P.s* pv. *morsprunorum* R2 leaf with pBBR-5	This study

Strain	Description	Reference
R2-leaf pBBR-6	*P.s* pv. *morsprunorum* R2 leaf with pBBR-6	This study
R2-leaf pBBR-7	*P.s* pv. *morsprunorum* R2 leaf with pBBR-7	This study
R2-leaf pBBR-8	*P.s* pv. *morsprunorum* R2 leaf with pBBR-8	This study
R2-leaf pBBR-9	*P.s* pv. *morsprunorum* R2 leaf with pBBR-9	This study
R2-leaf pBBR-10	*P.s* pv. *morsprunorum* R2 leaf with pBBR-10	This study
R2-leaf pBBR-11	*P.s* pv. *morsprunorum* R2 leaf with pBBR-11	This study
R2-leaf pBBR-12	*P.s* pv. *morsprunorum* R2 leaf with pBBR-12	This study
syr9644 pBBR-1	*P.s* pv. *syringae* 9644 with pBBR-1	This study
syr9644 pBBR-2	*P.s* pv. *syringae* 9644 with pBBR-2	This study
syr9644 pBBR-3	*P.s* pv. *syringae* 9644 with pBBR-3	This study
syr9644 pBBR-4	*P.s* pv. *syringae* 9644 with pBBR-4	This study
syr9644 pBBR-5	*P.s* pv. *syringae* 9644 with pBBR-5	This study
syr9644 pBBR-6	*P.s* pv. *syringae* 9644 with pBBR-6	This study
syr9644 pBBR-7	*P.s* pv. *syringae* 9644 with pBBR-7	This study
syr9644 pBBR-8	*P.s* pv. *syringae* 9644 with pBBR-8	This study
syr9644 pBBR-9	*P.s* pv. *syringae* 9644 with pBBR-9	This study
syr9644 pBBR-10	*P.s* pv. *syringae* 9644 with pBBR-10	This study
syr9644 pBBR-11	*P.s* pv. *syringae* 9644 with pBBR-11	This study
syr9644 pBBR-12	*P.s* pv. *syringae* 9644 with pBBR-12	This study
R1-5244 Δ*hrpA*	*P.s* pv. *morsprunorum* R1 5244 *hrpA* knockout mutant	This study
R2-leaf Δ*hrpA*	*P.s* pv. *morsprunorum* R2 leaf *hrpA* knockout mutant	This study
R2-leaf Δ*hopAR1*	*P.s* pv. *morsprunorum* R2 leaf *hopAR1* knockout mutant	This study
R2-leaf Δ*hopAZ1*	*P.s* pv. *morsprunorum* R2 leaf *hopAZ1* knockout mutant	This study
R2-leaf Δ*hopBB1*	*P.s* pv. *morsprunorum* R2 leaf *hopBB1* knockout mutant	This study
R2-leaf Δ*hopAR1*/Δ*hopAZ1*	*P.s* pv. *morsprunorum* R2 leaf *hopAR1*/*hopAZ1* double knockout mutant	This study

Strain	Description	Reference
R2-leaf Δ*hopAR1*/Δ*hopBB1*	*P.s* pv. *morsprunorum* R2 leaf *hopAR1*/*hopBB1* double knockout mutant	This study
R2-leaf Δ*hopAZ1*/Δ*hopBB1*	*P.s* pv. *morsprunorum* R2 leaf *hopAZ1*/*hopBB1* double knockout mutant	This study
R2-leaf Δ*hopAR1*/Δ*hopAZ1*/Δ*hopBB1*	*P.s* pv. *morsprunorum* R2 leaf *hopAR1*/*hopAZ1*/*hopBB1* triple knockout mutant	This study
R2-leaf Δ*hopAB3*	*P.s* pv. *morsprunorum* R2 leaf *hopAB3* knockout mutant	This study
syr9644 Δ*hopAR1*	*P.s* pv. *syringae* 9644 *hopAR1* knockout mutant	This study
R1-5300 Δ*hopAB1*	*P.s* pv. *morsprunorum* R1 5300 *hopAB1* knockout mutant	This study
R1-5300 Δ*avrPto1*	*P.s* pv. *morsprunorum* R1 5300 *avrPto1* knockout mutant	This study
R1-9657 Δ*hopAB1*	*P.s* pv. *morsprunorum* R1 9657 *hopAB1* knockout mutant	This study
PsavBPIC631 Δ*hopAB2*	*P.s* pv. *avellanae* BPIC631 *hopAB2* knockout mutant	This study

Table 5.1: All* P. syringae *strains and mutants used in this study.

Plasmid/*E. coli*	Description	Features	Reference
pBBR1MCS-5	Broad host range cloning vector	Gm^R	Kovach *et al.* 1995
pK18mobsacB	Broad host range suicide vector	Suc^S, Km^R	Shäfer *et al.* 1994
pRK2013	Triparental mating helper plasmid	Km^R,Tra^+, Mob^+, Col El replicon	Ditta *et al.* 1980
pRK2073	Triparental mating helper plasmid	Sp^R	Leong et al. 1982
pLAFR3	Cosmid vector	Tet^R, Tra^-, Mob^+, RK2 replicon, *cos*	Staskawicz *et al.* 1987
803	*E. coli* host for cosmid vectors		Wood *et al.* 1966
pBBR-1	pBBR1MCS-5 carrying $avrA_{R1\text{-}5300}$ CDS and promoter	Gm^R	This study
pBBR-2	pBBR1MCS-5 carrying $avrPto1_{R1\text{-}5300}$ CDS and promoter	Gm^R	This study
pBBR-3	pBBR1MCS-5 carrying $avrPto5_{RMA1}$ CDS and promoter	Gm^R	This study
pBBR-4	pBBR1MCS-5 carrying $avrRps4_{PsavBPIC631}$ CDS and promoter	Gm^R	This study
pBBR-5	pBBR1MCS-5 carrying $hopAA1_{R1\text{-}5300}$ CDS and promoter	Gm^R	This study

Plasmid/*E. coli*	Description	Features	Reference
pBBR-6	pBBR1MCS-5 carrying *hopAB1*$_{R1\text{-}5300}$ CDS and promoter	Gm^R	This study
pBBR-7	pBBR1MCS-5 carrying *hopAB2*$_{PsavBPIC631}$ CDS and promoter	Gm^R	This study
pBBR-8	pBBR1MCS-5 carrying *hopAB3*$_{RMA1}$ CDS and promoter	Gm^R	This study
pBBR-9	pBBR1MCS-5 carrying *hopAO2*$_{R1\text{-}9657}$ CDS and promoter	Gm^R	This study
pBBR-10	pBBR1MCS-5 carrying *hopAW1*$_{Pph1448A}$ CDS and promoter	Gm^R	This study
pBBR-11	pBBR1MCS-5 carrying *hopC1*$_{RMA1}$ CDS and promoter	Gm^R	This study
pBBR-12	pBBR1MCS-5 carrying *hopG1*$_{R1\text{-}5300}$ CDS and promoter	Gm^R	This study
pK18-1	pK18mobsacB::Δ*hrpA*$_{R1\text{-}5244}$	Km^R	This study
pK18-2	pK18mobsacB::Δ*hrpA*$_{R2\text{-}leaf}$	Km^R	This study
pK18-3	pK18mobsacB::Δ*hopAR1*$_{R2\text{-}leaf}$	Km^R	This study
pK18-4	pK18mobsacB::Δ*hopAZ1*$_{R2\text{-}leaf}$	Km^R	This study
pK18-5	pK18mobsacB::Δ*hopBB1*$_{R2\text{-}leaf}$	Km^R	This study
pK18-6	pK18mobsacB::Δ*hopAB3*$_{R2\text{-}leaf}$	Km^R	This study
pK18-7	pK18mobsacB::Δ*hopAR1*$_{syr9644}$	Km^R	This study
pK18-8	pK18mobsacB::Δ*hopAB1*$_{R1\text{-}5300}$	Km^R	This study
pK18-9	pK18mobsacB::Δ*avrPto1*$_{R1\text{-}5300}$	Km^R	This study
pK18-10	pK18mobsacB::Δ*hopAB1*$_{R1\text{-}9657}$	Km^R	This study
pK18-11	pK18mobsacB::Δ*hopAB2*$_{PsavBPIC631}$	Km^R	This study

Table 5.2: All Plasmids generated during this study, with names and features including antibiotic resistance. Km: kanamycin, GM: gentamycin, Tet: tetracycline, Sp: spectinomycin.

5.3.2 Candidate gene identification

To obtain candidate effector genes for cloning, bioinformatics was used to identify homologues of all known *P. syringae* effectors (downloaded from *pseudomonas-syringae*.org) in 104 strains across the species complex. A heatmap of presence and absence of effector genes in each strain's genome was generated using the R package ggplot2 (Wickham 2009) (more detail in section 4.3). Potential avirulence effectors were those that were absent from or pseudogenised in cherry pathogens, but had full-length homologues in close out-groups isolated from other plant species. Whereas, potential virulence-associated effectors were those that were statistically associated with cherry pathogenicity (in Bayestraits analysis, Chapter 4) and present in multiple cherry-infecting clades in the heatmap.

Protein and nucleotide sequences of candidate effectors were aligned and visualised using Geneious 7.1.9 (Kearse *et al.* 2012). The program Interproscan was used to identify important protein domains for annotation (Quevillon *et al.* 2005).

5.3.3 General DNA manipulations

All restriction enzymes and T4 DNA ligase were purchased from New England Biolabs (NEB). For cloning, the Phusion high-fidelity DNA polymerase (Thermofisher) was used, following manufactures instructions. For all routine PCRs, Taq 5X Master Mix (NEB) was used. All primers were purchased from Integrated DNA Technologies (IDT). Tables 5.3-5.4 list the primers used to generate and validate cloning vectors in this study. For gel electrophoresis, the gel was pre-stained with 1X GelRed (Biotium).

ID	Sequence	RS	Product	Vector	DNA source
*hopC1*_F1	GGTGGTCTCGAGGAGCTCCCCATAGTCCTTCC	XhoI	990	PBBR1MCS-5	RMA1
*hopC1*_R1	GGTGGTACTAGTGACATCCTTGCCCGAACAAT	SpeI		PBBR1MCS-5	
*hopAW1*_F1	GGTGGTCTCGAGTTTGAGCTGCCATGCCTTTT	XhoI	770	PBBR1MCS-5	Pph1448A
*hopAW1*_R1	GGTGGTACTAGTCCCCTGTTCACTACGCATAA	SpeI		PBBR1MCS-5	
*avrRps4*_F1	GGTGGTCTCGAGGGGACGGTACTGATAGCACG	XhoI	890	PBBR1MCS-5	PsavBPIC631
*avrRps4*_R1	GGTGGTACTAGTGCTTGAGGCAAGGGCTATGA	SpeI		PBBR1MCS-5	
*hopAA1*_F2	GGTGGTCTCGAGCTTAAAGCTCGCTTAATTTC	XhoI	18882	PBBR1MCS-5	R1-5300
*hopAA1*_R2	GGTGGTACTAGTCCCAACACACAGAGCTTTGAA	SpeI		PBBR1MCS-5	
*hopAB1*_F1	GGTGGTAAGCTTCTCAATGGGGAAGCCCTTGT	HindIII	1834	PBBR1MCS-5	R1-5300
*hopAB1*_R1	GGTGGTACTAGTCGATCACGGATTTCGCCTTG	SpeI		PBBR1MCS-5	
*hopAB2*_F1	GGTGGTAAGCTTGCGCTCCAGCTGGTTTTAAG	HindIII	2060/1976	PBBR1MCS-5	RMA1/PsavBPIC631
*hopAB2*_R1	GGTGGTACTAGTACATTGACGGGCTTTTGCTG	SpeI		PBBR1MCS-5	
*avrPto1*_F1	GGTGGTAAGCTTTTTCTGACAGTCTCCCGGGG	HindIII	728	PBBR1MCS-5	R1-5300
*avrPto1*_R1	GGTGGTACTAGTTGTATTTTCAGCTTAAAATTGCTTAGC	SpeI		PBBR1MCS-5	
*avrPto5*_F1	GGTGGTAAGCTTATGTCTCGAAATCATGCGCC	HindIII	629	PBBR1MCS-5	RMA1
*avrPto5*_R1	GGTGGTACTAGTATTGAGACTGCGGCGTCTC	SpeI		PBBR1MCS-5	
*avrA*_F1	GGTGGTCTCGAGACGCCGAATGTCGTTCTGAT	XhoI	3165	PBBR1MCS-5	R1-5300
*avrA*_R1	GGTGGTACTAGTCTTCCTGCCCATGACGGAAA	SpeI		PBBR1MCS-5	
*hopAO2*_F1	GGTGGTCTCGAGGGCTCCTCATCGTGGTGAAT	XhoI	1652	PBBR1MCS-5	R1-9657
*hopAO2*_R1	GGTGGTACTAGTGGCACTGGGCTTAAATGCAC	SpeI		PBBR1MCS-5	
*hopG1*_F1	GGTGGTCTCGAGTTGGTTGATCCTGGCTAGCG	XhoI	1709	PBBR1MCS-5	R1-5300
*hopG1*_R1	GGTGGTACTAGTCGGAGGGACTATCGGCATTC	SpeI		PBBR1MCS-5	

Table 5.3: Primer sequences used in this study for cloning genes into the pBBR1MCS-5 broad-host range expression vector.

ID	Sequence	RS	Product	SOE	Vector	DNA source
*hrpA*del_F1	GGTGGTGAATTCCTGGGTTCGACATCCACGATCAA	EcoRI	622/624	1192/1254	pK18mobsacB	R1-5244/R2-leaf
*hrpA*del_R1	CGCTCTTGCGGCCGCTTGGAACGGTACCAGCGTAATTGCTTGGTAC	none				
*hrpA*del_F2	CCGTTCCAAGCGGCCGCAAGAGCGCAGAGGGGGCTGCTACTTTG	none	592/654			
*hrpA*del_R2	GGTGGTGGATCCTCTTCCTTCAGTTCGTTCACCCA	BamHI				
*hopAR1*del_F1	GGTGGTGAATTCAAGGAGTCATGACCAACCGG	EcoRI	515/514	947/951	pK18mobsacB	R2-leaf/syr9644
*hopAR1*del_R1	CGCTCTTGCGGCCGCTTGGAACGGTGATCACTCTCGCTCGATGC	none				
*hopAR1*del_F2	CCGTTCCAAGCGGCCGCAAGAGCGGTTGGTAGGCTTCGTCCTGA	none	456/461			
*hopAR1*del_R2	GGTGGTGGATCCCGCCTTATGTTCAATCGATCCCAC	BamHI				
*hopBB1*del_F1	GGTGGTGAATTCCGCCAAGCCGAAAACTGTTT	EcoRI	578	1115	pK18mobsacB	R2-leaf
*hopBB1*del_R1	CGCTCTTGCGGCCGCTTGGAACGGAACACCTCTCGCATCTGGAC	none				
*hopBB1*del_F2	CCGTTCCAAGCGGCCGCAAGAGCGTCTGTTTGACTCCGCTGCAT	none	561			
*hopBB1*del_R2	GGTGGTGGATCCAAGAATTGGTTGCCAACGCG	BamHI				
*hopAZ1*del_leaf_F1	GGTGGTGAATTCCGGCCTGACAACCAAGATTC	EcoRI	507	1016	pK18mobsacB	R2-leaf
*hopAZ1*del_leaf_R1	CGCTCTTGCGGCCGCTTGGAACGGACCGTTAAGCCGCTTTACGT	none				
*hopAZ1*del_leaf_F2	CCGTTCCAAGCGGCCGCAAGAGCGTGGCCGTTGTTGATTAGGCT	none	533			
*hopAZ1*del_leaf_R2	GGTGGTGGATCCTAGATGATGGCGTTGGCGAT	BamHI				
*hopAB1*_del_5300_F1	GGTGGTGAATTCAGACTAGGGCCGTTTTCTGC	EcoRI	582	1188	pK18mobsacB	R1-5300/R1-9657
*hopAB1*_del_5300_R1	CGCTCTTGCGGCCGCTTGGAACGGCATGGTTGCGAACTGTTCCG	none				
*hopAB1*_del_5300_F2	CCGTTCCAAGCGGCCGCAAGAGCGCAAGGCGAAATCCGTGATCG	none	630			
*hopAB1*_del_5300_R2	GGTGGTGGATCCCCGGAGAAGCATTTCCGGAT	BamHI				
*hopAB2/3*_del_F1	GGTGGTGAATTCCAGCGTGGTGTTGTCTCTGA	EcoRI	518/521	939/942	pK18mobsacB	RMA1/PsavBPIC631
*hopAB2/3*_del_R1	CGCTCTTGCGGCCGCTTGGAACGGCTCTCCGATCTTCTGCCTGC	none				

*hopAB2/3*_del_F2	CCGTTCCAAGCGGCCGCAAGAGCGCAGCAAAAGCCCGTCAATGT	BamHI	445			
*hopAB2/3*_del_R2	GGTGGTGGATCCTCAGTGCCAGTGGTGTGTTT	none				
*hopC1*_del_F1	GGTGGTGAATTCCTTGGGGAGATCAAGGCCAC	EcoRI	544	1065	pK18mobsacB	RMA1
*hopC1*_del_R1	CGCTCTTGCGGCCGCTTGGAACGGTTGGTTCCCTGGCCTTTCTG	none				
*hopC1*_del_F2	CCGTTCCAAGCGGCCGCAAGAGCGGGATGTCCGCAGCTGAGTAG	none	545			
*hopC1*_del_R2	GGTGGTGGATCCGACGGCATCAGCTACGAGAA	BamHI				
*avrPto1*_del_F1	GGTGGTGAATTCACCTGATCTTCGCTGCTCAC	EcoRI	437	886	pK18mobsacB	R1-5300
*avrPto1*_del_R1	CGCTCTTGCGGCCGCTTGGAACGGGGTTCCACACCGATGGTAGA	none				
*avrPto1*_del_F2	CCGTTCCAAGCGGCCGCAAGAGCGTTCTAGGTTGGGTACGGCCT	none	473			
*avrPto1*_del_R2	GGTGGTGGATCCATTGGAGCGAATAGAGCGGG	BamHI				
*hopAB3*_del_leaf_F1	GGTGGTGAATTCCAGCGTGGTGTTGTCTCTGA	EcoRI	521	957	pK18mobsacB	R2-leaf
*hopAB3*_del_leaf_R1	CGCTCTTGCGGCCGCTTGGAACGGCTCTCCGATCTTCAGGCTGC	none				
*hopAB3*_del_leaf_F2	CCGTTCCAAGCGGCCGCAAGAGCGATCCGAGCAGCCCTGAATTT	none	460			
*hopAB3*_del_leaf_R2	GGTGGTGGATCCCACAGACCTACGAAGTGGCA	BamHI				

Table 5.4: Primer sequences used in this study to create pK18-mobsacB-based deletion vectors.

Candidate avirulence genes and their upstream *hrp*-promoters were inserted into the multiple-cloning site (MCS) of pBBR1MCS-5 using two different restriction sites to ensure correct orientation to allow expression via the *lac* promoter in the vector. Both the *hrp* box and *lac* promoter were used to ensure a high expression. Primers containing the desired restriction sites were used to amplify the insert by PCR. The product size was confirmed by gel electrophoresis and gel purified if non-specific bands were present using the NucleoSpin Gel and PCR Clean-up (Machery Nagel). Insert DNA and the cloning vector then underwent a restriction digest with the desired enzymes following manufacturer's instructions. The vector DNA was then dephosphorylated using shrimp-alkaline phosphorylase r(SAP) (NEB), to remove phosphorylated ends to prevent any re-ligation. Both vector and insert DNA were then purified using the NucleoSpin Gel and PCR Clean-up kit. Ligation was then performed with a 1:3 ratio of vector:insert at 16°C overnight. This ligation mix was then purified and 5 µl of DNA with a concentration ≥1 ng/µl was used for *E. coli* transformations.

The presence of an insert in each vector was validated by restriction digests of the plasmid and PCR using the M13 primers which are universal to all pUC-based cloning sites (Yanisch-Perron *et al.* 1985). The correct product size was confirmed by gel electrophoresis.

Effector deletions were generated by homologous recombination using the pK18mobsacB system. First 500bp flanking regions of the effector gene were PCR amplified and the correct band purified from a gel. The two PCR amplicons contained homologous regions in the R1 and F2 primer overhangs, allowing them to be joined by Splice Overlap Extension PCR. This PCR was performed in a total volume of 50 µl and a low annealing temperature of 50 °C. The product was then gel purified. The product and vector were then restriction digested using manufacturer's instructions and ligated at 16 °C overnight to ensure efficient ligation. After purification 5µl of DNA with a concentration ≥1 ng/µl was used for *E. coli* transformations. The presence of the insert was confirmed as above.

To transform these vectors into *E. coli*, the NEB High-affinity transformation protocol was used with C2987H/C2987I cells as per manufacturer's instructions. Blue-white screening was used to confirm the presence or absence of an insert followed by restriction digest and PCR validation.

5.3.4 *P. syringae* transformations

5.3.4.1 Conjugation

Both pBBR1MCS-5 and pK18mobsacB plasmids were conjugated into *Pseudomonas.* The pBBR1MCS-5 plasmid is replicative, whilst the pK18mobsacB plasmid is non-replicative and integrates via homologous recombination with the genome. Conjugation was achieved using tri-parental mating involving *P. syringae*, the cloned plasmid and a helper plasmid that facilitates transfer. For vectors containing the kanamycin resistance gene, such as pK18mobsacB, the streptomycin resistant helper pRK2073 was used. For all other vectors the helper strain pRK2013 was used. *Pseudomonas* and the *E. coli* strains were grown overnight in LB. The cells were pelleted by spinning at 3500 g for 10 mins. *E. coli* were washed twice with fresh LB to remove any antibiotics. The bacteria were then mixed in a 2:1:1 ratio (*P. syringae*:helper:vector). The bacteria were then spun down again (3500 g, 10 mins), the supernatant removed and spread onto KBA with no antibiotics to form a lawn. The plates were left at 25 °C for one day to allow conjugation to occur and then a 10 µl loop of bacteria was streaked onto KBA, nitrofurantoin and the desired antibiotic to select for transconjugants. After two days of incubation, colonies were transferred to plates consisting of KBA and the desired antibiotic to minimise time on Nitrofurantoin.

When performing deletions, the resulting *P. syringae* colonies had integrated the vector, but needed a second round of selection to delete the gene of interest. They were counter-selected by picking a colony, re-suspending in 1 ml of sterile distilled water and spreading 50 µl on no-salt LBA with 15% (w/v) sucrose (Hmelo *et al.* 2015). The expression of the *sacB* gene induced sensitive to sucrose, meaning that any colonies were assumed to have undergone a second homologous recombination to lose the *sacB* and kanamycin resistance cassette or to have lost the plasmid sequence

completely. These colonies were plated onto replicate plates of KBA or KBA and kanamycin. Colonies that failed to grow on kanamycin were then screened by a PCR. The PCR used primers that flanked the gene of interest, and the product size indicated if the gene had been successfully deleted. Where more than one independent gene deletion mutant could be obtained, all were kept to provide replicate mutants in pathogenicity assays.

5.3.4.2 Electroporation

Where conjugation attempts failed to produce transformants, an electroporation method as in Cadoret *et al.* (2014) was trialled. *P. syringae* was grown overnight in 10 ml of LB and then 1 ml was transferred into 9 ml of fresh LB and grown to an OD_{600} of 0.6. The bacteria were then spun down at 2,300 g for 10 mins and washed in an equal volume of cold sterile filter-sterilised 300mM sucrose solution. This process was repeated, gradually concentrating the cells by reducing the volume of sucrose solution to 0.5x and then 0.01x, producing a final volume of 100 µl. 10 µl of vector DNA (concentration not exceeding 1 µg/ml) suspended in deionised water was then mixed with 80 µl of bacteria cells and chilled for 30 mins on ice. Cells were then transferred to a 2 mm electroporation cuvette and electroporated using a *E.coli* pulser (Bio-Rad) set at 2.5 kV. Cells were then plated out onto KBA plus the desired antibiotic and incubated for two days at 25 °C.

5.3.5 PCR validation

The M13 PCR primers were used to confirm the presence of the pBBR1MCS-5 vector with the correct insert. To determine if gene deletions using pK18mobsacB had occurred, a PCR was performed using primers designed on regions outside of both the gene and homologous recombination flanking regions. Sanger sequencing was used to confirm that several of the deletions were correct.

5.3.6 Genomic libraries

5.3.6.1 Creation of genomic libraries

Genomic libraries were produced for several *P. syringae* strains using the cosmid vector pLAFR3 (Staskawicz *et al.* 1987). High-molecular weight bacterial genomic DNA was isolated using a CTAB-based method (Feil *et al.* 2012). Approximately 60 µg of DNA was then partially digested with 0.03U of Sau3a (NEB) at 37 °C for 30 mins followed by a 70 °C inactivation step for 10 mins. This was run slowly on a gel overnight at 2 V/cm. To predict the size of the fragmented DNA, two ladders consisting of Lambda DNA digested with either XhoI or EcoRI were used. DNA ranging from 15-30 Kb was then gel extracted using the Zymoclean Large Fragment Recovery Kit (Zymo Research). Sizes were confirmed by running 1 µl of DNA on a gel.

The vector pLAFR3 was digested with BamHI for 1 hour and dephosphorylated using r(SAP). Digested genomic DNA was ligated into pLAFR3 at a ratio of 3:1 (insert:vector) using T4 DNA ligase (NEB). The ligation was performed overnight at 16 °C and then stored at 4 °C for 24 h to allow complete ligation.

The cosmid DNA was then packaged using a MaxPlax packaging extract (Cambio) and transformed into *E. coli* 803. For each library, 10 colonies were mini-prepped (Macharey-Nagel) and cosmid DNA digested with BamHI and BglII to confirm the presence of pLAFR3 and variable insert sizes. Approximately 500-1000 clones were obtained for each library created.

5.3.6.2 Conjugation of genomic libraries into *P. syringae*

Genomic libraries were conjugated into a recipient *Pseudomonas* strain. After overnight incubation in LB and tetracycline, *E. coli* libraries were spun down (3500 g, 10 mins) and washed several times in fresh LB without antibiotics. Plate-mating was used where a 5 µl drop of *Pseudomonas* suspension was mixed with 2.5 µl of *E.coli* suspension harbouring each pLAFR3 clone and 2.5 µl with the helper plasmid pRK2013 on KBA without antibiotics. This mixture was incubated at 25 °C for 24 h to allow

mating to occur. The colonies were then transferred to KBA with nitrofurantoin and tetracycline to select for *Pseudomonas* containing the pLAFR3 plasmid. To confirm pLAFR3 was present, several colonies were tested by PCR of the *tetA* gene (tetracycline resistance). The *Pseudomonas* conjugants were then stored at -80 °C in 40% glycerol.

5.3.7 Pathogenicity assays

5.3.7.1 Plant material

All assays were performed on detached glasshouse grown leaves. This method was used because when trees are grown outside, leaves become quickly contaminated with bacteria and fungi which would prevent accurate population studies. The cherry cultivar Van was used as it is reportedly universally susceptible to bacterial canker (Long & Olsen 2013). In Winter/Spring 2016, trees were periodically moved into the 2 °C cold-store for 2-3 months to ensure dormancy. When leaves were required, the trees were brought into an unheated glasshouse (ambient summer temperatures and light were sufficient) to induce leaf formation. Young, fully-expanded 1-2 week old leaves were used in all pathogenicity tests.

5.3.7.2 Population counts

Mutant strains were tested for differences in virulence on cherry by measuring population counts in detached leaves. To prepare the inoculum, overnight bacterial cultures were spun down (3500 g, 10 mins) and re-suspended in 10mM $MgCl_2$ to an OD_{600} of 0.2. A 100-fold dilution was then performed to produce an inoculum corresponding to $2x10^6$ CFU/ml. Bacteria were syringe-infiltrated into detached cherry leaf material. The leaves were then placed in plastic trays containing a layer water agar (10 g/L) covered in damp paper towel to maintain a high humidity. The tray was sealed inside a transparent bag and incubated at 22 °C (16h light, 8h dark). Bacterial population counts were measured every 2 days, or at the start (0 days) and end (10 days). Day 0 populations were measured to check bacteria were present in the leaves and at a similar concentration. At least three replicate leaves were inoculated with each isolate, with the three leaves coming from different cherry trees. The

experimental design of population counts is detailed in Figure S3.2. A 10mM $MgCl_2$ control was included to check for contamination of leaf material.

To measure bacterial population growth, leaf disks were excised using a sterile cork-borer (0.5 cm) and homogenised in 10mM $MgCl_2$. A dilution series was plated out to confirm bacterial concentrations (CFU/ml). Three technical replicates of each dilution series were performed and each experiment was repeat twice per strain analysed.

5.3.7.3 Symptom scoring

The symptoms induced by bacterial mutants were also scored over time. These experiments were done as above, except the bacterial concentration was $2x10^8$ CFU/ml as at this concentration a hypersensitive response is visible in cherry leaves. Different isolates were inoculated onto the same leaf to allow comparisons in the timing of symptom development to be made. This involved taking pictures after 16 h and then every day. Disease lesions were scored as either 0 (no lesion), 1 (limited browning), 2 (browning in <50% of inoculated area), 3 (browning in >50% of inoculated area) and 4 (complete browning). This experiment was repeated twice.

5.3.7.4 Screening of genomic libraries for virulence phenotypes

To screen *P. syringae* strains containing genomic libraries a more high-throughput approach was taken. Bacteria were grown overnight in 1 ml of LB and tetracycline in 2.2 ml 96-well plates (Alpha Laboratories). They were then spun down (3500 g, 10 mins) and re-suspended in 10mM $MgCl_2$. To measure the concentration, optical density was calculated on the CLARIOstar plate reader (BMG LABTECH). Bacteria were then diluted down to an OD_{600} of 0.002 to achieve suspensions of approximately $2x10^6$ CFU/ml for inoculations.

The pathogenicity assay was similar to above, however due to the large number of isolates being tested, four separate clones were inoculated per leaf with two replicate leaves for each. Each leaf was also inoculated with the controls. These included the genomic library recipient *P. syringae* wild-type strain, the donor *P. syringae* strain, the

recipient *P. syringae* with the empty pLAFR3 vector and a 10mM $MgCl_2$ control. This experiment was only performed once.

5.3.7.5 Tobacco HR assay

Mutant strains were tested for the ability to induce a hypersensitive response in *Nicotiana tabacum* cv. Samsung. Six-week old plants were used for pathogenicity assays. Bacterial inoculum was prepared as above and adjusted to a concentration of $2x10^8$ CFU/ml. Fully expanded leaves were syringe-infiltrated with the inoculum and incubated at 22 °C for 1 day before symptoms were assessed. This experiment was only performed once.

5.3.8 Statistical analysis

All statistics were performed using R software (R Core Team 2012). Graphs were produced using the R package gplots (Warnes *et al.* 2016). Statistical analysis generally involved the use of ANOVA on raw or log_2 transformed data (in cases where the ANOVA residuals were not normal). For population counts, the strain was treated as a fixed factor, whilst leaf replicate, with technical replicate (individual 10µl droplets on the plate) nested within, were treated as random effects. The model: aov(log_2(cfu+0.1) ~ strain + leaf/replicate). Where more than one deletion mutant could be obtained from cloning, all were included in pathogenicity tests. Here the ANOVA grouped the strains with the same genotype together e.g. strains with the same gene deletion were all compared to the wildtype control. The model: aov(log_2(cfu+0.1) ~ genotype + leaf/replicate). Post-hoc pairwise comparisons between strains was performed using the Tukey-HSD test, which binned strains into groups based on similarity. To analyse the rate of symptom development on leaves, the slope between 0 and 24hpi was calculated and an ANOVA performed on slope values. All ANOVA tables are presented in the thesis appendix.

5.4 Results

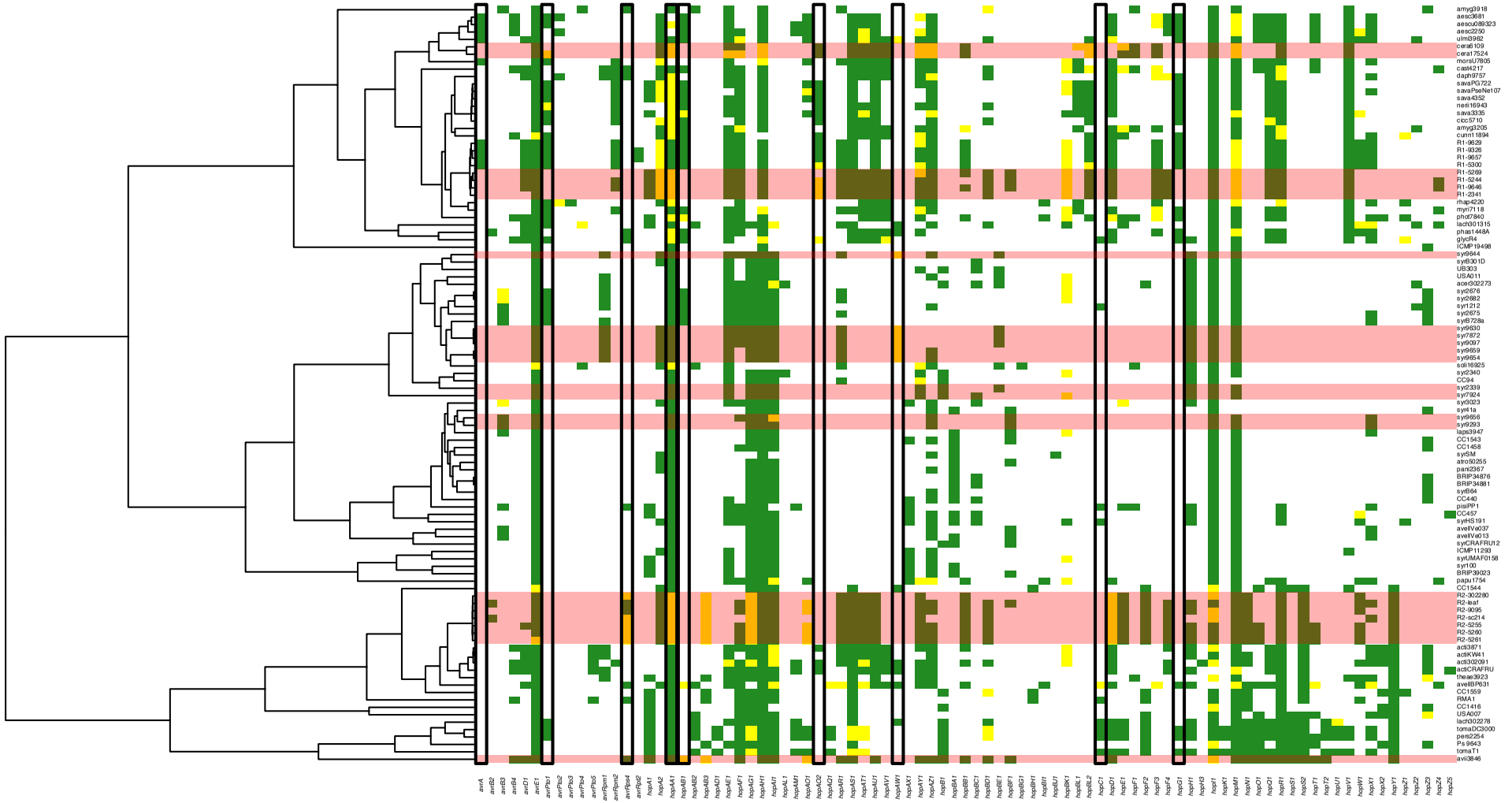
In this study, the host specificity of *P. syringae* on cherry was studied. First candidate effector genes involved in host specificity were identified using bioinformatics. Effectors from non-pathogenic strains isolated from other plant species were tested to

determine if they are avirulence factors in cherry and therefore prevent non-pathogen population growth. Those effectors revealed to be avirulence factors, were deleted in the non-pathogen strains to determine if deletion allowed these strains to achieve greater populations in cherry. Putative virulence-associated genes in cherry pathogens were also deleted to determine their importance in virulence. Finally, genomic-libraries were used to try and identify host-specificity factors in an unbiased approach.

5.4.1 Identification of candidate avirulence effectors

Bioinformatics was used to identify candidate effectors genes that may influence host specificity and virulence of *P. syringae* on cherry. To do this a heatmap of gene presence and absence, including putative pseudogenes, was constructed. By comparing cherry-pathogenic strains with close out-groups isolated from other plant species, nine effectors were identified (*avrA1*, *avrPto1*, *avrRps4*, *hopAA1*, *hopAB1*, *hopAO2*, *hopAW1*, *hopC1* and *hopG1*), that were present in closely-related out-groups but absent from or pseudogenised in cherry pathogens (Figure 5.1). The absence of these effector genes in the cherry pathogen genomes suggests that the products of these genes may be avirulence effectors detected by the cherry immune system. Cloning was first focused on *Psm* R1 as this highly homogenous clade contains both pathogens and non-pathogens of cherry. Any variation in effector complement between strains within the clade are likely to be due to differences in host specificity as opposed to phylogenetic distance. Potential avirulence effectors included *avrA1*, *avrPto1*, *hopAA1*, *hopAB1*, *hopAO2* and *hopG1*, which were present in non-pathogenic members of *Psm* R1 but absent from pathogenic strains. These effector genes were cloned from the non-pathogenic strain R1-5300 (except *hopAO2* which was cloned from R1-9657).

Figure 5.1 (overleaf): Heatmap of effector gene presence across* P. syringae *strains with candidate avirulence genes highlighted. *The heatmap was generated using R gplots. The dendrogram clustered strains based on similarity of effector content. The green squares indicate presence of a full-length homologue of the effector gene, whereas yellow squares indicate that the gene is disrupted or truncated in some way. Strains pathogenic to cherry are highlighted in red. Putative avirulence genes are outlined in black.*


amyg3918
aesc3681
aesc089323
aesc2250
ulmi3962
cera6109
cera17524
morsU7805
cast4217
daph9757
savaPG722
savaPseNe107
sava4352
neii16943
sava3335
ctcc5710
amyg3205
cunn1894
R1-9629
R1-9326
R1-9657
R1-5300
R1-5289
R1-5244
R1-9646
R1-2341
rhap4220
myri7118
phot7840
lach301315
phae1448A
glycR4
ICMP19498
syr9644
syrB301D
UB303
USA011
aceri302273
syr2676
syr2682
syr1212
syr2675
syr8728a
syr9630
syr7872
syr9097
syr9659
syr9654
soli16925
syr2340
CC94
syr2339
syr7924
syr3023
syr1a
syr9656
syr9293
laps3947
CC1543
CC1458
syr5M
atro50255
pani2367
BRIP34876
BRIP34881
syr564
CC440
pisiPP1
CC457
syrHS191
avelIVe037
avelIVe013
sycCRAFRUI2
ICMP11293
sytUMAF0158
syr100
BRIP39023
papu1754
CC1544
R2-302280
R2-leaf
R2-9095
R2-sc214
R2-5255
R2-5260
R2-5261
acti3871
actKW41
acti302091
actiCRAFRU
theae3923
aveiBRI931
CC1559
RMA1
CC1416
USA007
lach302278
tomaDC3000
pers2254
Ps 9643
tomaT1
avii3846
avrA
avrB2
avrB3
avrB4
avrD1
avrE1
avrPto1
avrPto2
avrPto3
avrPto4
avrPto5
avrRpm1
avrRpm2
avrRps4
hopA1
hopA2
hopAA1
hopAB1
hopAB2
hopAB3
hopAD1
hopAE1
hopAF1
hopAG1
hopAH1
hopAI1
hopAL1
hopAM1
hopAO1
hopAO2
hopAQ1
hopAR1
hopAS1
hopAT1
hopAU1
hopAV1
hopAW1
hopAX1
hopAY1
hopAZ1
hopB1
hopBA1
hopBB1
hopBC1
hopBD1
hopBE1
hopBF1
hopBG1
hopBH1
hopBI1
hopBJ1
hopBK1
hopBL1
hopBL2
hopC1
hopD1
hopE1
hopF1
hopF2
hopF3
hopF4
hopG1
hopH1
hopH3
hopI1
hopK1
hopM1
hopN1
hopO1
hopO1
hopR1
hopS1
hopS2
hopT1
hopT2
hopU1
hopV1
hopW1
hopX1
hopX2
hopY1
hopZ1
hopZ2
hopZ3
hopZ4
hopZ5

Several effectors from other *P. syringae* strains were also examined. One was *avrRps4* which was absent or pseudogenised in most cherry-infecting strains. Three cherry-infecting strains possessed the full-length effector gene (R2-leaf, R2-9095 and *P.s* pv. *avii* 3846), however they lacked the KRVY domain shown to be important for virulence functions (Figure 5.2) (Sohn *et al.* 2009). Therefore, a full-length homologue that contained the KRVY domain was cloned from *P.s* pv. *avellanae* (*Psav*) BPIC631, which is closely related to *Psm* R2.

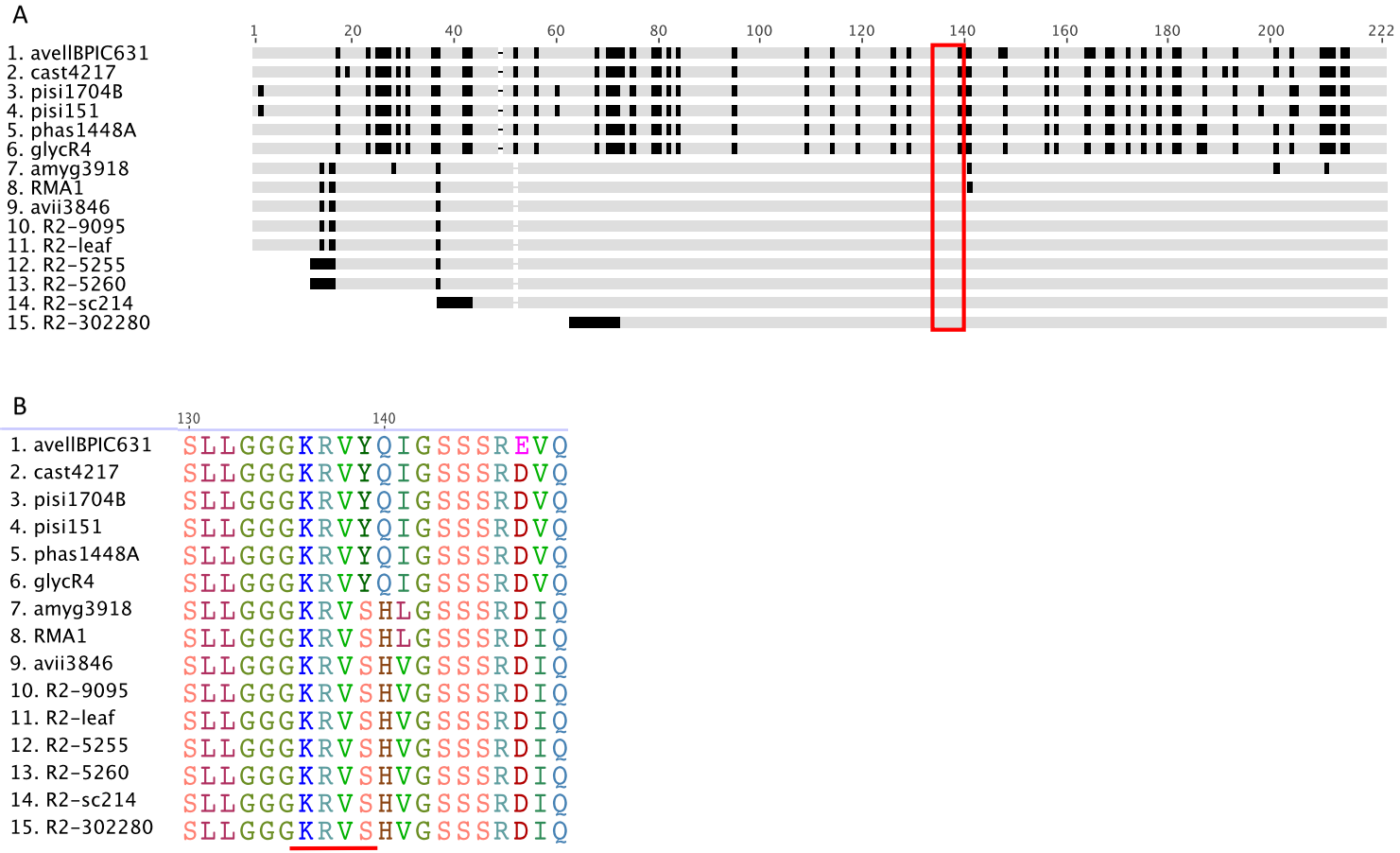


Figure 5.2: AvrRps4 alignment. *A: Amino acid alignment of AvrRps4. The alignment is colour-coded based on similarity, where identical residues are in grey, whereas dissimilar residues appear in black. This protein is truncated in several Psm R2 strains (12-15). The red box indicates the location of the important KRVY motif. B: Close-up of alignment showing that the KRVY motif is not present in all Psm R2 strains and several other pathovars.*

The *hopAW1* effector gene was also cloned to study its role in host specificity. This was because *hopAW1* is absent from most cherry pathogens. The gene is present in six *Pss* strains, but they have experienced two independent mutations that disrupt the beginning of the gene. *Pss* strains syr9654, syr9659 and syr9644 are missing the first 74bp, whilst in syr9630, syr9097 and syr7872 a single point mutation at position 18 (C to A) has led to a stop codon (TAA), meaning that an alternative start codon downstream may act as the beginning of the *hopAW1* gene. Figure 5.3 shows the nucleotide alignment and putative HopAW1 proteins produced if the genes are expressed. The full-length *hopAW1* gene was cloned from Pph1448A to see if the protein is detected by the cherry immune system.

Finally, *hopC1* was cloned as this effector is absent from all cherry-infecting strains and has been shown to be an avirulence factor in bean species (Baltrus *et al.* 2012), suggesting that a similar immune response may be present in cherry. The *hopC1* gene was cloned from the *Aquilegia vulgaris* pathogen RMA1 as this strain is basal to the clade containing *Psm* R2, and this effector may have been lost in lineages leading to cherry pathogens.

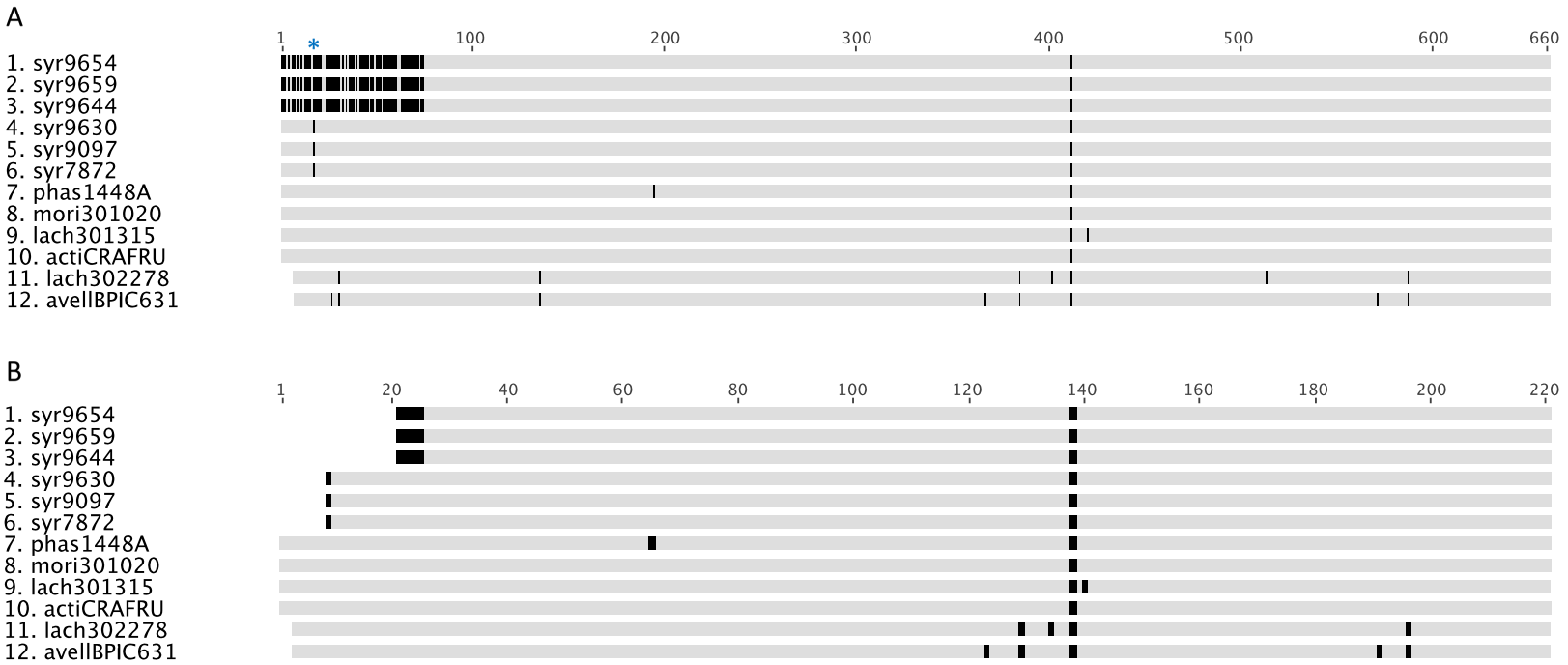


Figure 5.3: HopAW1 alignment*. A: Nucleotide alignment of the* hopAW1 *gene in* Pss *strains and other pathovars.* Pss *strains syr9654, 9659 and 9644 have a divergent sequence at the start of the gene.* Pss *strains syr9630, 9097 and 7872 have a point mutation leading to a stop codon (*). The alignment is colour-coded based on similarity; identical bases/amino acids are in grey, whereas dissimilar bases/amino acids appear in black. B: Alignment of amino acids of putative HopAW1 proteins, truncated at the N-terminus in all* Pss *strains.*

5.4.2 Two effector genes act as avirulence factors in cherry

The nine candidate effectors were cloned into pBBR1MCS-5 as described (see 5.3.3). PCR was used to validate that the gene was inserted into pBBR1MCS-5 (Figure S5.2). The constructs were conjugated into three pathogenic strains (R1-5244, R2-leaf and syr9644). It was originally planned to transform these vectors into syr9097 as this was the most well characterised *Pss* strain - sequenced with PacBio, and included in all previous pathogenicity tests (Chapter 3 and 4). However, this strain was recalcitrant to transformation via either conjugation or electroporation. Therefore, the strain syr9644 was used. As a control, Δ*hrpA* mutants were obtained for R1-5244 and R2-leaf. These mutants could not produce a type III secretion system so would therefore be non-pathogenic in cherry leaves, as they would fail to secrete effectors. To validate that these strains could not secrete effectors, they were inoculated into *Nicotiana tobacum*. The wild-type strains secreted effectors which induced a hypersensitive response (HR) in tobacco leaves. As expected, the Δ*hrpA* mutants both failed to induce the HR (Figure S5.1). The Δ*hrpA* deletion mutants were included as a control in the later avirulence effector screens to see if expression of any putative *avr* genes could reduce pathogenicity to a similar level as non-pathogenic mutants.

An initial population count was conducted on R2-leaf mutants and transconjugants over time for ten days (Figure 5.4). All strains except the Δ*hrpA* mutant, pBBR1MCS-5 with *hopAB1* and pBBR1MCS-5 with *hopC1* produced disease lesions and grew to exceed 10^6 CFU/ml. An ANOVA was performed on the $\log_2$ transformed CFU/ml for day 10 populations. There were significant differences between strains ($p<0.01$, $df=11$) and also a significant difference between leaves ($p=0.03$, $df=2$). The three replicate leaves were obtained from different cherry trees and therefore subtle physiological differences could have led to variations in bacterial population growth. A post-hoc Tukey-HSD analysis showed that the Δ*hrpA* mutant and strains containing the pBBR1MCS-5 with *hopAB1*, *hopC1* and *hopAW1* were significantly different from the WT control. Bacteria expressing *hopAB1* and *hopC1* reached final populations not significantly different from the Δ*hrpA* mutant indicating that expression of these genes severely reduced the virulence of the R2-leaf strain *in planta*.

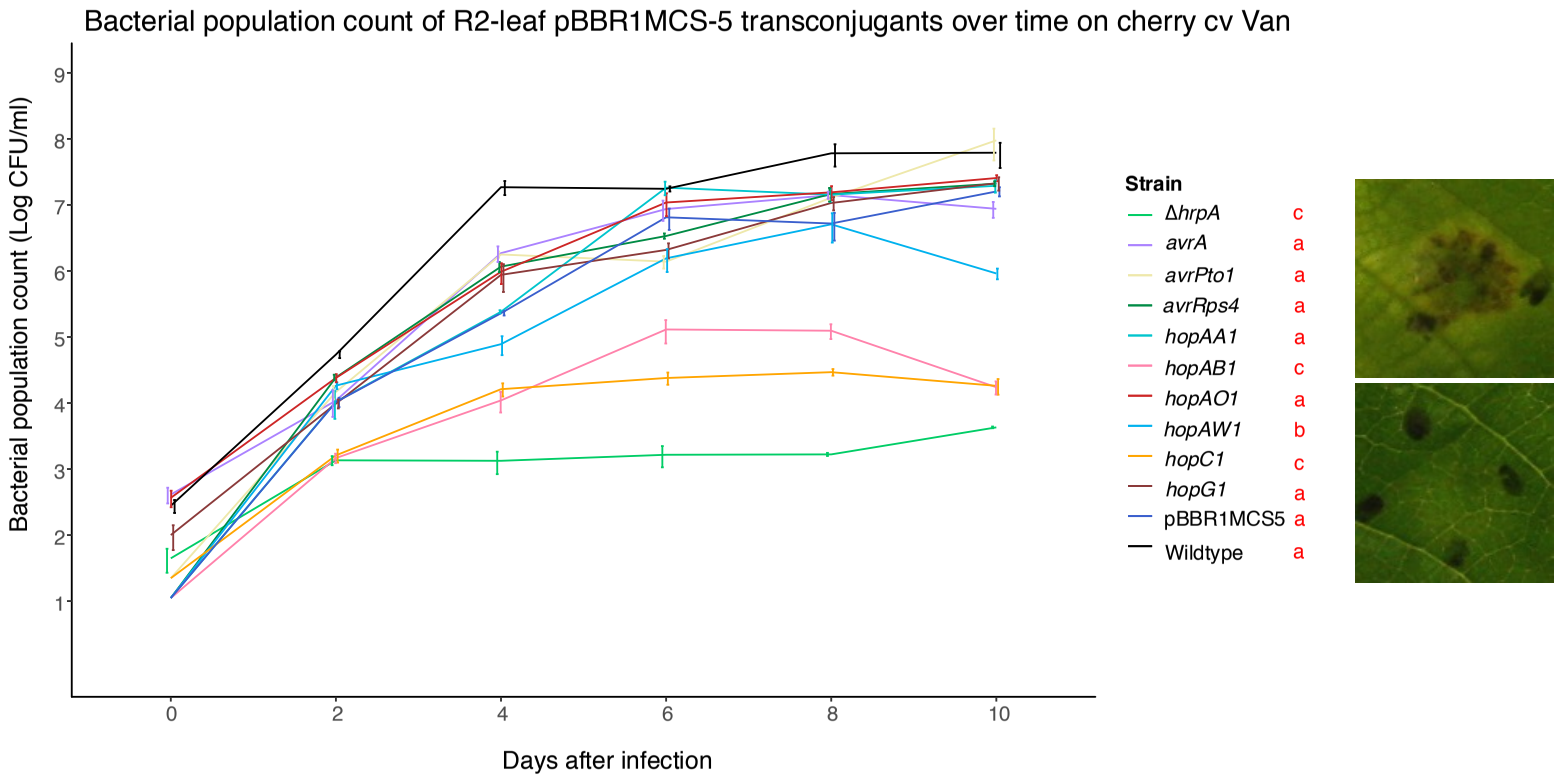


Figure 5.4: Time-course of in planta population growth of* Psm *R2-leaf wildtype, Δ*hrpA *mutant and R2-leaf expressing candidate avirulence effector genes. *Values represent the mean of three replicates and error bars shows standard error above and below the mean. Controls included the wildtype strain, a strain containing the empty pBBR1MCS-5 vector (to check the vector alone does not affect population growth) and a Δ*hrpA *deletion mutant (which is non-pathogenic as it cannot secrete effectors). Representative images of the disease lesion and no lesion are presented. Black marks are where a pen was used to mark the location of inoculation. Tukey-HSD significance groups (p=0.05, confidence level: 0.95) for the day 10 populations of each strain are presented in red.*

To see if the expression of these genes also had an effect on the virulence of the other cherry pathogens R1-5244 and syr9644, additional population counts were performed. Day 0 population counts for all leaf population assays are presented in Figures S5.6-S5.11). Figure 5.5 shows the 10-day post-inoculation bacterial counts for all three pathogens. There were significant differences between strains for both R1-5244 ($p<0.01$, $df=11$) and syr9644 ($p<0.01$, $df=10$). Several effector genes showed variation in the ability to reduce pathogen growth *in planta*. In R1-5244, slight but significant reductions were seen in strains expressing *avrA1* and *avrRps4*, whilst in R2-leaf, expression of *hopAW1* caused a significant reduction in growth. The expression of *hopAB1* and *hopC1* consistently reduced bacterial growth in all three pathogens and prevented symptom development, indicating the products of these effector genes may be avirulence factors detected by the cherry immune system. Therefore, further work focused on these two effectors and confirmed that they reduced population growth in four independent experiments.

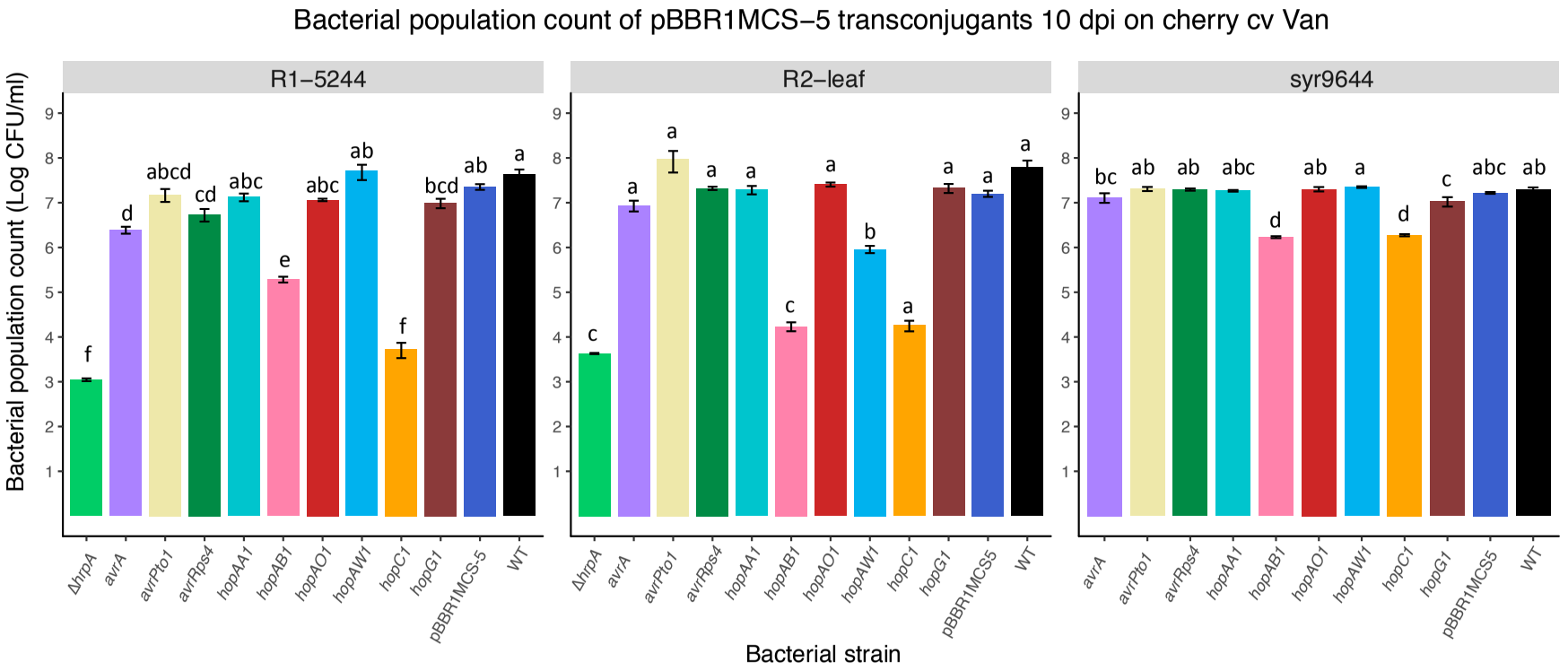


Figure 5.5: Ten-day population counts of cherry pathogens (R1-5244, R2-leaf and syr9644) expressing candidate avirulence genes. *Values represent the mean of three replicates and error bars shows standard error above and below the mean. Controls included the wildtype strain, a strain containing the empty pBBR1MCS-5 vector and a Δ*hrpA *deletion mutant (for R1-5244 and R2-leaf). A separate ANOVA was performed for each cherry pathogen (R1-5244, R2-leaf and syr9644) and the Tukey-HSD significance groups (p=0.05, confidence level: 0.95) for each strain are presented above each bar.*

HopAB1 belongs to the HopAB family of effector proteins. To see if other members of HopAB reduced bacterial population growth in cherry leaves, the *hopAB2* and *hopAB3* genes were cloned from *P.s.* pv. *avellanae* BPIC631 and RMA1, and expressed in pathogenic strains R1-5244, R2-leaf and syr9644. The virulence of these strains was assessed by measuring day 10 population counts. Figure 5.6 shows that compared to the empty vector control, all three pathogens expressing the *hopAB1* and *hopC1* genes exhibited reduced population growth. The ANOVA revealed for all three cherry pathogens, there were significant differences between strains ($p<0.01$, $df=4$ for all). Those strains expressing *hopAB2* and *hopAB3* grew to statistically higher levels than *hopAB1*-expressing mutants, and in syr9644, the expression of *hopAB2* led to a final population count not significantly different from the control. The different alleles of HopAB vary widely at the amino acid level (Figure 5.7). Therefore, allelic differences between members of HopAB, may mean that HopAB2 and HopAB3 may not be inducing immune response as strongly as HopAB1, or perhaps the bacterial populations are able to overcome immunity towards these effectors.

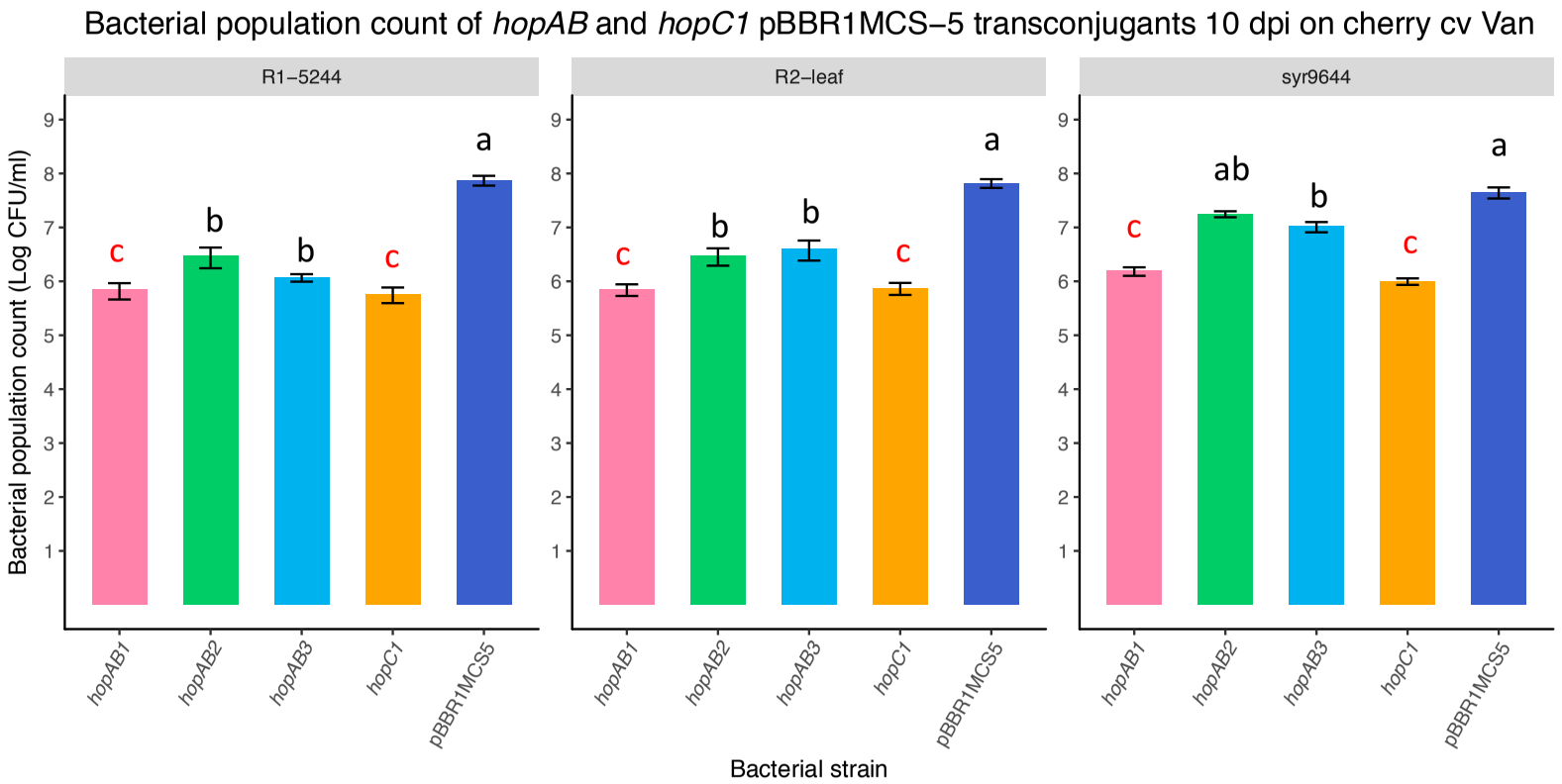


***Figure 5.6: Ten-day population counts of cherry pathogens (R1-5244, R2-leaf and syr9644) expressing different alleles of* hopAB.** *Values represent the mean of three replicates and error bars shows standard error above and below the mean. The empty vector control and* hopC1 *strains were included for comparison. A separate ANOVA was performed for each cherry pathogen (R1-5244, R2-leaf and syr9644) and the Tukey-HSD significance groups (p=0.05, confidence level: 0.95) for each genotype are presented above each bar. The expression of* hopC1 *and* hopAB1 *led to populations significantly different from the empty vector control in all three cherry pathogens, so the Tukey-HSD groups are indicated in red.*

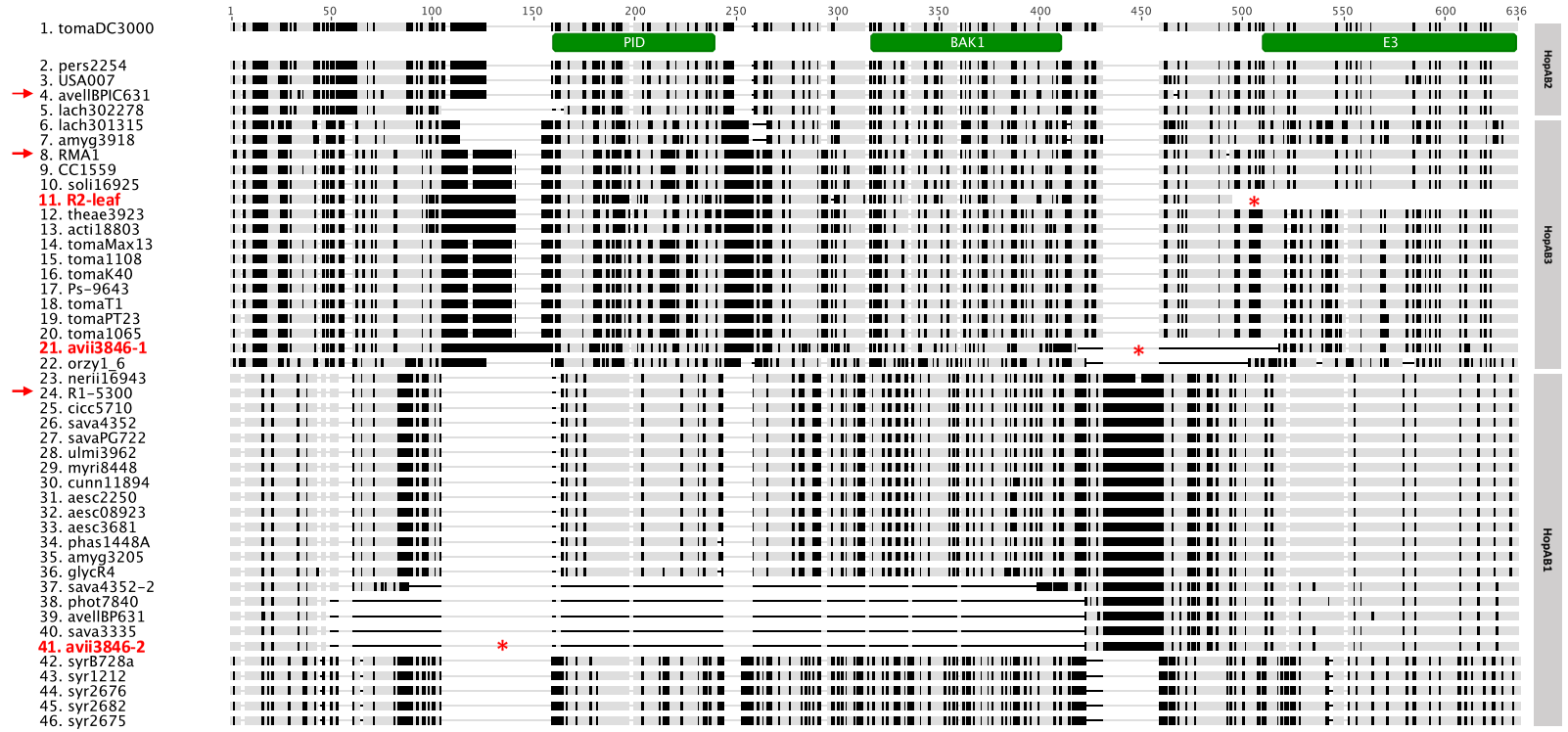
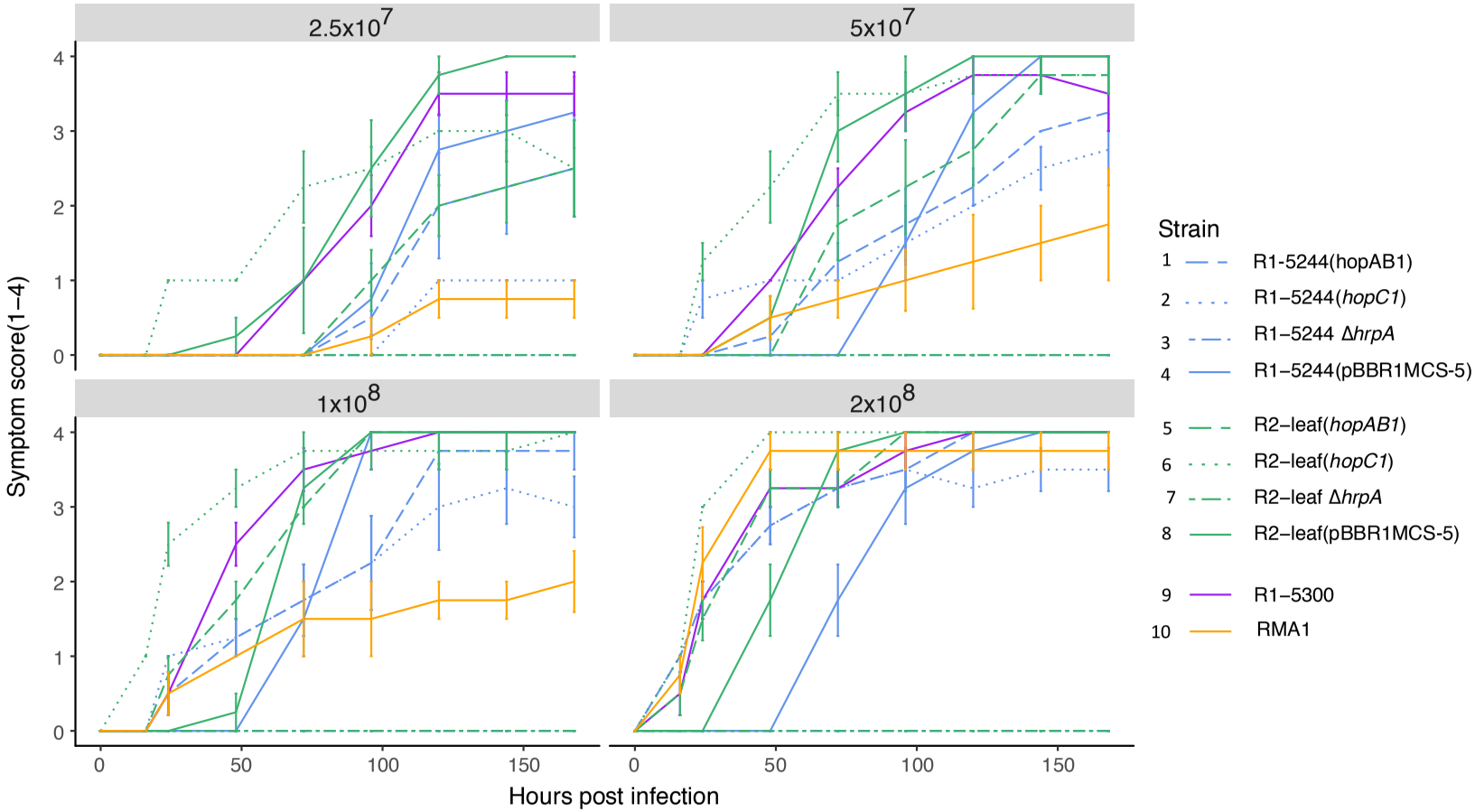
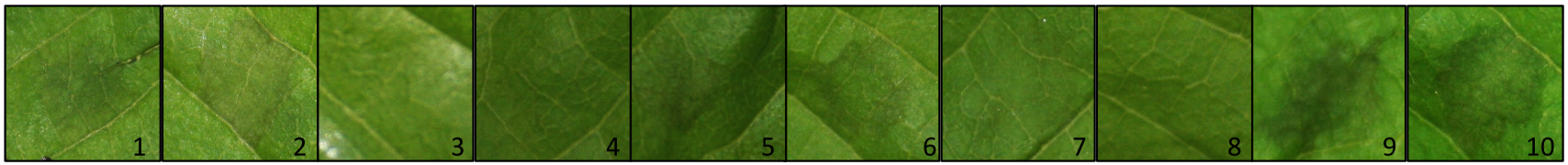


Figure 5.7: Protein alignment of members of the HopAB effector family. *Members of HopAB1, HopAB2 and HopAB3 are included. Putative domains identified with InterProScan – PID (Pto-interacting domain), BAK1 (BAK1-interacting domain) and E3 (E3-ubiquitin ligase domain) are annotated in the* P.s *pv.* tomato *DC3000 HopAB2 sequence. Cherry-infecting strains are highlighted in red. * shows the disrupted versions of HopAB found in cherry-infecting strains – the truncated version of HopAB3 in* Psm *R2-leaf, the large deletion in* P.s *pv.* avii *HopAB3 and the large deletion in* P.s *pv.* avii *HopAB1. Red arrows indicate the HopAB alleles that were cloned into pBBR1MCS-5 in this study.*

The detection of HopAB alleles and HopC1 by the cherry immune system may be inducing a hypersensitive response, which would prevent bacteria achieving wild-type population growth. To determine if a HR was occurring, symptom development was scored over time in cherry leaves. At the concentrations used for population counts ($2x10^6$ CFU/ml) there was clear differentiation in symptoms between the wild-type strains and mutants expressing *hopAB1*/*hopC1*, as only the wild-type strains produce symptoms (representative images are presented in Figure 5.4). The non-pathogens R1-5300 and RMA1 from which the *hopAB1* and *hopC1* genes were cloned also failed to cause symptoms. To see if a HR was visible at higher inoculum levels, symptoms were scored over time using progressively higher concentrations starting at $2.5x10^7$ CFU/ml and doubling the concentration until $2x10^8$ CFU/ml. Figure 5.8 shows that at these higher concentrations, all strains produce symptoms in the leaves. Lesions did not appear to vary in colour or size in pathogens or non-pathogens. When inoculated at the highest concentration ($2x10^8$ CFU/ml), a concentration 100x of that used in the population analysis, the timing of symptom development could be used to discriminate disease lesions from the putative HR. The non-pathogens and strains expressing *hopAB1*/*hopC1* all produced symptoms within 24 hours, whilst the pathogens generally did not produce symptoms until at least 48 h post-inoculation. The Δ*hrpA* deletion mutants were included as a negative control as they failed to induce symptoms at all as they cannot secrete effectors involved in either virulence (disease) or avirulence (HR).

Figure 5.8 (overleaf): Symptom development in cherry leaves of* P. syringae *strains expressing different effectors. *Values represent the mean of three replicates and error bars shows standard error above and below the mean. Lines are coloured based on* P. syringae *strain, whilst mutants of these strains have different line types. Representative symptoms 24 hpi are presented below for each strain labelled 1-10, at $2x10^8$ CFU/ml: 1: R1-5244(*hopAB1*), 2: R1-5244(*hopC1*), 3: R1-5244 Δ*hrpA*, 4: R1-5244 pBBR1MCS-5 empty vector, 5: R2-leaf(*hopAB1*), 6: R2-leaf(*hopC1*), 7: R2-leaf Δ*hrpA*, 8: R2-leaf pBBR1MCS-5 empty vector, 9: R1-5300, 10: RMA1.*


2.5x10^7
5x10^7
1x10^8
2x10^8
Symptom score(1–4)
Hours post infection
Strain
1 R1-5244(hopAB1)
2 R1–5244(hopC1)
3 R1–5244 ΔhrpA
4 R1–5244(pBBR1MCS-5)
5 R2–leaf(hopAB1)
6 R2–leaf(hopC1)
7 R2–leaf ΔhrpA
8 R2–leaf(pBBR1MCS-5)
9 R1–5300
10 RMA1

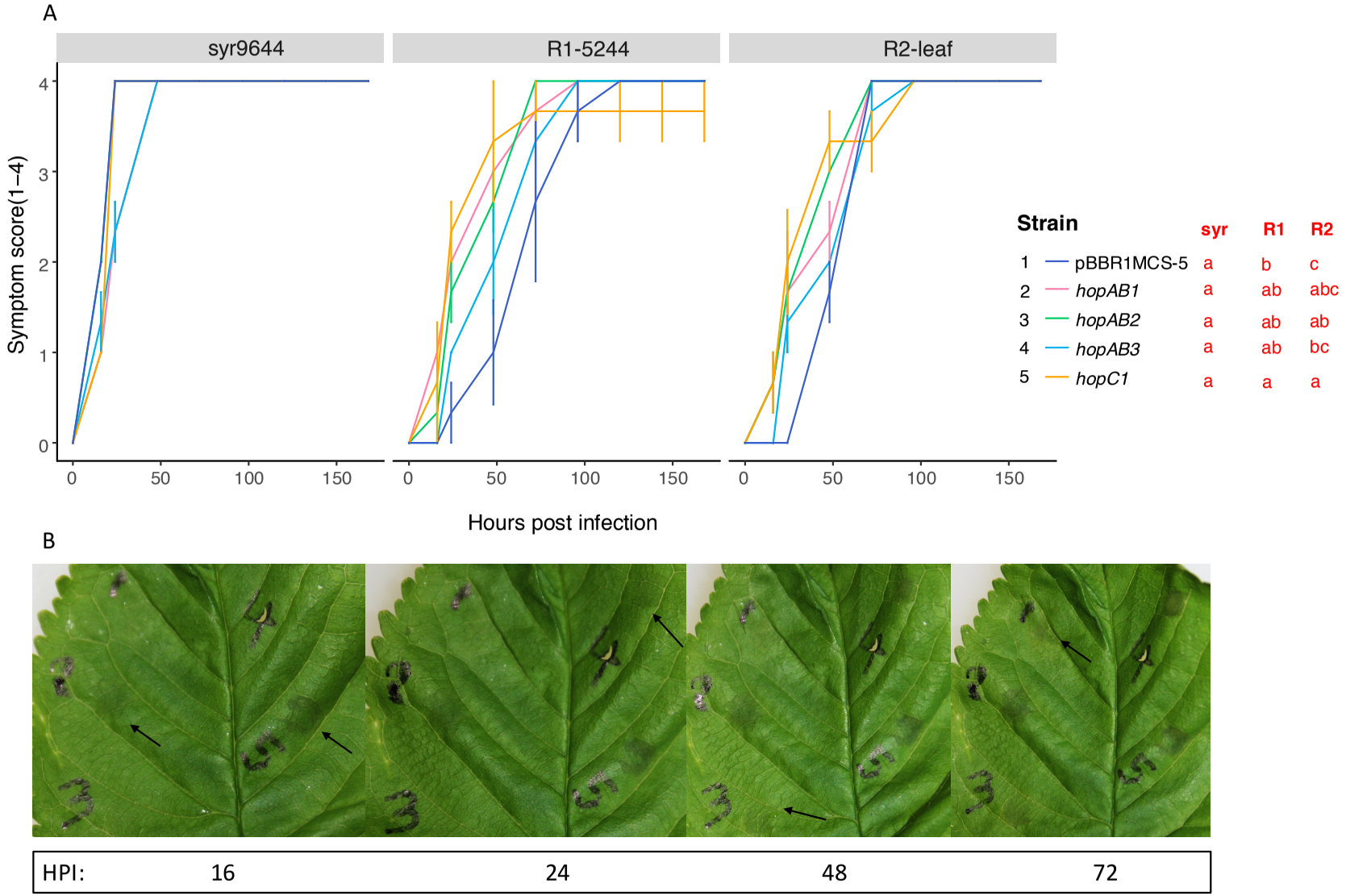

1
2
3
4
5
6
7
8
9
10

To statistically analyse these results, an ANOVA was performed comparing the slope of lesion development from 0-24 hours (as this is the time period in which a HR appears). There were significant differences between strains ($p<0.01$, df-9) and concentrations ($p<0.01$, df=3) as well as an interaction between strain and concentration ($p<0.01$, df=27), showing that the different strains behaved differently depending on inoculum concentration. At the lowest concentration ($2.5x10^7$ CFU/ml), only R2-leaf expressing *hopC1* produced lesions significantly faster than other strains, at $5x10^7$ CFU/ml both strains expressing *hopC1* produced lesions faster than other strains. At $1x10^8$ CFU/ml all non-pathogens and those pathogens expressing *avr* candidates showed faster lesion development from the controls but were not always significantly different. Whilst at the highest concentration $2x10^8$ CFU/ml all non-pathogens and pathogens expressing *avr* effectors were significantly different from the controls. This showed that, at higher inoculum concentrations the pathogens expressing *avr* effectors may be triggering a visible HR that can be differentiated from disease symptoms by timing of lesion onset. The HR induced by HopC1 was particularly strong, occurring rapidly at all concentrations.

This assay was repeated at the highest concentration for all three cherry pathogens expressing the three alleles of *hopAB* and *hopC1*, comparing symptom development to the empty vector control (Figure 5.9). In *Psm* R1 and R2 all strains expressing avirulence effectors produced symptoms before the control. In *Pss*, the empty vector control also produced symptoms within 24 hours and therefore could not be easily differentiated from those expressing avirulence effectors. An ANOVA of the slope of lesion development from 0-24 hours confirmed that there were differences between the three cherry pathogen clades ($p<0.01$, df=2) and a significant interaction between clade and mutant strain ($p<0.01$, df=12) showing that that the expression of *avr* effectors was affecting symptom development differently in each clade. R1-5244 transconjugants expressing *hopC1* and *hopAB1* were significantly different from the control in symptom development, but although those expressing *hopAB2* and *hopAB3* induced symptoms slightly quicker than the control, the difference in slope was not significant. R2-leaf transconjugants expressing *hopC1*, *hopAB1* and *hopAB2* were significantly different from the control in symptom development. Whilst in syr9644,

no strains expressing effectors induced symptom development at a significantly different rate to the control strain. Again, this showed that the HR induced by *hopC1* is rapid, whilst the HR induced by the different *hopAB* alleles was variable and sometimes could not be differentiated from the onset of disease. This meant that if members of *hopAB* are triggering a HR, it may be a weaker plant response to that induced by *hopC1* or maybe that the pathogenic strain is able to overcome this response and continue to suppress immunity, leading to a delayed symptom development that almost matches the empty vector control.

Figure 5.9: Symptom development of pathogens expressing* hopAB *alleles and* hopC1 *in cherry leaves. *A: Symptom development in cherry leaves for infections with R1-5244, R2-leaf, syr9644 and their pBBR1MCS-5 transconjugants. Values represent the mean of three replicates and error bars shows standard error above and below the mean. A separate ANOVA was performed on the 24-hour slope of symptom development for each cherry pathogen (R1-5244, R2-leaf and syr9644) and the Tukey-HSD significance groups (p=0.05, confidence level: 0.95) for each genotype are presented in red. B: Symptom development over time on a leaf inoculated with R1-5244 and its mutants. HPI: Hours post inoculation. The order of strains: 1: R1-5244 empty vector, 2: R1-5244 (hopAB1), 3: R1-5244 (hopAB2), 4: R1-5244 (hopAB3), 5: R1-5244 (hopC1). Arrows indicate the first appearance of symptoms associated with each strain.*


A
syr9644
R1-5244
R2-leaf
Symptom score(1–4)
Hours post infection
Strain
syr
R1
R2
1 pBBR1MCS-5 a b c
2 hopAB1 a ab abc
3 hopAB2 a ab ab
4 hopAB3 a ab bc
5 hopC1 a a a
B
HPI: 16 24 48 72

These effectors may therefore be triggering the HR when cherry is inoculated with non-pathogens such as R1-5300 and RMA1, and may contribute to the immune responses that limit the multiplication of these strains on cherry. These effectors were then deleted in the non-pathogens to see if this had an impact on their growth *in planta*. Without the HR-inducing effectors these strains may be able to reach higher population levels. *hopAB1* was deleted from both R1-5300 and another *Psm* R1 isolate, R1-9657, which is non-pathogenic, but originally isolated from cherry. The *hopAB2* gene was also deleted from *P.s* pv. *avellanae* BPIC631. However, when trying to delete the *hopC1* gene from RMA1, this strain was found to be recalcitrant to transformation using both conjugation and electroporation, meaning the gene could not be deleted. Figure 5.10 shows the bacterial population counts for all avirulence gene mutants and the wild-type controls. The R2-leaf pathogen which was inoculated in the same experiment was included for comparison. Individual ANOVAs on the two *Psm* R1 *hopAB1* mutants showed that there was no difference between genotypes (mutant vs. wildtype) (R1-5300: $p=0.67$, $df=1$ and R1-9657: $p=0.12$, $df=1$). Whilst, the deletion of *hopAB2* in *P.s* pv. *avellanae* BPIC631 results in a small (almost significant) increase in growth ($p=0.055$, $df=1$), however the mutant strains were unable to reach population levels similar to the *Psm* R2 pathogen (Figure 5.10C).

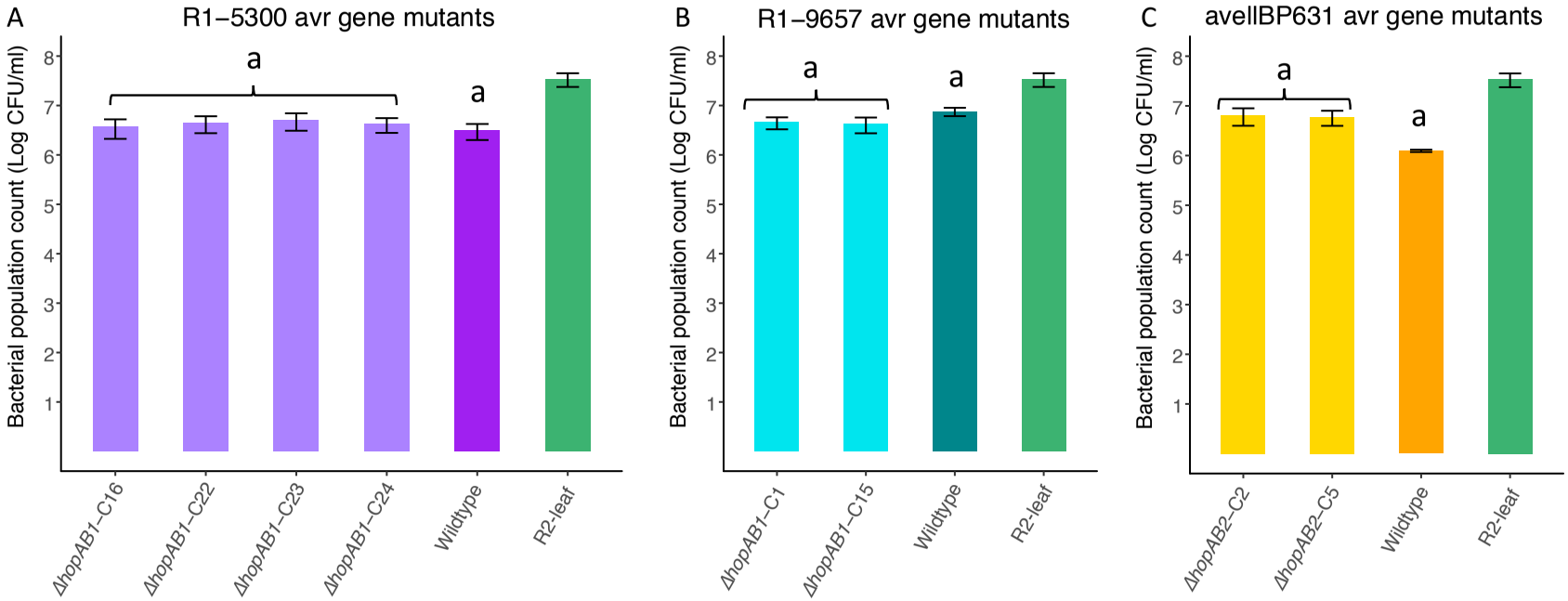


Figure 5.10: Ten-day population counts of avirulence gene deletion mutants and wildtype strains in cherry leaves for the three non-pathogens of cherry (A: R1-5300, B: R1-9657 and C: avellBPIC631). *Values represent the mean of three replicates and error bars shows standard error above and below the mean. A separate ANOVA was performed for each strain R1-5300, R1-9657 and avellBPIC631 and their mutants (The R2-leaf pathogenic strain in included in the graph for visual comparison of population levels, but was not included in the ANOVA). The ANOVA compared genotype (WT vs. mutant), with replicate genotypes with the same genes deleted, being grouped together. Tukey-HSD significance groups (p=0.05, confidence level: 0.95) are presented above the bar for each genotype.*

Cherry pathogens belonging to *Psm* R2 possess a truncated version of the *hopAB3* gene. Figure 5.7 shows an alignment of HopAB proteins from various *P. syringae*. The E3 ubiquitin ligase domain (involved in ubiquitination of protein substrates) in the R2-leaf HopAB3 protein is absent due to this truncation. To determine if this effector is important for the virulence of *Psm* R2 on cherry, two deletion mutants were created. Figure 5.11 shows the population counts of two Δ*hopAB3* mutants compared to the wild-type strain. Strains with the deleted gene were able to exceed 10^7 CFU/ml and cause disease symptoms, similar to the wildtype. An ANOVA revealed no significant difference between genotype (wildtype vs. mutant) ($p=0.64$, $df=1$). The two mutants varied in final population levels with C13 reaching a lower final population than the other mutant and WT. The ANOVA revealed that there were significant differences ($p<0.01$, $df=2$) between leaves (the three replicate leaves each came from separate cherry trees) indicating that this reduced population growth of C13 may have been influenced by physiological differences in leaf material. Therefore, deletion of *hopAB3* had little or no impact on the ability of *Psm* R2 to cause disease in cherry leaves.

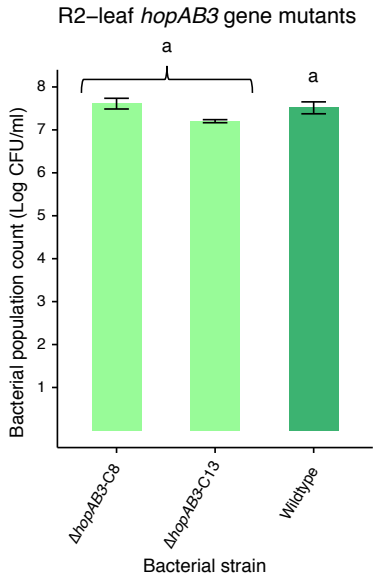
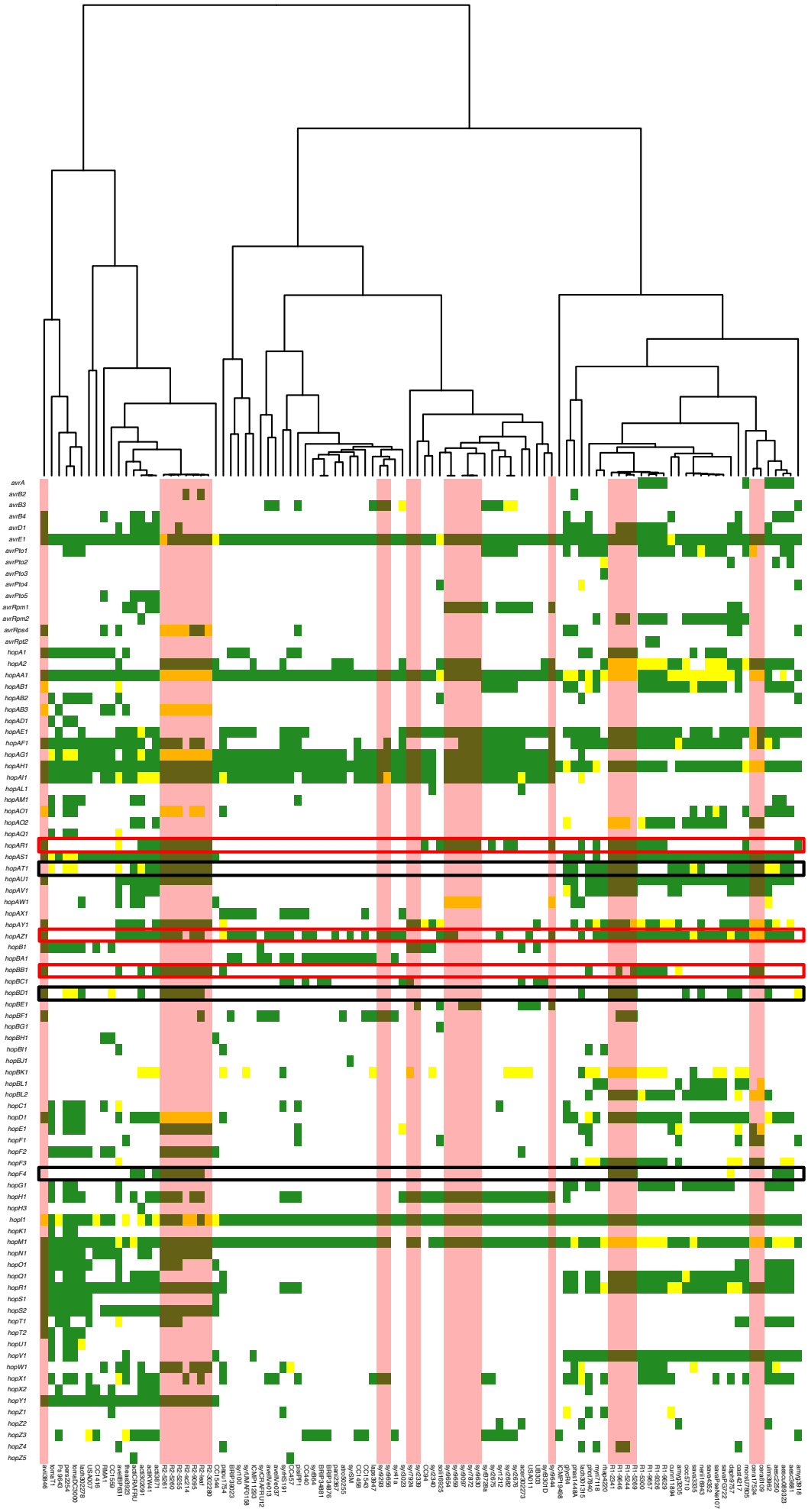


Figure 5.11: Ten-day population counts of* hopAB3 *gene deletion* Psm *R2-leaf mutants. *Values represent the mean of three replicates and error bars shows standard error above and below the mean. The wildtype strain was included for comparison. An ANOVA compared genotype (WT vs. mutant) and Tukey-HSD significance groups (p=0.05, confidence level: 0.95) are presented for WT vs. mutant.*

5.4.3 Identification of candidate virulence genes

In a similar approach to identifying *avr* genes, the presence and absence matrix of effector genes in cherry-infecting strains and the rest of the *P. syringae* complex was studied. Full-length effector genes that were present in members of each cherry-infecting *P. syringae* clade, but absent from closely-related non-pathogenic strains, were putatively important in virulence. Figure 5.12 shows a heatmap with candidate virulence genes highlighted. Bayestraits analysis in Chapter 4 (section 4.45) revealed several candidate effectors whose presence was statistically associated with cherry pathogenicity. The heatmap was also examined to determine if these effectors were present in multiple cherry-infecting pathogen clades, indicating that they may be vital for pathogenicity. For example, the *hopAR1* effector gene was statistically associated with pathogenicity ($p<0.05$) and was present in members of *Psm* R1, *Psm* R2, *Pss* and *P.s* pv. *avii* but not in all close out-groups. The effectors *hopAT1*, *hopBB1*, *hopBD1* and *hopF4* also scored highly in the Bayestraits analysis. They were absent from *Pss,* but present in most members of *Psm* R1 and *Psm* R2, despite these two clades being distantly-related, indicating that these effectors may be important for virulence. Another effector gene, hopAZ1 was also statistically associated with disease during Bayestraits analysis using the RealPHY phylogeny (Figure 4.18, $p=0.055$). Full-length homologues of this effector gene were present in *Psm* R1, *Psm* R2, *P.s* pv. *avii* and multiple clades of cherry pathogenic *P.s* pv. *syringae.* As it was found in nearly all the cherry-infecting *P. syringae* clades it was also a candidate virulence factor.

Figure 5.12 (overleaf): Heatmap of effector presence across* P. syringae *with candidate virulence genes highlighted. The heatmap was generated using the R gplots library*. The dendrogram clustered strains based on similarity of effector content. The green squares indicate presence of a full-length homologue of the effector gene whereas yellow squares indicate that the gene is disrupted or truncated in some way. Strains pathogenic to cherry are highlighted in red shade (in rows). Putative virulence genes are outlined. Cloned effectors are in red boxes (columns) whilst those that were not cloned are in black.*


avrA
avrB2
avrB3
avrB4
avrD1
avrE1
avrPto1
avrPto2
avrPto3
avrPto4
avrPto5
avrRpm1
avrRpm2
avrRps4
avrRpt2
hopA1
hopA2
hopAA1
hopAB1
hopAB2
hopAB3
hopAD1
hopAE1
hopAF1
hopAG1
hopAH1
hopAI1
hopAL1
hopAM1
hopAO1
hopAO2
hopAQ1
hopAR1
hopAS1
hopAT1
hopAU1
hopAV1
hopAW1
hopAX1
hopAY1
hopAZ1
hopB1
hopBA1
hopBB1
hopBC1
hopBD1
hopBE1
hopBF1
hopBG1
hopBH1
hopBI1
hopBJ1
hopBK1
hopBL1
hopBL2
hopC1
hopD1
hopE1
hopF1
hopF2
hopF3
hopF4
hopG1
hopH1
hopH3
hopI1
hopK1
hopM1
hopN1
hopO1
hopQ1
hopR1
hopS1
hopS2
hopT1
hopT2
hopU1
hopV1
hopW1
hopX1
hopX2
hopY1
hopZ1
hopZ2
hopZ3
hopZ4
hopZ5
amyg3918
aleic5891
aleic6893
aleic2250
urin3862
urin17504
mori17895
clajh2527
savaf382
savaf382
savaf382
nerii6843
savaf382
savaf382
savaf382

5.4.4 Single effector deletions

To determine if these effectors are important for virulence, *hopAR1*, *hopAZ* and *hopBB1*) were deleted in *Psm* R2 (R2-leaf) using the suicide vector pK18mobsacB. The *hopAR1* effector genes was also deleted in syr9644. PCR was used to validate that the vector inserts were correct and that the genes had been successfully deleted (Figures S5.3 and S5.4). Sanger sequencing across the deleted region was used to confirm the deletion in several mutants. The virulence of these mutants was then tested by measuring population growth in cherry leaves. Figure 5.13 shows that for all mutants there was no difference in final population counts compared with the wild-type strain. Mutants also produced similar symptoms to the wild-type. An ANOVA showed that there were significant differences between strains ($p<0.01$, $df=3$). However, the post-hoc Tukey-HSD test revealed no mutants were in different significance groups to the WT control. The deletion of *hopAR1* actually lead to slightly higher population growth. Whilst, the deletion of *hopAR1* in 9644 had no impact on pathogen virulence, with an ANOVA revealing no significant difference between strains ($p=0.89$, $df=1$).

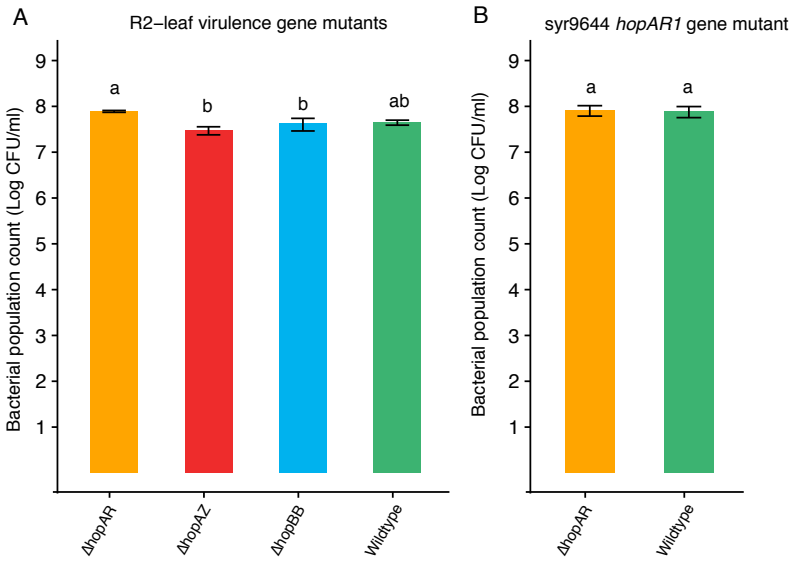


Figure 5.13: Ten-day population counts of* Psm *R2-leaf and syr9644 virulence gene deletion mutants compared to the wildtype strains. *Values represent the mean of three replicates and error bars shows standard error above and below the mean. The wildtype strains were included for comparison. A separate ANOVA was performed for each strain R2-leaf and syr9644 and their effector gene deletion mutants. Tukey-HSD significance groups (p=0.05, confidence level: 0.95) for the day 10 populations of each strain are presented above each bar.*

Clearly, the deletion of individual effectors did not significantly reduce virulence. This is often the case with cloning studies, as due to functional redundancy of effectors in *P. syringae*, single deletions usually do not affect virulence (Kvitko *et al.* 2009). Therefore, single, double and triple mutants were created of these three effectors and all mutants were screened on leaves. All deletion mutants grew to exceed 10^7 CFU/ml and cause lesion development. The ANOVA revealed that there were no significant differences between genotypes ($p=0.28$, $df=7$). As with some previous leaf experiments there were significant differences between leaves ($p=0.02$, $df=2$). Figure 5.14 shows that there were subtle differences in population growth of individual mutants. The differences observed were likely due to factors such as inoculum preparation or slight differences in leaf material.

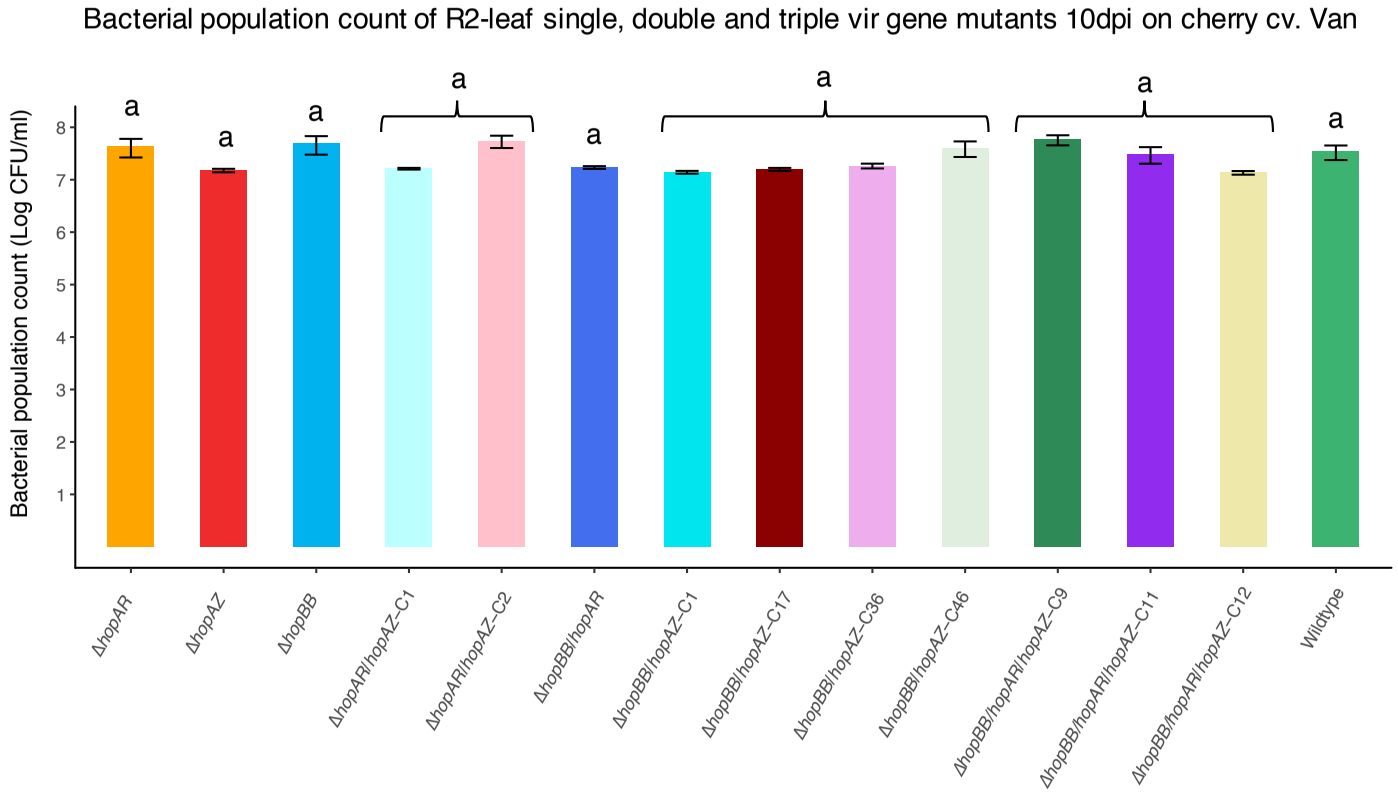


Figure 5.14: Ten-day population counts of all* Psm *R2-leaf single, double and triple virulence gene deletion mutants and the wildtype. *Values represent the mean of three replicates and error bars shows standard error above and below the mean. The ANOVA compared genotype (WT vs. mutant), with replicate genotypes with the same genes deleted, being grouped together. Tukey-HSD significance groups (p=0.05, confidence level: 0.95) are presented above the bar for each genotype.*

5.4.5 Genomic library screening

To take an unbiased approach to identify candidate avirulence genes involving host specificity for cherry, genomic libraries were created. To look for avirulence genes, genomic libraries of the *Aquilegia vulgaris* pathogen RMA1 and the non-pathogenic *Psm* R1 strain R1-5300 were created. Genomic DNA was partially digested and 20-30kb fragments were ligated into the cosmid vector pLAFR3. The cosmid library was then packaged and transfected into *E. coli* for storage. The presence of inserts was confirmed by restriction digests of the cosmid vectors (Figure S5.5). The RMA1 and R1-5300 libraries consisted of 960 and 480 clones respectively, each containing approximately 25kb of the donor strains genomic DNA. This number of clones corresponded to 3.88x coverage of the RMA1 genome and 1.89x coverage of the R1-5300 genome. Genomic coverage is important to ensure that there is a high probability that the whole genome is represented at least once in the library. With these clone numbers, the probabilities of successfully identifying specific genes were 97.9% and 84.6% for the RMA1 and R1-5300 libraries respectively.

The RMA1 and R1-5300 genomic libraries were conjugated into the closest related cherry pathogenic strains, R2-leaf and R1-5244, using tri-parental plate mating. A high-throughput plate mating method was used which allowed the conjugation of hundreds of cosmids into *P. syringae*. This involved mating on KBA and then transferring to KBA-Tet20-Nit100 to select for *P. syringae* containing the cosmid, as *E. coli* are counter-selected by nitrofurantoin. The resulting colonies were then transferred to KBA-Tet20. It was easy to differentiate R2-leaf transconjugant colonies from any contaminant *E. coli* that survived the nitrofurantoin selection as R2-leaf is fluorescent on KBA. However, R1-5244 is a non-fluorescent strain, making it difficult to distinguish from any *E. coli* contaminants. This led to a small number of false non-pathogenic reactions during virulence screening which were subsequently identified as *E. coli* cultures that had escaped selection.

The resulting R2-leaf(RMA1) and R1-5244(R1-5300) transconjugants each contained a small fragment of the non-pathogen's genome. These libraries were screened for virulence changes on cherry leaves. When inoculated at 10^6 CFU/ml both R2-leaf and

R1-5244 give clear disease symptoms after 7-10 dpi. If a gene or specific gene cluster on the cosmid vector is expressing gene products that trigger a resistance response in cherry leaves, such as the responses triggered by HopC1 and HopAB1, this would lead to reduced bacterial growth and a reduction in disease symptoms. As candidate *avr* effectors had been identified in RMA1 (HopC1) and R1-5300 (HopAB1), these could be identified through this approach and therefore act as positive controls. Both libraries were screened on cherry leaves for changes in disease symptoms, using the low inoculum concentration ($2x10^6$ CFU/ml) used in population counts. Each leaf was also inoculated with the pathogen containing an empty pLAFR3 vector and the donor non-pathogen to control for between leaf variation.

Unfortunately, the results of these screens were inconclusive, as all transconjugants produced disease symptoms *in planta*. It was surprising, particularly for the higher coverage RMA1 library, that HopC1-associated avirulence was not detected. The failure to detect avirulent clones may be because genome coverage was relatively low for both genomic libraries; less than the >5x coverage that is thought to be optimal for library creation (Boysen *et al.* 1997). Also, the leaves were screened based entirely on the presence and absence of disease symptom development and may not have picked up subtle differences in virulence due to avirulence factors.

5.5 Discussion

5.5.1 Identification of two avirulence effectors on cherry

In this study, candidate *P. syringae* effectors involved in host specificity for cherry were identified. First, candidate avirulence effectors involved in non-host resistance of cherry towards non-pathogenic strains of *P. syringae* were determined. During specialisation towards cherry, bacterial lineages may have lost or modified effector genes present in close out-groups that are detected by the cherry immune system. As in a similar study on bean (Baltrus et al. 2012), this question was approached using knowledge of the phylogenetic relationships between cherry-infecting and non-pathogen strains. This allowed informed predictions to be made about the likely genomic changes that have facilitated host jumps onto cherry within *P. syringae*. The

analysis was aided by the large number of genome sequences now available for *P. syringae*. The effector repertoires of many strains across the species complex could be compared, allowing hypotheses to be formed about the relationship between plant host and effector content.

Using a presence and absence matrix, putative avirulence genes for cloning were identified. This mainly focused on *Psm* R1, which is a highly homogenous clade containing both pathogenic and non-pathogenic isolates that differ in effector repertoires. The differences in effector content are likely to reflect host specificity. The effector genes *avrA1*, *avrPto1*, *hopAA1*, *hopAB1*, *hopAO2* and *hopG1* were present in non-pathogenic strains and close out-groups of *Psm* R1, but were absent from or pseudogenised in cherry *Psm* R1 strains. Several other candidate avirulence effectors were cloned from other clades. The *hopC1* gene was cloned as this effector had been previously shown to act as an avirulence factor in soybean (Baltrus *et al.* 2012). It was cloned from a strain isolated from *Aquilegia vulgaris*, which is basal to the clade containing *Psm* R2. Either this strain has gained this effector for pathogenicity or potentially it has been lost in the lineages leading to cherry pathogens. Effectors were also examined for allelic variations that could be important for host specificity, as allelic variation has been shown to affect the ability of the host to detect effectors (Stevens *et al.* 1998; Sohn *et al.* 2009). Two effectors (AvrRps4 and HopAW1) showed allelic variation in cherry pathogens compared to other strains, that could have led to pseudogenisation or differences in function.

Candidate avirulence effector genes were cloned into the broad-host range expression vector pBBR1MCS-5 and transferred into three pathogenic strains of *Psm* R1, *Psm* R2 and *Pss*. Population counts were used to initially test for differences in pathogen virulence. Any avirulence effectors would trigger immune responses in cherry, which could lead to reduced pathogen growth *in planta*. These tests were performed on detached leaves as it had previously been shown that leaves rapidly differentiated pathogens and non-pathogens (Chapter 3). It should be noted that statistical analysis sometimes revealed that there were significant differences between leaf replicates inoculated with the same bacterial strain. Although this was controlled for by placing

dormant trees in the glasshouse to ensure similarly aged, contamination-free leaf material, differences in physiology could have influence bacterial populations.

The expression of several effectors reduced pathogen growth *in planta* of one pathogenic clade but not of the others (Figure 5.5). In *Psm* R1, AvrA1 and AvrRps4 expression caused small but significant reductions in final population levels compared to the WT and empty vector controls, whilst in the other two pathogens these effectors did not affect virulence. Likewise, in *Psm* R2 the effector HopAW1 reduced the final population count but this effector had no effect in the other pathogens. These three effectors did not appear to influence symptom development and the reduction in final populations were small but significant. It would be interesting to repeat these experiments to see if these differences can be replicated, or if they are the product of experimental variation in bacterial inoculum and leaf material. If these effectors are triggering a host immune response it is likely to relatively weak as the pathogens were still able to grow and induce disease symptoms. The differences between cherry pathogens may be a product of their different effector repertoires. For example, in the case of AvrA1, this effector may be triggering immunity when expressed by *Psm* R1, but the other two pathogens are able to overcome this immune response e.g. by having effectors that suppress AvrA1-induced ETI. If AvrA1 expression does induce a weak immune response towards *Psm* R1, natural selection may have led to the loss of this effector from cherry infecting strains within this clade. The observation that HopAW1 reduced population growth in *Psm* R2 is also interesting as this effector is present in a closely related strain *P.s* pv. *avellanae* BPIC631 which is an out-group to the clade containing *Psm* R2. One hypothesis is that loss of this effector in lineages leading to *Psm* R2 was driven by its induction of immunity in cherry.

The expression of two effectors (HopC1 and HopAB1) consistently reduced bacterial population growth and inhibited symptom development in all three cherry pathogens. In *Psm* R2, final population levels were not significantly different from those recorded for a Δ*hrpA* mutant, indicating that these effectors are severely reducing virulence *in planta*. The ability of both effectors to induce a HR on cherry leaves was measured. At high inoculum concentrations the HR becomes macroscopic, as a large number of

plant cells undergo cell death (Mur *et al.* 2008). On cherry leaves, disease symptoms and the HR were similar in appearance but the putative HR appeared much more rapidly. In *Psm* R1 and R2, the expression of these two effectors induced symptom development significantly faster than the empty vector control, indicating that they may be inducing a HR. HR and disease could not be differentiated for *Pss*, as this strain appeared to be more necrotrophic on cherry leaves, causing disease symptoms within 24 hours.

The effector HopC1 has been shown to induce a HR and reduces bacterial population growth when expressed by pathogens in bean (Arnold *et al.* 2001b; Baltrus *et al.* 2012). A similar immune response may be occurring in cherry towards this effector. However, little is known about virulence and avirulence functions of this effector, apart from that it is a cysteine protease that resembles other effectors such as HopAR1 and HopN1 (Dowen *et al.* 2009).

The HopAB1 effector, cloned from R1-5300, is a homologue of the well-characterised HopAB2 (AvrPtoB) effector from *P.s* pv. *tomato* DC3000. HopAB2 was originally identified as it triggers immunity in Pto-expressing tomato lines (Kim *et al.* 2002). Diverse HopAB homologues, including HopAB1, from various *P. syringae* clades have also been shown to trigger immunity, indicating that detection of members of this effector family is associated with both host and non-host resistance in tomato (Lin *et al.* 2006). This effector contains five known domains and interacts with several immune proteins such as the receptor-like kinases BAK1 and FLS2, which are involved in PAMP-induced signalling during PTI. At the C-terminus there is an E3 ubiquitin ligase, which has been shown to be important for virulence (Göhre *et al.* 2008; Cheng *et al.* 2011; Chien *et al.* 2013). The immune responses induced by HopAB2 have been studied. The interaction of HopAB2 with the serine-threonine kinase Pto triggers the activation of the NB-LRR R protein Prf, leading to ETI. HopAB2 is detected by another Pto-related kinase called Fen that also triggers Prf-associated ETI. However, when a functional E3 ubiquitin ligase domain is present, Fen is degraded (Rosebrock *et al.* 2007). Studies on the interaction of *P.s* pv. *tomato* DC3000 and tomato, have shown that another effector, AvrPto, interacts with Pto and induces similar resistance

mechanisms as activated by HopAB2 (Lin & Martin 2005). However, the expression of AvrPto1 did not reduce population growth when expressed by cherry pathogens. This was intriguing as it suggests that the cherry immune response towards HopAB1 may not involve a homologue of Pto, but rather another factor is detecting HopAB1 and inducing ETI.

5.5.2 Alleles of HopAB vary in ability to trigger immune responses in cherry

Alleles of HopAB are highly diverse, with large scale mutational events occurring quite commonly in this effector family (Lindeberg *et al.* 2005). These events may be due to the selective pressures exerted by host immunes responses. To see if other members of this family trigger ETI in cherry, two additional alleles of HopAB were cloned; HopAB2 from *P.s* pv. *avellanae* BPIC631 and HopAB3 from RMA1. The expression of both HopAB2 and HopAB3 reduced bacterial populations *in planta* and prevented disease symptoms in *Psm* R1 and *Psm* R2. However, the reduction in growth compared to the empty vector control was less marked and growth of HopAB2-expressing *Pss* was not significantly different from the control strain. These two effectors were found to induce a HR, but at a slower rate than HopAB1 and HopC1, when expressed by *Psm* R1 and R2. This suggested that there is variation in the strength of cherry immune responses towards different members of the HopAB family, probably due to allelic variation within the protein domains involved in recognition.

Next, avirulence factors were deleted from the parental non-pathogenic strains to see if this had any effect on fitness in cherry. It was hypothesised that the deletion of these avirulence effectors in these strains could allow them to reach higher population levels *in planta*, as these effectors would no longer induce ETI. *hopAB1* was deleted from the plum isolate R1-5300 and another non-pathogen R1-9657, which was originally isolated from a cherry leaf wash. *hopAB2* was deleted in *P.s* pv. *avellanae* BPIC631. Attempts to delete *hopC1* from RMA1 were unsuccessful due to this strain's intolerance to transformation. The two *Psm* R1 *hopAB1* deletion mutants showed no significant differences in growth compared with the wild-type strains. The *ΔhopAB2 P.s* pv. *avellanae* BPIC631 mutants did show small and almost significant increases in population counts compared to the wildtype. This increase in population

growth may be due to a lack of HopAB2-induced immunity. However, the deletion of this gene did not allow *P.s* pv. *avellanae* BPIC631 to reach a similar final population to the *Psm* R2 pathogen. Overall the deletion of these avirulence genes had little or no impact on population growth. This is likely to be because these strains express other avirulence factors triggering immune responses in cherry that are sufficient to limit population growth. These strains may also lack the required virulence factors possessed by cherry pathogens that allow high population levels to be achieved.

Two cherry pathogenic clades possessed alleles of HopAB. The wild cherry pathogen *P.s* pv. *avii* possessed a truncated version of *hopAB1* and the *hopAB3* gene with a large deletion affecting the start of the E3-ubiquitin ligase domain. *Psm* R2 possessed a truncated homologue of *hopAB3,* which completely lacks the E3 ubiquitin ligase domain. To determine if *Psm* R2 has retained this effector for virulence functions, a deletion mutant was generated. The deletion of *hopAB3* had no impact on *Psm* R2 population growth. This indicated that this effector may not play an important role in disease on cherry, potentially due to the lack of the E3 ubiquitin ligase domain (Janjusevic *et al.* 2006) or functional redundancy with other effectors. As *Psm* R2 possessed this effector it would be interesting to see if it is expressed and translocated during infection. Variation in accumulation of this effector protein has been reported, which could be a possible mechanism to prevent detection (Lin *et al.* 2006; Kunkeaw *et al.* 2010). If the effector protein accumulates in bacterial cells and is translocated into plant cells, does it induce host immunity like other members of this family? Allelic variation may allow the *Psm* R2 HopAB3 protein to evade host immune responses. Or perhaps, *Psm* R2 possesses another effector involved in suppressing HopAB3-associated ETI.

If the *Psm* R2 HopAB3 effector fails to induce host immunity it would imply that convergent mechanisms to evade ETI have occurred during specialisation for cherry. *Psm* R1 has lost the hopAB1 effector gene completely, whilst *Psm* R2 could have modified the hopAB3 gene to evade recognition of the protein.

Studies of the HopAB family have shown it be an important effector in various diseases and part of a redundant effector group (REG) with AvrPto, required for full pathogenicity of the model strain *P.s* pv. *tomato* DC3000 Δ*hopQ1-1* on *N. benthamiana* (Jackson *et al.* 1999; De Torres *et al.* 2006; Kvitko *et al.* 2009; Cunnac *et al.* 2011). Comparative analysis of effector complements across *P. syringae* showed that most crop-associated strains (Chapter 4) possessed at least one member of this REG, which is involved in early suppression of PTI. It is intriguing that cherry-infecting strains either lack this REG entirely or possess truncated versions of these effectors. The avirulence activity of HopAB members on cherry may have driven the loss of this effector family during specialisation. However, AvrPto1 did not trigger immune responses, so why do no cherry pathogens possess a full-length version of this effector? Maybe effector targets of HopAB/AvrPto are not present in cherry so these effectors have been lost. Or perhaps, cherry pathogens rely on other effectors with similar functions to suppress early PTI responses.

5.5.3 Deletion studies of candidate virulence genes failed to alter the virulence phenotype

Candidate virulence-associated effectors involved in host specificity for cherry were also identified. Three effectors were deleted from a pathogenic *Psm* R2 strain. The effector gene *hopAR1* is present in the majority of cherry-infecting *P. syringae* strains so it was hypothesised that it may be important for virulence. HopAR1 (AvrPphB) is a papain-like cysteine protease which has been shown to disrupt PTI in *Arabidopsis* by interacting with receptor-like kinases such as BIK1, PBL1 and PBL2 (Zhang et al. 2010). It is reported to be an important virulence factor in bean. This is because even when strains carrying this gene were passaged through resistant plants able to detect HopAR1 and trigger immunity, bacterial populations still carried the effector gene at a low frequency. This is thought to be due to the negative frequency-dependent fitness benefit it provides on plants lacking the corresponding *R* gene (Neale *et al.* 2016). The effector *hopAZ1* was also deleted as this effector is present in multiple cherry-infecting *P. syringae* clades. This effector supresses both PTI-induced callose deposition and ETI in *Nicotiana tabacum* (Matas *et al.* 2014). Finally, *hopBB1* was deleted. Although this effector is absent from *Pss* strains, it is present in all other

cherry-infecting clades and appears to have been gained multiple times as close out-groups often lacked the *hopBB1* gene (Figure 5.12). HopBB1 is a member of the HopF super-family of effectors, which contains chimeric proteins classified as the same family due to similarity in the N-terminus. The N-terminus contains putative myristoylation sites involved in the localisation of the effector to the host cell membrane (Lo *et al.* 2016). The rest of the protein however is divergent from other HopF members so may contain domains of unknown function. HopBB1 has been found primarily in the genomes of *P. syringae* that infect woody hosts, so may be important for disease on woody tissues (Lo *et al.* 2016).

Population growth of the three deletion mutants was measured in cherry leaves. None were found to be significantly different from the wild-type control. Double and triple mutants of these effectors also showed no significant reduction in ability to reach high population levels and cause disease symptoms. In some ways this result was not surprising as functional redundancy within the effector repertoire likely allowed the bacteria to continue to cause disease without these effectors (Wei *et al.* 2007). In particular, the deletion of HopBB1 may have had little impact on virulence due to functional redundancy with other members of the HopF family. *Psm* R2 putatively expresses both HopF2 and The use of detached leaves to determine bacterial virulence may also prevent the detection of woody-tissue specific effects. Tissue-specific effectors have not been documented for *P. syringae*, however variation in immune responses between plant tissues could mean that different effectors are more important for immune suppression in particular tissues (Ryan *et al.* 2011). Several effectors have been reported to be predominantly present in *P. syringae* infecting woody plant species (Matas *et al.* 2014; Lo *et al.* 2016; Nowell *et al.* 2016). Therefore, particular effectors could be more important for bacterial fitness and ability to cause disease on the woody-tissue of the plant.

5.5.4 Genomic library screening

To determine host specificity factors in an unbiased approach, genomic libraries were created of the two non-pathogenic strains R1-5300 and RMA1. In the past, genomic libraries have been used successfully to identify candidate avirulence effectors that

trigger non-host resistance (Kobayashi *et al.* 1989; Arnold *et al.* 2001a). The two libraries were conjugated into the closely related pathogenic strains *Psm* R1 and *Psm* R2 and screened for changes in symptom development on detached leaves that could be due to the expression of avirulence factors from the genomic library. Unfortunately, no clones experienced a change in virulence on cherry leaves. As the HopAB1 and HopC1 effectors from R1-5300 and RMA1 had already been identified to be avirulence factors that prevent symptom development, these were expected to be identified as positive controls.

There are several possible reasons why no avirulent clones were identified. Firstly, the genome coverage of both libraries was quite low (1.89x and 3.88x for R1-5300 and RMA1 respectively), this would have reduced the possibility of candidate avirulence genes being represented in the library. Other studies of *Pseudomonas* have used thousands of clones to achieve higher coverage and identify candidate genes (Staskawicz *et al.* 1987; Ma *et al.* 1988). With a greater number of clones, candidate genes may have been identified. Any future library preparations would therefore aim to generate a much higher number of clones to work with. This would require optimising the protocol of cosmid transfection into *E. coli* to improve transfection efficiency. Also, virulence screening involved visualisation of changes in symptom development on leaves. This method may not be sensitive enough to detect subtle differences in virulence that may have been evident from population counts. For example, the initial screen of candidate effectors showed that expression of AvrA1, AvrRps4 in *Psm* R1 and HopAW1 in *Psm* R2 significantly reduced population levels but did not affect symptom development. Avirulence factors that weakly induce immune responses would therefore be missed. Also, by utilising a pathogen to deliver putative avirulence factors it is difficult to study the role of single effectors in immune activation as the pathogen also secretes many other effectors which could be interfering with the host response (Guo *et al.* 2009).

5.5.5 Conclusions

In this study, bioinformatics was used to identify several candidate avirulence genes from non-pathogens that trigger the ETI response in cherry. The absence or

modification of these avirulence effectors in cherry pathogens confirmed that *P. syringae* has fine-tuned its effector repertoire during host specialisation to evade host immunity. There has been dynamic evolution of members of the HopAB family of effectors during specialisation for cherry. The *hopAB1* effector gene has been completely lost in *Psm* R1, *hopAB3* is truncated in *Psm* R2 and there are large-scale mutations of both *hopAB1* and *hopAB3* genes in *P.s* pv. *avii*. These changes may be the result of the strong selective pressures exerted on pathogens to evade host immune responses (Kunkeaw *et al.* 2010). Further work on alleles of HopAB possessed by cherry pathogens could reveal if they are inducing immunity or have mutated to evade recognition.

It was determined that ETI and the HR may operate in cherry during non-host resistance towards non-pathogenic strains. Similar effector-induced immune responses could occur during host resistance in resistant cherry cultivars. By utilising the knowledge of the effectors found in cherry-infecting *P. syringae,* effectors could be used to screen for resistance as has been done with oomycete pathogens of potato (Vleeshouwers & Oliver 2014). This would allow the identification of *R* genes within cherry germplasm whose products are involved in both broad and clade-specific resistance towards *P. syringae*. As *P. syringae* has converged onto cherry at least five times, with the different clades containing divergent effector repertoires, identifying broad-spectrum resistance towards effectors may be challenging. Comparative genomic analysis revealed several convergently acquired effectors that may be important for disease, but their deletion did not affect virulence on cherry leaves. By using a polymutant approach, as in Cunnac *et al.* (2011), the problems associated with functional redundancy could be overcome. This would allow the identification of core effectors that are important in virulence on cherry and thus potential targets for effector-informed breeding.

Chapter 5: Supplementary results

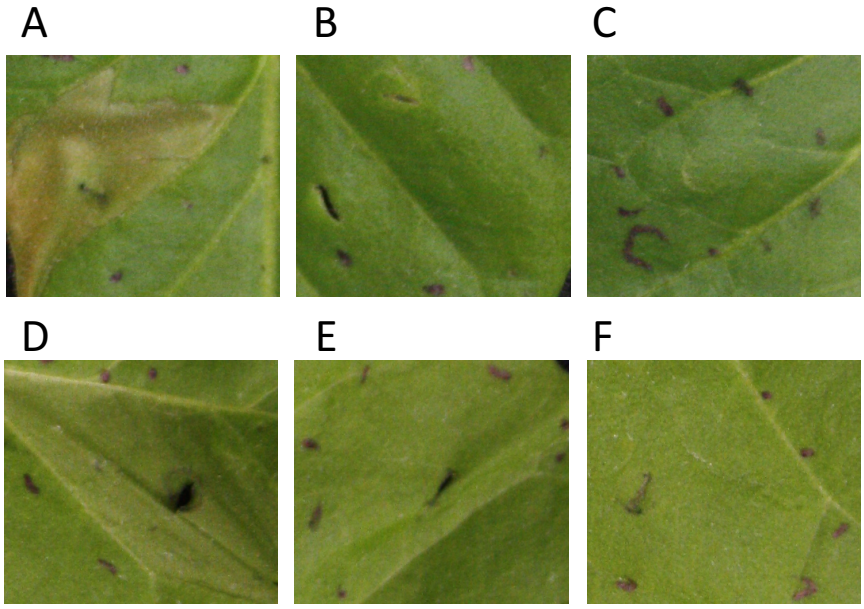
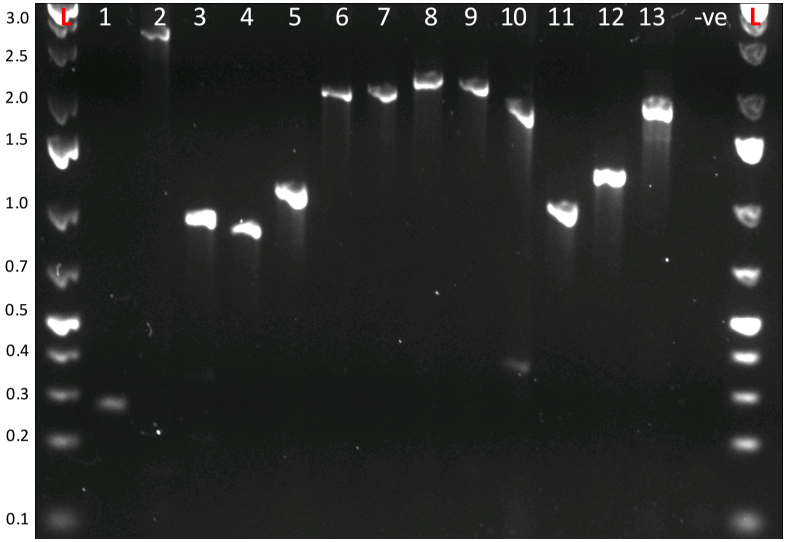
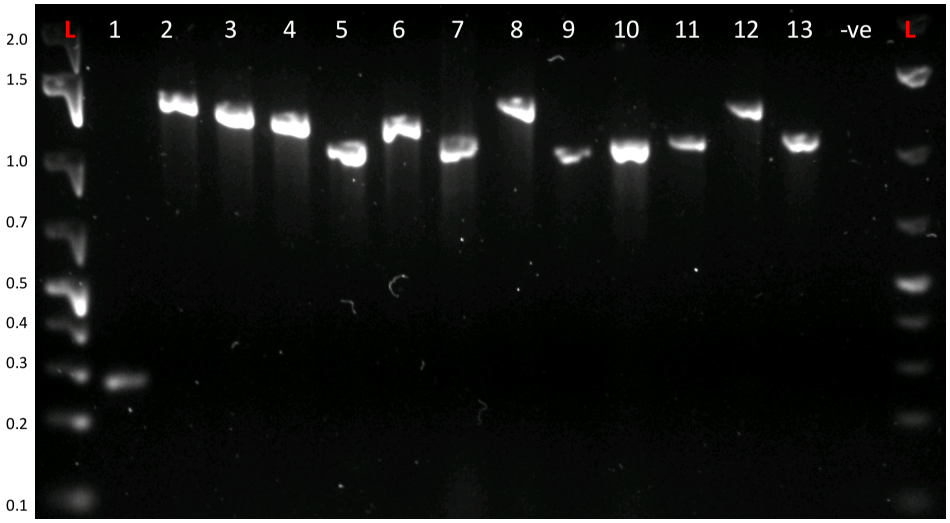


***Figure S5.1: Representative images of the tobacco hypersensitive response assay for R1-5244 and R2-leaf Δhrp**A **mutants**. Pictures were taken 24 hours after inoculation. A: 5244 WT, B: 5244 Δ*hrpA*, C: Negative control. D: R2-leaf wildtype, E: R2-leaf Δ*hrpA*, F: Negative control.*



Lane	Amplicon	Product size
1	pBBR1MCS-5	261
2	*avrA*	3921
3	*avrPto1*	931
4	*avrPto5*	860
5	*avrRps4*	1103
6	*hopAA1*	1938
7	*hopAB1*	2044
8	*hopAB2*	2184
9	*hopAB3*	2183
10	*hopAO1*	1862
11	*hopAW1*	982
12	*hopC1*	1200
13	*hopG1*	1919

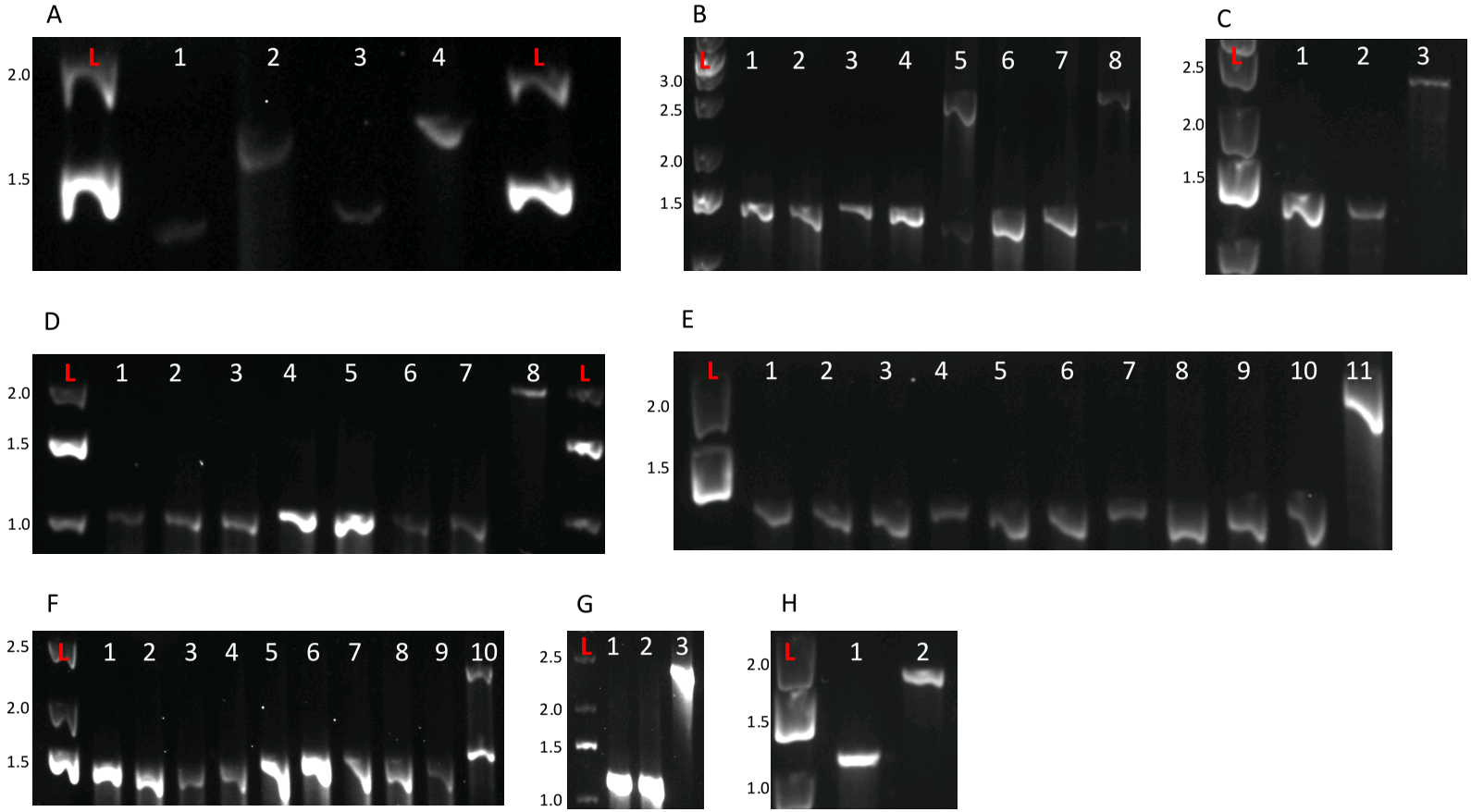
Figure S5.2: PCR products of an M13-PCR to check the insert size of all pBBR1MCS-5 vectors generated in this study. *The table reports putative product sizes. Ladder sizes are indicated per kb.*



Lane	Amplicon	Product size
1	pK18mobsacB	261
2	*hrpA*$_{R2\text{-}leaf}$	1348
3	*hrpA*$_{R1\text{-}5244}$	1286
4	*hopAZ1*$_{R2\text{-}leaf}$	1187
5	*hopAR1*$_{R2\text{-}leaf}$	1024
6	*hopBB1*$_{R2\text{-}leaf}$	1218
7	*hopAR1*$_{syr9644}$	1046
8	*hopAB1*$_{R1\text{-}5300}$	1279
9	*hopAB2*$_{avellBPIC631}$	1052
10	*hopAB3*$_{RMA1}$	1037
11	*hopAB3*$_{R2\text{-}leaf}$	1034
12	*hopC1*$_{RMA1}$	1161
13	*avrPto1*$_{R1\text{-}5300}$	958

Figure S5.3: PCR products of an M13-PCR to check the insert size of all pK18mobsacB vectors generated in this study. *The table reports putative product sizes. Ladder sizes are indicated per kb.*

Figure S5.4 (overleaf): PCR products to validate all deletions performed in this study*. The PCR involved amplifying the region containing the gene deletion. The Wildtype (WT) strain was included for comparison. Ladder sizes are indicated per kb.* ***A:*** **hrpA *deletion*** *– 1:R1-5244* Δhrp*A, 2: R1-5244 WT, 3: R2-leaf* Δhrp*A, 4: R2-leaf WT.* ***B:*** **hopAB1 *deletion*** *– 1: R1-5300* Δhop*AB1 C16, 2: R1-5300* Δhop*AB1 C22, 3: R1-5300* Δhop*AB1 C23, 4: R1-5300* Δhop*AB1 C24, 5: R1-5300 WT, 6: R1-9657* Δhop*AB1 C1, 7: R1-9657* Δhop*AB1 C15, 8: R1-9657 WT.* ***C:*** **hopAB2 *deletion*** *– 1: avellBPIC631* Δhop*AB2 C2, 2: avellBPIC631* Δhop*AB2 C5.* ***D:*** **hopAR1 *deletion*** *– 1: R2-leaf* Δhop*AR1, 2: R2-leaf* Δhop*AR1/*hop*AZ1 C1, 3: R2-leaf* Δhop*AR1/*hop*AZ1 C2, 4: R2-leaf* Δhop*BB1/*hop*AR1, 5: R2-leaf* Δhop*BB1/*hop*AR1/* hop*AR1 C9, 6:* R2-leaf Δhop*BB1/*hop*AR1/ hopAR1 C11, 7: R2-leaf* Δhop*BB1/*hop*AR1/* hop*AR1 C12, 8: R2-leaf WT.* ***E:*** **hopAZ1 *deletion*** *– 1: R2-leaf* Δhop*AZ1, 2: R2-leaf* Δhop*AR1/*hop*AZ1 C1, 3: R2-leaf* Δhop*AR1/*hop*AZ1 C2, 4: R2-leaf* Δhop*BB1/*hop*AZ1 C1, 5: R2-leaf* Δhop*BB1/*hop*AZ1 C17, 6: R2-leaf* Δhop*BB1/*hop*AZ1 C36, 7: R2-leaf* Δhop*BB1/*hop*AZ1 C46, 8: R2-leaf* Δhop*BB1/*hop*AR1/ hopAR1 C9, 9: R2-leaf* Δhop*BB1/*hop*AR1/*hop*AZ1 C11, 10: R2-leaf* Δhop*BB1/*hop*AR1/* hop*AZ1 C12, 11: R2-leaf WT.* ***F:*** **hopBB1 *deletion*** *– 1: R2-leaf* Δhop*BB1, 2: R2-leaf* Δhop*BB1/*hop*AR1, 3: R2-leaf* Δhop*BB1/*hop*AZ1 C1, 4: R2-leaf* Δhop*BB1/*hop*AZ1 C17, 5: R2-leaf* Δhop*BB1/*hop*AZ1 C36, 6: R2-leaf* Δhop*BB1/*hop*AZ1 C46, 7: R2-leaf* Δhop*BB1/*hop*AR1/* hop*AZ1 C9, 8: R2-leaf* Δhop*BB1/*hop*AR1/*hop*AZ1 C11, 9: R2-leaf* Δhop*BB1/*hop*AR1/* hop*AR1 C12, 10: R2-leaf WT.* ***G:*** **hopAB3 *deletion*** *– 1: R2-leaf* Δhop*AB3 C8, 2: R2-leaf* Δhop*AB3 C13, 1: R2-leaf WT.* ***H: syr9644*** **hopAR1 *deletion*** *– 1: syr9644* Δhop*AR1, 2: syr9644 WT.*


A
L 1 2 3 4 L
2.0
1.5
B
L 1 2 3 4 5 6 7 8
3.0
2.5
2.0
1.5
C
L 1 2 3
2.5
2.0
1.5
D
L 1 2 3 4 5 6 7 8 L
2.0
1.5
1.0
E
L 1 2 3 4 5 6 7 8 9 10 11
2.0
1.5
F
L 1 2 3 4 5 6 7 8 9 10
2.5
2.0
1.5
G
L 1 2 3
2.5
2.0
1.5
1.0
H
L 1 2
2.0
1.5
1.0

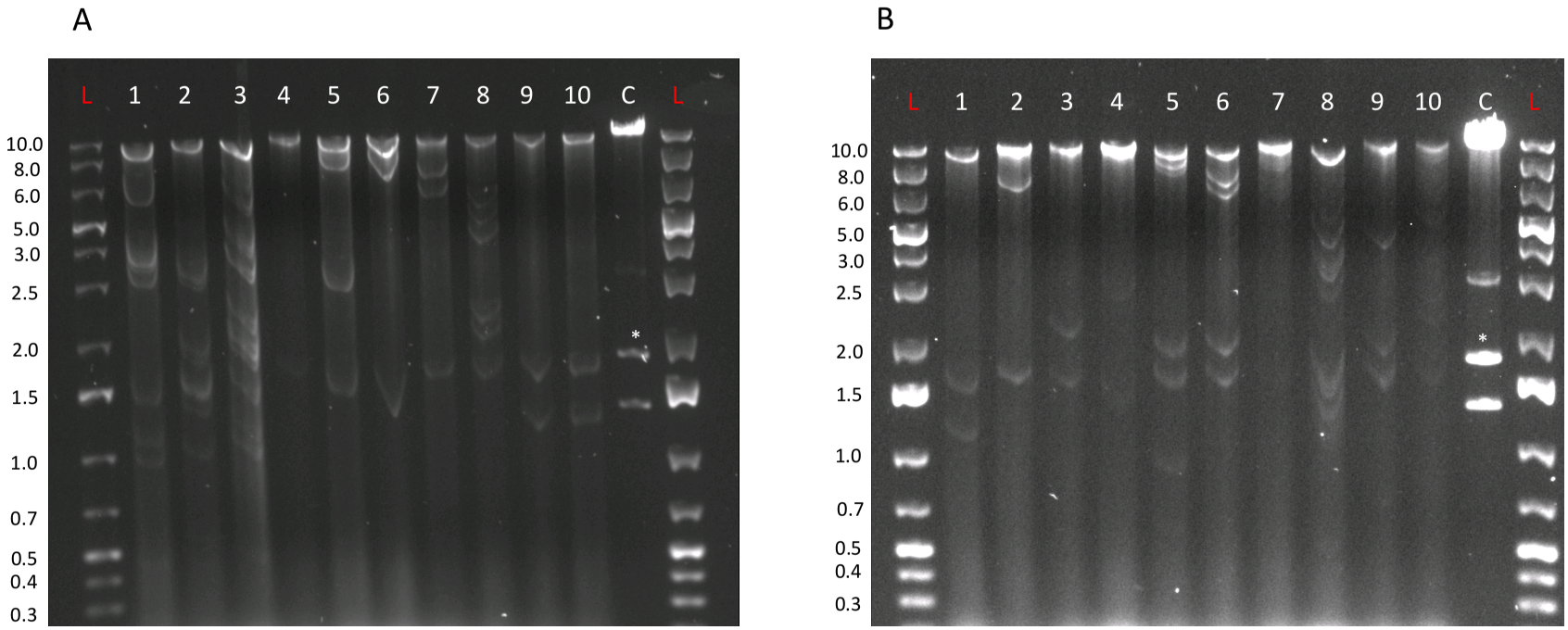


Figure S5.5: Restriction digest of 10 genomic library clones from the RMA1 (A) and R1-5300 (B) libraries*. Plasmids were extracted and digested with BamHI and BglII. 10 separate clones from each library were tested and pLAFR3 was included as a control. Gel images are labelled 1-10 for the separate clones and C for the pLAFR3. Ladder sizes are indicated per kb. *: The 1739bp band is due to BglII digestion of the plasmid backbone so is present in all clones. The presence of an insert in each genomic library clone is supported by additional bands of variable sizes compared to the pLAFR3 control, showing that the restriction enzymes are cutting within the insert.*

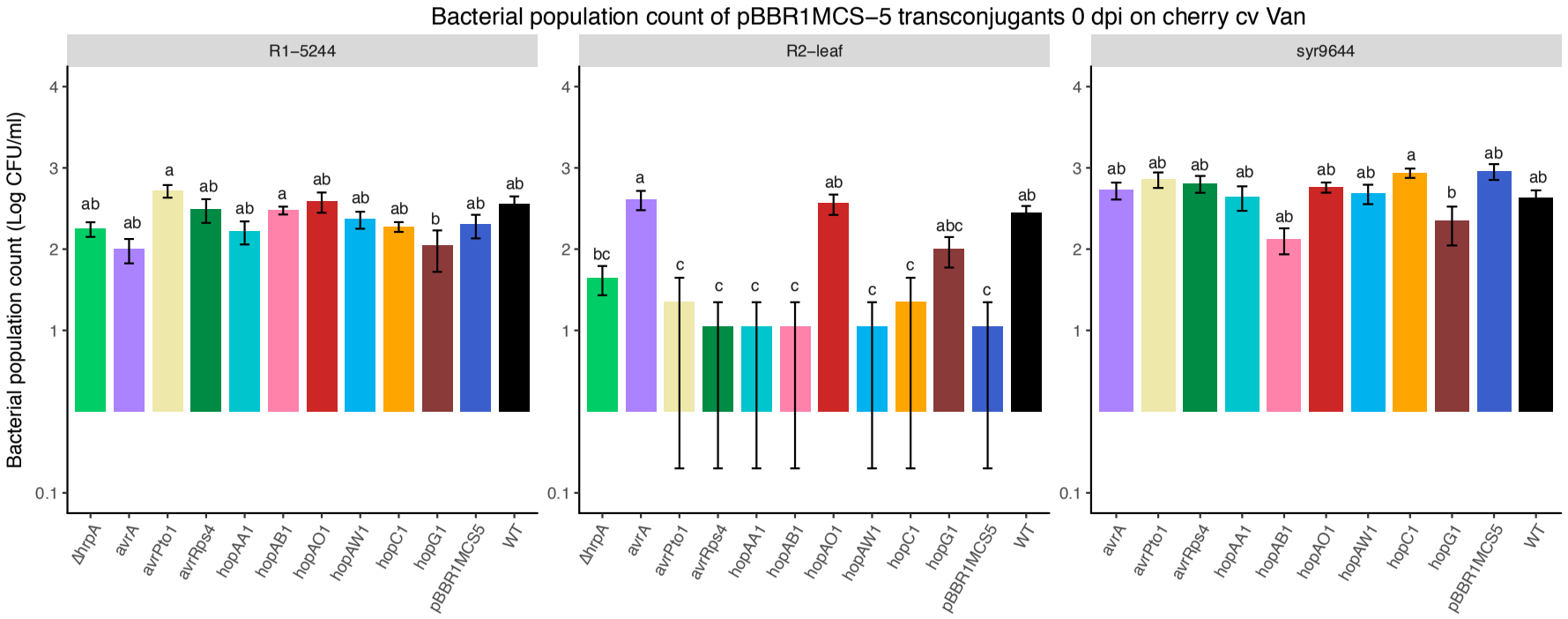


Figure S5.6: 0-day population counts of cherry pathogens (R1-5244, R2-leaf and syr9644) expressing candidate avirulence genes (for Figure 5.5)*. Values represent the mean of three replicates and error bars shows standard error above and below the mean. Controls included the wildtype strain, a strain containing the empty pBBR1MCS-5 vector and a* Δ*hrpA deletion mutant (for R1-5244 and R2-leaf). A separate ANOVA was performed for each cherry pathogen (R1-5244, R2-leaf and syr9644) and the Tukey-HSD significance groups (p=0.05, confidence level: 0.95) for each strain are presented above each bar.*

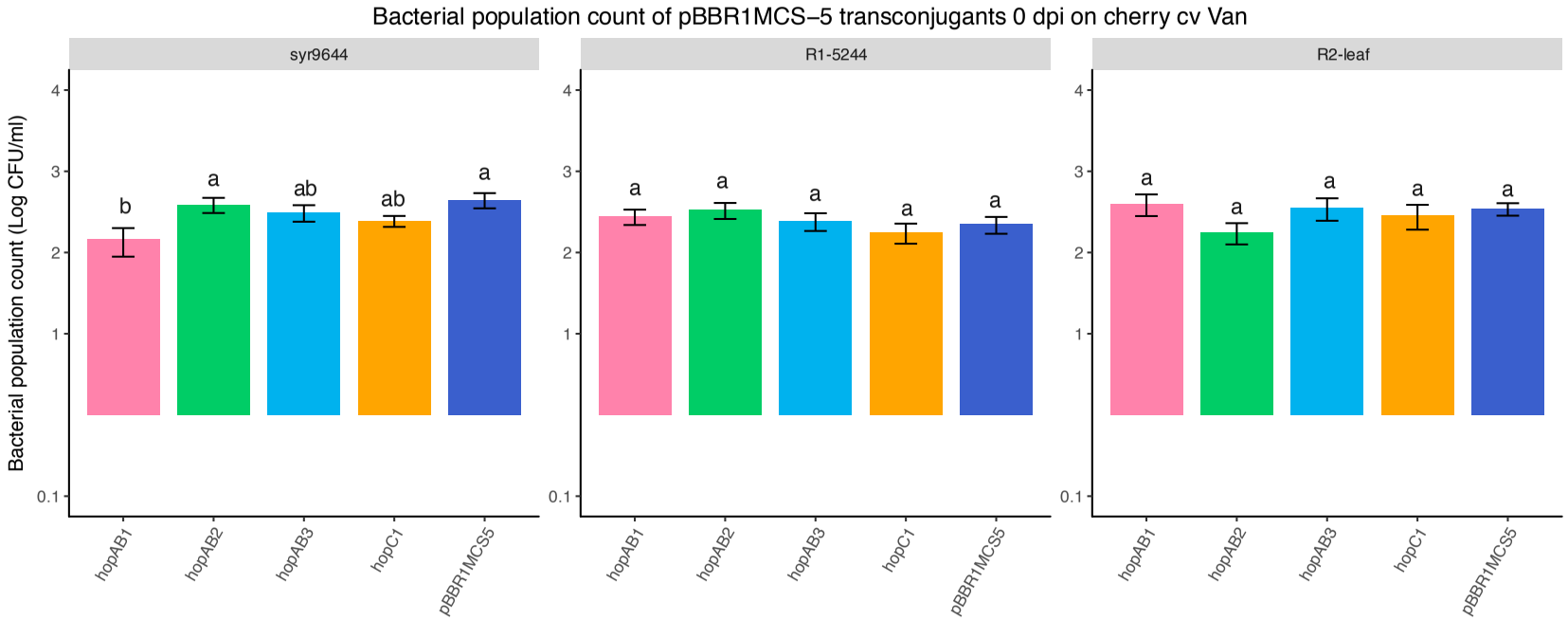


Figure S5.7: 0-day population counts of cherry pathogens (R1-5244, R2-leaf and syr9644) expressing different alleles of* hopAB *(for Figure 5.6). *Values represent the mean of three replicates and error bars shows standard error above and below the mean. The empty vector control and* hopC1 *strains were included for comparison. A separate ANOVA was performed for each cherry pathogen (R1-5244, R2-leaf and syr9644) and the Tukey-HSD significance groups (p=0.05, confidence level: 0.95) for each genotype are presented above each bar.*

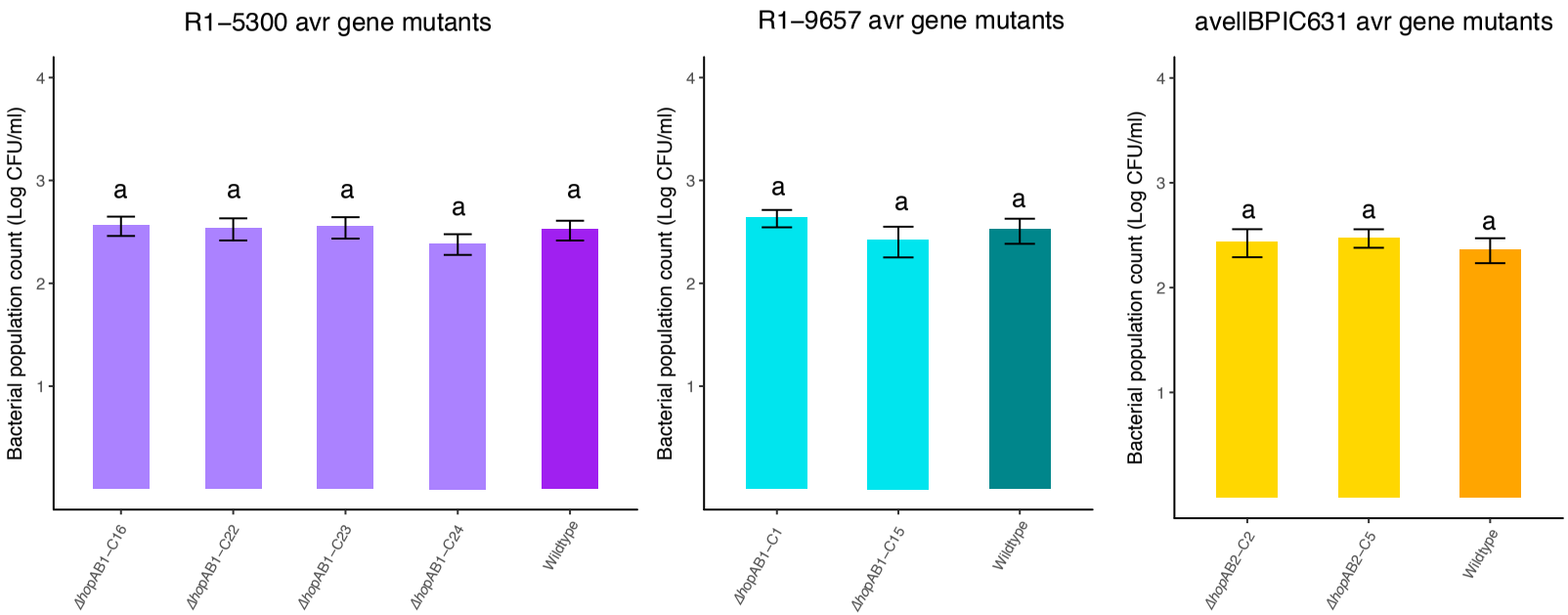


Figure S5.8: 0-day population counts of avirulence gene deletion mutants and wildtype strains in cherry leaves for the three non-pathogens of cherry (A: R1-5300, B: R1-9657 and C: avellBPIC631) (for Figure 5.10). *Values represent the mean of three replicates and error bars shows standard error above and below the mean. A separate ANOVA was performed for each strain R1-5300, R1-9657 and avellBPIC631 and their mutants. The ANOVA compared all strains inoculated in this experiment. Tukey-HSD significance groups (p=0.05, confidence level: 0.95) are presented above the bar for each strain.*

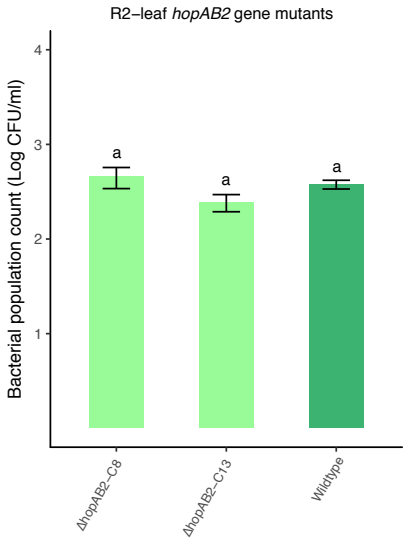


Figure S5.9: 0-day population counts of* hopAB3 *gene deletion* Psm *R2-leaf mutants (for Figure 5.11). *Values represent the mean of three replicates and error bars shows standard error above and below the mean. The wildtype strain was included for comparison. An ANOVA compared all strains inoculated. Tukey-HSD significance groups (p=0.05, confidence level: 0.95) are presented.*

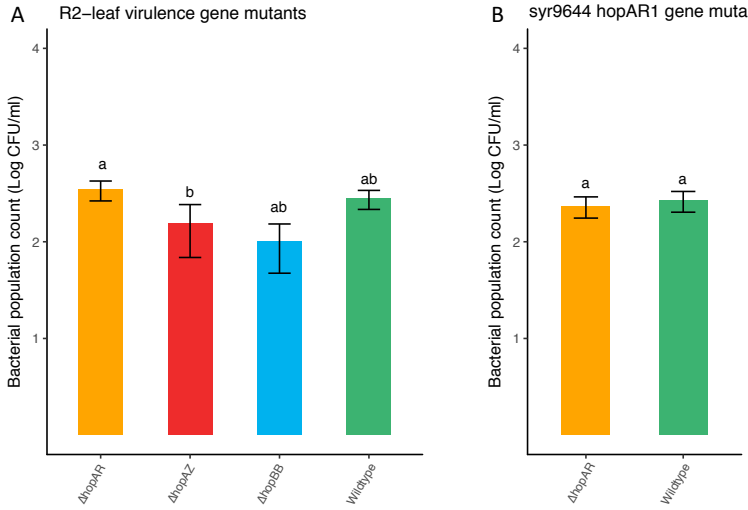


Figure S5.10: 0-day population counts of Psm R2-leaf and syr9644 virulence gene deletion mutants compared to the wildtype strains (for Figure 5.13). *Values represent the mean of three replicates and error bars shows standard error above and below the mean. The wildtype strains were included for comparison. A separate ANOVA was performed for each strain R2-leaf and syr9644 and their effector gene deletion mutants. Tukey-HSD significance groups (p=0.05, confidence level: 0.95) for the day 10 populations of each strain are presented above each bar.*

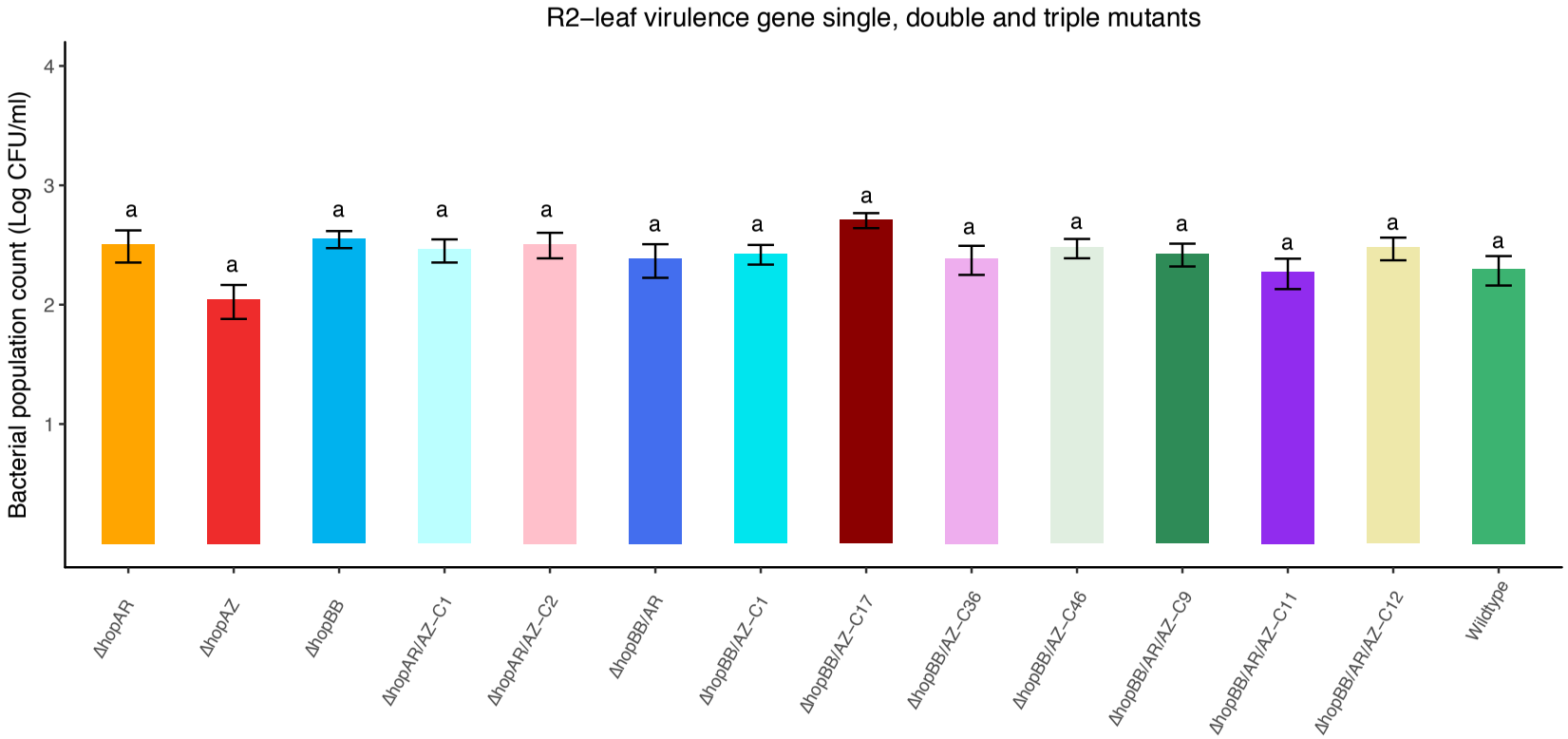


Figure S5.11: 0-day population counts of all* Psm *R2-leaf single, double and triple virulence gene deletion mutants and the wildtype (for Figure 5.14). *Values represent the mean of three replicates and error bars shows standard error above and below the mean. The ANOVA compared all strains inoculated in this experiment. Tukey-HSD significance groups (p=0.05, confidence level: 0.95) are presented above the bar for each genotype.*

Chapter 6: General discussion

This thesis presents an in-depth study of the cherry – *Pseudomonas syringae* pathosystem. *Pseudomonas syringae* is a widely-studied model organism, however, most studies have focused on strains infecting herbaceous plant species. In this study, the extensive knowledge obtained from a century of *P. syringae* research was applied to the economically important disease of stone fruits, bacterial canker. Firstly, the pathogenicity of various strains on cherry and plum was characterised in detail and variations in host cultivar susceptibility in field conditions were identified. Using this knowledge of pathogenicity, the genomes of various *P. syringae* isolated from cherry and plum were sequenced and comparative genomics used to elucidate candidate effectors important during interactions with cherry. Finally, gene cloning explored the role of these effectors *in planta* and identified several effectors that are putatively recognised by the cherry immune system. The combination of high throughput sequencing and well developed cloning systems therefore provided an insight into this complex biotic interaction.

Comparative genomics is becoming a common tool in the study of important microorganisms. However, phenotypic characterisation is vital for the creation of accurate hypotheses based on genomic comparisons. Recent genomics studies of *P. syringae* have utilised pathotype strains, so make assumptions about pathogenicity (Baltrus *et al.* 2012; Nowell *et al.* 2016). However, the strains in this study were uncharacterised, so their pathogenicity on *Prunus* was unknown. To make valid inferences about which genes products contribute to or prevent pathogenicity, it was deemed crucial to first characterise each strain's ability to cause disease. Therefore, before genome sequencing, a range of pathogenicity tests using *P. syringae* isolated from *Prunus* and some close out-groups were performed in Chapter 3. A pathogenicity assay on whole trees revealed that several non-pathogen strains isolated from other plant species were indeed unable to cause canker, conforming with the pathovar concept of *P. syringae*. All strains of *Psm* R1 isolated from plum and one isolated from cherry exhibited low virulence equivalent to non-pathogens, whilst in *Pss,* strains

isolated from plum and cherry were all capable of causing disease symptoms. This supported previous studies of *Pss* that found it had a broader host range than other *P. syringae* pathovars (Gilbert *et al.* 2009). Testing the pathogenicity of *Psm* R1 isolates in the field showed that a strain from plum was non-pathogenic on cherry but pathogenic to plum, indicating that a race structure may exist within this pathovar, with individual strains showing specificity towards different *Prunus* species.

A leaf-based population count assay was used to determine that non-pathogens (including *Psm* R1 from plum), failed to grow to high levels in cherry leaves and produced a putative hypersensitive response (HR). The HR has been reported to be the result of ETI (Thomma *et al.* 2011), and therefore the non-host resistance response of cherry towards these strains may have involved recognition of effectors. The timing of lesion appearance caused by *Pss* was found to not be significantly different from the onset of the HR caused by non-pathogens. The pathogenic strategy of *Pss* was therefore quite different from both *Psm* races in cherry leaves. The *Psm* races conformed to the hemi-biotrophic model of *P. syringae*-plant interactions and did not induce symptoms for several days, whilst *Pss* was more necrotrophic, inducing symptoms rapidly. The concentration used to measure symptom development in leaves was high (10^8 CFU/ml), and likely did not represent natural inoculum levels. It would therefore be interesting to apply these results to disease in the field and see how the symptom onset due to different cherry pathogens vary. This result is not only interesting from an ecological aspect, but by better understanding pathogenic strategies and population dynamics of different pathogenic *P. syringae* during infection, disease control measures could be optimised.

Once virulence levels were established for a set of strains, the susceptibility of several cherry and plum cultivars was determined under field conditions. Bacterial canker resistance has been a major objective of many cherry breeding programmes (Matthews 1979; Tobutt 1985). However, the breeding of resistant cultivars has been hindered due to the variability of different *P. syringae* clades that cause this disease (Kappel *et al.* 2012). The geographic prevalence of *Pss* compared to *Psm* (which is mainly reported in Europe, although known to occur elsewhere), has meant some

studies have only focused on *Pss* (Young 1987; Spotts *et al.* 2010; Farhadfar *et al.* 2016). However, by not testing the other known pathogenic clades, breeders may produce cultivars that are still susceptible to canker. For example, before the discovery of *Psm* R2 by Garrett (1978), breeding efforts produced cultivars exhibiting some resistance to *Psm* R1, which were later found to be susceptible to *Psm* R2 (Garrett 1986). Focusing on cherry, four cultivars were screened that were reported to exhibit either full susceptibility, resistance or race-specific resistance using strains from the three major clades that cause bacterial canker. The cultivar Merton Glory exhibited broad resistance to all strains, whilst Napoleon and Van were resistant to *Psm* R2. These three cultivars could be further studied to identify the genetic mechanisms underlying broad and race-specific resistances to canker. Currently, it is thought that broad resistance to all races of a pathogen is likely the product of various QTL, whilst race-specific resistance involves single major genes involved in specific R protein – avirulence effector interactions (Poland *et al.* 2009). It would be interesting to determine if resistance in Van and Napoleon to *Psm* R2 is mediated by the detection of a particular *Psm* R2 effector. Is the resistance of Merton Glory due to the ability to detect conserved effectors found in all three pathogenic clades? Maybe resistance of this cultivar is due to allelic variation of a whole set of genes, whose products are involved in immunity. Other genes, whose products are not directly involved immunity may contribute; for example, more vigorously growing varieties that exit dormancy earlier in spring have been shown to be more resistant to bacterial canker (Wilson 1939). Due to the less durable nature of race-specific resistance (Michelmore *et al.* 2013), future breeding efforts should focus on identifying the mechanisms of broad host resistance, such as that seen in Merton Glory.

Several different laboratory-based pathogenicity tests were also assessed. Rapid phenotyping is very important for resistance breeding, and therefore whole tree inoculations are not a viable, high throughput strategy. Detached leaves and cherry fruits provided a rapid means of assessing virulence and were able to differentiate pathogens and non-pathogens. However, when screening the four cherry cultivars, both assays generally failed to replicate the results observed in the field. For example, there was no significant difference between population growth of pathogens on

Merton Glory leaves compared to the other cherry cultivars, so any resistance of this cultivar was not picked up using this screen. This assay showed that all pathogens were capable of growing and causing symptoms on leaves. However, this rapid, *in vitro* assay likely fails to pick up the subtle differences in susceptibility of the different cultivars that were observed in the field. In field conditions, this plant-pathogen interaction lasts for many months. Perhaps, factors important not just for pathogenicity, but for the ability of bacterial populations to successfully colonise and persist through the season dictate the outcome of this interaction. Resistance mechanisms that reduce persistence in woody tissue, e.g. responses that block bacteria spreading to new tissues and acquiring nutrients, may prevent a pathogenic strain from causing severe disease.

Cut shoots were found to be better correlated with the field results, and therefore this assay could be implemented in large scale screening for bacterial canker resistance. As the field results could only be replicated using woody plant tissue, perhaps the resistance mechanisms operating towards *P. syringae* are tissue-specific. The woody cambial tissues that *P. syringae* populations occupy during winter months are in a dormant state and therefore the expression of resistance may differ substantially from leaf tissue. Cankers arise in spring when temperatures increase. Afterwards, the tissue becomes resistant to further spread of the disease during the growing season due to active periderm formation (Crosse 1966; Hinrichs-Berger 2004). Very little is known about the molecular basis of immune responses occurring in woody tissues. Does a hypersensitive response occur upon infection of resistant cultivars? Certainly, induced defences towards canker-causing bacteria and fungi, involve the formation of phenolic structures such as lignin and suberin, that act as barriers to prevent pathogen spread (Rioux 1996; Abe *et al.* 2007). Further study is therefore required to identify the molecular mechanisms involved in resistance to bacterial canker.

With an understanding of pathogenicity, genome sequencing of the *P. syringae* strains was performed. Genome sequencing confirmed that the three major bacterial canker pathogens (*Psm* R1, *Psm* R2 and *Pss*) were found in separate *P. syringae* phylogroups. *Psm* R1 and *Psm* R2 formed discrete monophyletic clades, whereas *Pss* strains were

interspersed with other strains in phylogroup 2. If separate *Pss* clades are assumed to have adapted independently, and the additional cherry-infecting pathovars *P.s* pv. *avii* and *P.s* pv. *cerasicola* are included, at least seven distinct clades have converged onto cherry. Phylogenetic analysis of the core genome revealed that strains of *Psm* R1 cluster based on pathogenicity for cherry. All plum isolates and a lower virulence cherry isolate (R1-9657), which grew to similar levels to the plum isolates in cherry leaves, formed a distinct clade from the cherry pathogens. This suggested that a race-structure based on host may be present in this clade, however further sampling would be needed to confirm this. This finding was significant as due to the close relationship between these strains, any differences in virulence factors likely reflected recent host adaptations. Phylogenetic analysis of *Pss* revealed that cherry pathogenicity may have evolved multiple times in phylogroup 2. However, pathogenicity data of the other strains within this clade would be needed to confirm this hypothesis. Due to confusion with nomenclature, many strains within phylogroup 2 have been named *Pss* on the basis of pathogenicity to lilac (Young 1991). Strains within this pathovar have been isolated from many different plant species. It would be interesting to see if other strains within phylogroup 2 can infect and cause disease on cherry. Perhaps lineages within phylogroup 2 exhibit a less specialised lifestyle and have broader host ranges than other *P. syringae* phylogroups?

Overall, differences in methods applied to produce phylogenetic trees meant that the placement the cherry-infecting strains within each phylogroup varied. This was not surprising as the different methods utilised different regions of the genome to assess strain similarity. The use of genes or particular genomic regions that have undergone recombination may lead to conflicting signals during phylogenetic construction, leading to different tree topologies. Which methods are most suitable for inferring phylogenetic relationships? Bacterial genomes are often divided into a core and flexible genome based on homology with other strains. Recent *P. syringae* studies have built core genome phylogenies based on single-copy orthologous genes (Nowell *et al.* 2016; Monteil *et al.* 2016). To build the core genome tree, a concatenated alignment of 902 core genes present in all strains was used. However, concatenation can lead to incorrect phylogenies when individual gene histories are discordant due to

recombination. A more informed approach would be to generate phylogenies for each gene and then combine these trees into a supertree, removing those trees that are incongruent (Moriarty Lemmon & Lemmon 2013; Daubin *et al.* 2002). The core genome tree contained 902 genes, which collectively represented 0.4Mb of the 6Mb *P. syringae* genome. Would increasing the amount of sequence data and including non-core genes produce improved phylogenies or just add noise due to conflicting phylogenetic patterns? Also, should recombination be ignored completely when creating phylogenies, as it is such an important process in shaping bacterial evolution? The RealPhy approach involved producing a 1.5Mb alignment based on mapping sequences to reference genomes. Within phylogroup 2, it clustered bean and cherry strains separately, whilst they were interspersed in the core genome tree. Host-specific polymorphisms were therefore being detected by using a larger alignment. Recombination of regions in this alignment occurring between lineages occupying the same ecological niche may have caused this host-specific clustering.

The effector analysis revealed that the different clades have gained several common non-core effectors that may have been important during specialisation on cherry. These included *hopAR1*, *hopBF1* and *hopBB1*. Homologues of the *hopAB*/*avrPto* effector families were found to have been lost or disrupted in multiple cherry-infecting clades. The matrix of effector presence and absence allowed inference of gene gain and loss across the phylogeny and the correlation of particular effector genes with cherry pathogenicity. However, despite several common effector genes, the different cherry-infecting clades appeared to possess distinct effector repertoires. They have therefore convergently evolved to cause disease using different effectors that may provide similar functions *in planta*. Overall, all cherry-infecting clades, apart from *Pss,* possessed large effector repertoires which they likely use to subvert immunity and cause disease. Meanwhile, *Pss* strains possessed very few known effector genes, but all possessed at least two toxin biosynthesis pathways. This finding supported the idea that *Pss* may be less specialised and exhibit a more necrotrophic lifestyle, possessing non-host specific toxins to achieve a broad host range. The negative correlation between effector genes and toxin biosynthesis genes suggested there may be a fitness cost to producing both. Without a large repertoire of effectors,

Pss may be able to occupy a broader host range as it would be less likely to induce ETI in new hosts. However, one could hypothesise that with a small set of core effectors, *Pss* would be unable to suppress immune responses and remain biotrophic for as long as other *P. syringae*. Necrosis-inducing toxins would therefore be required to kill plant cells before the immune response can limit the spread of *Pss*. However, this generalist pathogenic strategy may prevent *Pss* from occupying and fully exploiting a particular niche for as long as specialists.

How did these two different pathogenic strategies (effector and toxin-driven) evolve in *P. syringae*? The presence of phylogroup 2 strains with more effectors that lacked the complete set of genes in toxin biosynthesis clusters, such as *P.s* pv. *papulans*, suggested that reversion towards a more effector-driven strategy may occur.

Despite this concept of generalism, genomics did reveal that some specialisation may have occurred in *Pss*. For example, *Pss* strains pathogenic to bean all possessed *hopAB1,* which is known to be important for disease on bean (De Torres *et al.* 2006). Meanwhile, most cherry-infecting strains possessed *hopAR1*, *hopBE1* and *hopAW1* which were absent from most other strains in phylogroup 2. Perhaps these effectors are particularly important for disease on cherry. Several cherry-infecting *Pss* strains also possessed genes whose products are involved in the degradation of aromatic compounds commonly found in woody tissues, which are present in the other cherry- and wood-infecting pathovars. These pathways may provide fitness benefits to these strains on cherry, indicating that a degree of specialisation has occurred.

There was substantial evidence of horizontal transfer of effector genes between the cherry-infecting clades *Psm* R1, *Psm* R2, *P.s* pv. *avii* and *P.s* pv. *cerasicola*. PacBio long-read sequencing led to the assembly of complete genomes, revealing that many gained effectors were present on plasmids. Many of the gained genes that differentiate *Psm* R1 pathogens from non-pathogens in this clade, such as coronatine and certain effectors, have been acquired on plasmids. Therefore, acquisition of virulence genes via plasmids may have been important during the adaptation of *Psm* R1 towards cherry. Surprisingly, all sequenced *Prunus*-isolated *Pss* strains, apart from

syr9644, lacked plasmids. Recently gained effectors and lineage-specific genes have instead been acquired on genomic islands inserted into the chromosome e.g. within phage sequences. The lack of shared virulence genes between *Pss* and other cherry-infecting clades suggested that *Pss* has independently acquired pathogenicity, without gene acquisitions from other cherry-infecting clades. Although they occupy the same niche, *Pss* may be less capable of sharing genes with these other phylogroups. Or maybe due to differing pathogenic lifestyles, any shared effectors are not retained in *Pss* populations as they rely on a contrasting mechanism of virulence.

The genomic investigations revealed candidate genes involved in host specificity for cherry. Unfortunately, when single, double and triple deletion mutants of these candidate virulence effector genes were created for a strain of *Psm* R2, their deletion had no impact on virulence. Further studies may reveal if this result is due to functional redundancy with other effectors. Several genes that were absent or had been recently lost in cherry pathogens were also cloned to determine if they act as avirulence factors. The effectors HopC1 and members of the HopAB family were confirmed to trigger a hypersensitive response and reduce bacterial population growth when expressed by pathogenic strains. This suggested that an effector-triggered immune response is operating towards non-adapted *P. syringae* in cherry leaves. Interesting, members of HopAB differed in their ability to induce immune responses, likely due to allelic variation at the protein sites that interact with cherry immune proteins. The selection pressures exerted by the host immune response towards HopAB, may therefore have driven the convergent loss and disruption of hopAB homologues in cherry pathogens. The *hopAB1* gene has been completely lost in *Psm* R1, possibly due to movement of mobile elements, whilst other members of this effector family have been truncated in *P.s* pv. *avii* and *Psm* R2. HopAB forms a redundant effector group (REG) with AvrPto, which is required for full virulence of the model strain *P.s* pv. *tomato* DC3000 Δ*hopQ1-1* on *Nicotiana benthamiana* (Cunnac *et al.* 2011). This REG is involved in suppression of early immune responses triggered by the detection of conserved PAMPs. The absence/disruption of this REG in cherry pathogens was surprising as all other crop-associated strains in phylogroup 1 and 3 in this analysis possessed at least one full-length REG member. Perhaps the immune

targets of these effectors are absent from cherry, or other effectors are fulfilling the role of this REG?

Future directions

The detailed analysis of cherry pathogens presented in this thesis provides a baseline of knowledge to facilitate further study. Future research could focus on further characterising the differences in epidemiology between the different *Prunus*-infecting clades of *P. syringae.* Pathology and genomics predicted that *Pss* possesses a more necrotrophic, toxin-facilitated generalist lifestyle compared to *Psm*. Further work could therefore determine if this clade is less adapted to the canker niche than the highly specialised *Psm.* This could involve sampling from cankers to determine relative population levels throughout the season and comparing the frequencies at which *Pss* and *Psm* naturally cause canker. Also, as the different clades vary in virulence and likely induce contrasting host resistance mechanisms (particularly in cultivars such as Van that exhibits differential susceptibility), can these clades co-exist naturally in cankers, and if so is this coexistence beneficial or detrimental for the survival of different genotypes?

Further study would also be required to dissect the genetics of resistance to bacterial canker. The screening of mapping populations with parents of known susceptibility could be utilised to identify genetic markers, and eventually particular genes, associated with resistance. The potential tissue-specificity of resistance responses observed in this study is also an interesting avenue to explore. Future work could use gene expression assays of the different tissues inoculated with the pathogen to determine if resistance in woody tissues is correlated with the expression of particular genes associated with immunity. Varietal differences in resistance in woody tissues may also be linked to complex traits such as the length of dormancy, rootstock interactions and the timing of autumn leaf drop. Accurate phenotyping and genetic mapping could again be utilised to determine how and if these traits are linked.

Genomics revealed that the different *P. syringae* clades may utilise common and diverse mechanisms of pathogenicity. One area to explore further is the host

specificity of *Pss*. It would be interesting to dissect the host range of this clade and if strains show any specificity for their host of isolation. Evolution experiments, involving the passage of a strain on a particular plant species, could be used to see if *Pss* possesses a greater potential to expand its host range due to its smaller effector repertoire which may prevent avr recognition. In addition, the *Pss* cherry strains sequenced in this study were often closely related to those isolated from bean. If future pathogenicity tests reveal host specificity differences, the effector genes of these strains could be reciprocally cloned and expressed to identify virulence and avirulence factors. The different pathogenic lifestyles of *Pss* and *Psm* was reflected in a lack of horizontal gene transfer between these clades. Further sampling of isolates from orchards and genome sequencing could be used to examine the frequency of genetic transfer that has occurred between bacterial clades and determine if *Pss* is genetically isolated from *Psm* and other cherry pathogens.

The comparative genomics mainly focused on the identification of known type III effector genes. However, it is unlikely that effectors alone dictate the outcome of this disease, as bacterial lineages must adapt to survive on various plant tissues with contrasting physiology. Future bioinformatics could focus more broadly on the identification of candidate metabolic subsystems that are important for colonisation and persistence on cherry. An alternative, unbiased approach could be to utilise transposon mutant libraries to identify candidates genes (Manoharan *et al.* 2015). This involves the random integration of transposons into the bacterial genome, followed by screening of mutant libraries to identify mutants with altered virulence phenotypes.

In terms of identifying the effectors involved in host specificity for cherry further cloning and functional analysis would be required. Chapter 5 revealed that the HopAB effector is putatively recognised by the cherry immune system to trigger resistance. As the different alleles of this effector varied in their ability to induce immunity, it would be interesting to determine the mechanism of detection in cherry and whether the truncated version of HopAB3, possessed by pathogenic *Psm* R2 strains, can elicit an immune response. The identification of putative virulence factors was potentially

limited to due to functional redundancy. The production of polymutant strains deficient in all type II effectors would allow the impact of single effectors on virulence to be assessed (Cunnac *et al.* 2011). High-throughput methods of gene deletion and disruption have been recently optimised in other bacterial species, and therefore could provide a rapid means to generate effector polymutants (Nyerges et al. 2016).

Outlook

An overall aim of many plant pathogen genomics studies is to identify common, indispensable pathogen virulence factors, which could provide a basis for effector-informed breeding (Dangl *et al.* 2013; Burdon *et al.* 2016). The idea being that by understanding pathogen genetics, breeding strategies can be optimised to provide durable disease resistance. One aim is to use effectors as markers during breeding to identify sources of resistance, a method that has been successfully utilised in potato (Rietman *et al.* 2012). Also, by understanding the pathogens requirement for particular effectors and how functional redundancy is operating within effector repertoires, we can assess how durable such resistances are likely to be (Vleeshouwers & Oliver 2014) This study has revealed that at least seven *P. syringae* clades have converged to cause disease on cherry. By further increasing the sample size, further distinct pathogenic clades may have been identified. Broad resistance may occur towards common effectors shared by the different cherry-infecting clades. However, even though common effector genes, such as *hopAR1*, are present in many pathogenic strains and could be targets of resistance, not all pathogenic clades possessed these effectors. Hypothetically, an effector-recognition mediated resistance towards the various cherry-infecting lineages could provide broad-spectrum resistance, however it may fail to prevent disease associated with clades lacking these effectors. Sampling cherry-infecting strains from across *P. syringae* has revealed the versatility of these plant pathogens, which can use various distinct pathogenic strategies to cause disease on a common host. The way forward for breeding resistance to bacterial canker should therefore involve using a set of diverse bacterial strains during screening to capture the whole range of possible host responses. Resistance breeding could then be achieved by combining loci involved in both clade-specific and broad spectrum resistances.

In natural environments, pathogen genetic diversity is matched by a similar level of diversity in host populations which often keeps diseases in check. The increased use of genetically identical monocultures in agriculture in the last few decades has been a contributing factor to the emergence of various disease epidemics, due to a lack of host resistance in selected genotypes. The kiwifruit bacterial canker pandemic which has affected Europe, South America, South East Asia and New Zealand in recent years, is a prime example of how extensive use of only a few host genotypes can drive the rapid spread of highly virulent pathogen lineages and lead to crop devastation (McCann *et al.* 2016). By appreciating the evolutionary potential and diversity exhibited by plant pathogens in agriculture, growing systems can be designed that reduce the selective forces that drive the emergence of disease epidemics (McDonald & Linde 2002). In terms of bacterial canker, studies have shown that using mixed plantations with cherry interspersed with other tree species can reduce the incidence of this disease in timber producing plantations (Loewe *et al.* 2013). The wider ecological niches occupied by pathogen populations must also be considered, as environmental reservoirs may be the source of epidemic lineages (Monteil *et al.* 2013; Morris *et al.* 2017). With the continuous improvements in sequencing technologies and ability to perform large-scale genomic studies at low cost, the opportunities to address complex questions about bacteria have therefore never been greater.

Abe, K. *et al.*, 2007. Resistance sources to *Valsa* canker (*Valsa ceratosperma*) in a germplasm collection of diverse *Malus* species. *Plant Breeding*. 126(4), pp.449–453

Abramovitch, R.B. *et al.*, 2003. *Pseudomonas* type III effector AvrPtoB induces plant disease susceptibility by inhibition of host programmed cell death. *EMBO J.* 22(1), pp.60-69

Albacete, A. *et al.*, 2015. Unravelling rootstock×scion interactions to improve food security. *J Exp Bot*. 66(8), pp.2211–2226

Alfano, J.R. *et al.*, 2000. The *Pseudomonas syringae* Hrp pathogenicity island has a tripartite mosaic structure composed of a cluster of type III secretion genes bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity in plants. *P Natl Acad Sci USA*. 97(9), pp.4856–4861

Alfano, J.R., 2009. Roadmap for future research on plant pathogen effectors. *Mol Plant Pathol*. 10(6), pp.805–813

Almeida, N.F. *et al.*, 2009. A draft genome sequence of *Pseudomonas syringae* pv. *tomato* T1 reveals a type III effector repertoire significantly divergent from that of *Pseudomonas syringae* pv. *tomato* DC3000. *Mol Plant Microbe In*. 22(1), pp.52–62

Altschul, S.F. *et al.*, 1990. Basic local alignment search tool. *J Mol Biol*. 215(3), pp.403-410

APS, 1966. Merton cherries from England. *J Fruit Var Hortic Dig*, 20(46)

Aritua, V. *et al.*, 2015. Genome sequencing reveals a new lineage associated with lablab bean and genetic exchange between *Xanthomonas axonopodis* pv. *phaseoli* and *Xanthomonas fuscans* subsp. *fuscans*. *Front Microbiol*. 6(Oct), pp.1–18

Arnold, D.L. *et al.*, 2001a. Molecular characterization of *avrPphD*, a widely-distributed gene from *Pseudomonas syringae* pv.*phaseolicola* involved in non-host recognition by pea (*Pisum sativum*). *Physiol Mol Plant P.* 58(2), pp.55–62

Arnold, D.L. *et al.*, 2001b. Highly conserved sequences flank avirulence genes: Isolation of novel avirulence genes from *Pseudomonas syringae* pv. pisi. *Microbiol-UK*. 147(5), pp.1171–1182

Arnold, D.L. & Jackson, R.W., 2011. Bacterial genomes: evolution of pathogenicity. *Curr Opin Plant Biol*. 14(4), pp.385–91

Arrebola, E. *et al.*, 2009. Contribution of mangotoxin to the virulence and epiphytic fitness of *Pseudomonas syringae* pv. *syringae*. *Int Microbiol*. pp.87–95

Ashfield, T. *et al.*, 2003. Genetic and physical localization of the soybean Rpg1-b disease resistance gene reveals a complex locus containing several tightly linked

families of NBS-LRR genes. *Mol Plant Microbe In*. 16(9), pp.817–826

Ashfield, T. *et al.*, 2014. Evolutionary relationship of disease resistance genes in soybean and Arabidopsis specific for the *Pseudomonas syringae* effectors AvrB and AvrRpm1. *Plant Physiol*. 166(1), pp.235–251

Aziz, R.K. *et al.*, 2008. The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics*. 9, p.75

Van Baarlen, P. *et al.*, 2007. Molecular mechanisms of pathogenicity: How do pathogenic microorganisms develop cross-kingdom host jumps? *FEMS Microbiol Rev*. 31(3), pp.239–277

Baltrus, D.A. *et al.*, 2011. Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. *PLoS Pathogens*. 7(7), p.e1002132

Baltrus, D.A. *et al.*, 2012. The molecular basis of host specialization in bean pathovars of *Pseudomonas syringae*. *Mol Plant Microbe In*. 25(7), pp.877–888

Baltrus, D.A. *et al.*, 2014. Incongruence between multi-locus sequence analysis (MLSA) and whole-genome-based phylogenies: *Pseudomonas syringae* pathovar *pisi* as a cautionary tale. *Mol Plant Pathol*. 15(5), pp.1–5

Baltrus, D.A., 2016. Divorcing strain classification from species names. *Trends in Microbiol*. 24(6), pp.431–439

Baltrus, D.A., McCann, H.C & Guttman, D., 2017. Evolution, genomics and epidemiology of *Pseudomonas syringae*. *Mol Plant Pathol*. 18(1), pp.152-168

Bankevich, A. *et al.*, 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 19(5), pp.455–77

Barash, I. & Manulis-Sasson, S., 2007. Virulence mechanisms and host specificity of gall-forming *Pantoea agglomerans*. *Trends Microbiol*. 15(12), pp.538–45

Barker, D. & Pagel, M., 2005. Predicting functional gene links from phylogenetic-statistical analyses of whole genomes. *PLoS Comput Biol*, 1(1), p.e3

Bartoli, C. *et al.*, 2015. A framework to gauge the epidemic potential of plant pathogens in environmental reservoirs: The example of kiwifruit canker. *Mol Plant Pathol*. 16(2), pp.137–149

Bartoli, C., Roux, F. & Lamichhane, J.R., 2016. Molecular mechanisms underlying the emergence of bacterial pathogens: An ecological perspective. *Mol Plant Pathol*. 17(2), pp.303–310

Bender, C.L., Alarcón-chaidez, F. & Dennis, C., 1999. *Pseudomonas syringae* Phytotoxins : Mode of Action , Regulation , and Biosynthesis by *Pseudomonas syringae*. *Microbiol Mol Biol Rev*. 63(2), pp.266-292

Berge, O. *et al.*, 2014. A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex. *PLoS One*, 9(9), p.e105547

Berlin, K. *et al.*, 2015. Assembling large genomes with single-molecule sequencing and locality-sensitive hashing. *Nat Biotech*. 33(6), pp.623–630

Bertels, F. *et al.*, 2014. Automated reconstruction of whole-genome phylogenies from short-sequence reads. *Mol Biol Evol*. 31(5), pp.1077–1088

Binnewies, T.T. *et al.*, 2006. Ten years of bacterial genome sequencing: Comparative-genomics-based discoveries. *Funct Integr Genomics*. 6(3), pp.165–185

Block, A. *et al.*, 2010. The *Pseudomonas syringae* type III effector HopG1 targets mitochondria, alters plant development and suppresses plant innate immunity. *Cellular Microbiol*. 12(3), pp.318–330

Block, A. *et al.*, 2014. The *Pseudomonas syringae* type III effector HopD1 suppresses effector-triggered immunity, localizes to the endoplasmic reticulum, and targets the *Arabidopsis* transcription factor NTL9. *New phytol*. 201(4), pp.1358–1370

Børve, J. *et al.*, 2003. Rain protective covering of sweet cherry trees—effects of different covering methods on fruit quality and microclimate. *Hortechnnology*. 13(March), pp.143–148

Boysen, C., Simon, M.L. & Hood, L., 1997. Analysis of the 1.1 Mb Human alpha/beta T-cell receptor locus with bacterial artificial chromosome clones. *Genome Res*. 7(206), p.330–338

Buell, C.R. *et al.*, 2003. The complete genome sequence of the *Arabidopsis* and tomato pathogen *Pseudomonas syringae* pv. *tomato* DC3000. *P Natl Acad Sci USA*. 100(18), pp.10181–10186

Bulgarelli, D. *et al.*, 2013. Structure and functions of the bacterial microbiota of plants. *Ann Rev Plant Biol*. 64, pp.807–38

Bull, C.T. *et al.*, 2010. *Pseudomonas cannabina* pv. *cannabina* pv. nov., and *Pseudomonas cannabin*a pv. *alisalensis* (Cintas Koike and Bull, 2000) comb. nov., are members of the emended species *Pseudomonas cannabina* (ex Sutic & Dowson 1959) Gardan, Shafik, Belouin, Brosch, Grimont & Grimont 1999. *Syst Appl Microbiol*. 33(3), pp.105–115

Bultreys, A. & Kaluzna, M., 2010. Bacterial cankers caused by *Pseudomonas syringae*

on stone fruit species with special emphasis on the pathovars *syringae* and *morsprunorum* Race 1 and Race 2. *J Plant Pathol*. 92(1), pp.S1-21-S1.33

Buonaurio, R. *et al.*, 2015. The olive knot disease as a model to study the role of interspecies bacterial communities in plant disease. *Front Plant Sci.* 6(June), p.434

Burdon, J.J. *et al.*, 2016. Addressing the challenges of pathogen evolution on the world's arable crops. *Phytopathology*. 106(10), pp.1115-1127

Caballo-Ponce, E. *et al.*, 2016. WHOP, a genomic region associated with woody hosts in the *Pseudomonas syringae* complex contributes to the virulence and fitness of *Pseudomonas savastanoi* pv. *savastanoi* in olive plants. *Mol Plant Microbe In*. DOI: 10.1094/MPMI-11-16-0233-R

Cadoret, F., Soscia, C. & Voulhoux, R., 2014. Gene transfer: Transformation/electroporation. *In:* Filloux, A. & Ramos, J. (eds), *Pseudomonas Methods and Protocols.* New York: Springer Science, pp.11-15

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evo*. 17(4), pp.540–552

Cerboneschi, M. *et al.*, 2016. Indole-3-acetic acid in plant-pathogen interactions: A key molecule for *in planta* bacterial virulence and fitness. *Res Microbiol*. 167(9–10), pp.774–787

Chakraborty, S. & Newton, A.C., 2011. Climate change, plant diseases and food security: An overview. *Plant Pathol*. 60(1), pp.2–14

Chen, W.H. *et al.*, 2009. ColorTree: A batch customization tool for phylogenic trees. *BMC Res Notes*. 2(1), p.155

Cheng, W. *et al.*, 2011. Structural analysis of *Pseudomonas syringae* AvrPtoB bound to host BAK1 reveals two similar kinase-interacting domains in a type III effector. *Cell Host Microbe*. 10(6), pp.616–626

Chien, C.F. *et al.*, 2013. Nonhost resistance of tomato to the bean pathogen *Pseudomonas syringae* pv. *syringae* B728a is due to a defective E3 ubiquitin ligase domain in avrptobb728a. *Mol Plant Microbe In.* 26(4), pp.387–97

Clarke, C.R. *et al.*, 2013. Allelic variation in two distinct *Pseudomonas syringae* flagellin epitopes modulates the strength of plant immune responses but not bacterial motility. *New Phytol*. 200(3), pp.847–60.

Cohen, O. *et al.*, 2010. GLOOME: Gain loss mapping engine. *Bioinformatics*. 26(22), pp.2914–2915

Cook, D.E., Mesarich, C.H. & Thomma, B.P.H.J., 2014. Understanding plant immunity

as a surveillance system to detect invasion. *Annu Rev Phytopathol*. 53(May), pp.541-563

Corwin, J.A. *et al.*, 2016. The quantitative basis of the *Arabidopsis* innate immune system to endemic pathogens depends on pathogen genetics. *PLoS Genet*, 12(2), pp.1–29

Crosse, J.E., 1959. Bacterial canker of stone-fruits. IV. Investigation of a method for measuring the inoculum potential of cherry trees. *Ann Appl Bio.* 47(1), pp.306–317
Crosse, J.E., 1966. Epidemiological relations of the *Pseudomonad* pathogens of deciduous fruit trees. *Annu Rev Phytopathol*. 4(1), pp.291–310

Crosse, J.E. & Garrett, C.M.E., 1966. Bacterial canker of stone-fruits. VII. Infection experiments with *Pseudomonas morsprunorum* and *Ps. syringae*. *Ann Appl Biol*. 58(1), pp.31–41

Crosse, J.E. & Garrett, C.M.E., 1970. Pathogenicity of *Pseudomonas morsprunorum* in relation to host specificity. *J Gen Microbiol*. 62(1), pp.315–327

Cui, H. *et al.*, 2010. *Pseudomonas syringae* effector protein AvrB perturbs *Arabidopsis* hormone signaling by activating MAP kinase 4. *Cell Host Microbe*, 7(2), pp.164–75

Cunnac, S., Lindeberg, M. & Collmer, A., 2009. *Pseudomonas syringae* type III secretion system effectors: Repertoires in search of functions. *Curr Opinion Microbiol*. 12(1), pp.53–60

Cunnac, S. *et al.*, 2011. Genetic disassembly and combinatorial reassembly identify a minimal functional repertoire of type III effectors in *Pseudomonas syringae*. *P Natl Acad Sci USA*. 108(7), pp.2975–2980

Cuppels, D.A. & Ainsworth, T., 1995. Molecular and physiological characterization of *Pseudomonas syringae* pv. tomato and *Pseudomonas syringae* pv. *maculicola* strains that produce the phytotoxin coronatine. *Appl Env Microbiol*. 61(10), pp.3530–3536

Dangl, J.L., Horvath, D.M. & Staskawicz, B.J., 2013. Pivoting the plant immune system from dissection to deployment. *Science*. 341(6147), pp.746–751

Darling, A.E., Mau, B. & Perna, N.T., 2010. Progressivemauve: Multiple genome alignment with gene gain, loss and rearrangement. *PLoS One*, 5(6), p.e11147

Daubin, V., Gouy, M. & Perrière, G., 2002. A phylogenomic approach to bacterial phylogeny: Evidence of a core of genes sharing a common history. *Genome Res*. 12(7), pp.1080–1090

Daudi, A. & O'Brien, J.A., 2012. Detection of hydrogen peroxide by DAB Staining in *Arabidopsis* leaves. *Bio Protoc*. 2(18), p.e263

Debener, T. *et al.*, 1991. Identification and molecular mapping of a single *Arabidopsis thaliana* locus determining resistance to a phytopathogenic *Pseudomonas syringae* isolate. *Plant J*. 1(3), pp.289–302

Denman, S. *et al.*, 2012. *Brenneria goodwinii* sp. nov., associated with acute oak decline in the UK. *Int J Syst Evol Microbiol*. 62(10), pp.2451–2456

Dhillon, B.K. *et al.*, 2015. IslandViewer 3: More flexible, interactive genomic island discovery, visualization and analysis. *Nucleic Acids Res*. 43(W1), pp.W104–W108

Dowen, R.H. *et al.*, 2009. A family of bacterial cysteine protease type III effectors utilizes acylation-dependent and -independent strategies tolocalize to plasma membranes. *J Biol Chem*. 284(23), pp.15867–15879

Dudnik, A. & Dudler, R., 2014. Genomics-based exploration of virulence determinants and host-specific adaptations of *Pseudomonas syringae* strains isolated from grasses. *Pathogens*. 3(1), pp.121–148

Edwards, D.J. & Holt, K.E., 2013. Beginner's guide to comparative bacterial genome analysis using next-generation sequence data. *Microb Inform Exp*. 3(1), p.2

Ellinger, D. & Voigt, C.A., 2014. Callose biosynthesis in *Arabidopsis* with a focus on pathogen response: What we have learned within the last decade. *Ann Bot*. 114(6), pp.1349–1358

FAO, 2014. FAOSTAT. Available at: http://www.fao.org [Accessed February 8, 2017]

Farhadfar, S. *et al.*, 2016. Susceptibility of cherries to bacterial canker (*Pseudomonas syringae* pv. *syringae*) in field and laboratory, *Int J Agric For*. 6(1), pp.20–27

Feil, H. *et al.*, 2005. Comparison of the complete genome sequences of *Pseudomonas syringae* pv. *syringae* B728a and pv. *tomato* DC3000. *P Natl Acad Sci USA*. 102(31), pp.11064–11069

Feil, W., Feil, H. & Copeland, A., 2012. Bacterial genomic DNA isolation using CTAB. Available at: http://1ofdmq2n8tc36m6i46scovo2e.wpengine.netdna-cdn.com/wp-content/uploads/2014/02/JGI-Bacterial-DNA-isolation-CTAB-Protocol-2012.pdf

Ferrante, P. *et al.*, 2009. Contributions of the effector gene *hopQ1-1* to differences in host range between *Pseudomonas syringae* pv. *phaseolicola* and *P. syringae* pv. *tabaci*. *Mol Plant Pathol*. 10(6), pp.837–842

Fillingham, A.J. *et al.*, 1992. Avirulence genes from *Pseudomonas syringae* pathovars *phaseolicola* and *pisi* confer specificity towards both host and non-host species. *Phys Mol Plant Pathol*. 40(1), pp.1–15

Flor, H., 1941. Inheritance of rust reaction in a cross between the flax varieties Buda

and J.W.S. *J of Agric Res*. 63(1), pp.369–388

Flor, H., 1971. Current status of the gene-for-gene concept. *Ann Rev Phytopathol*. 9(1), pp.275–296

Fouts, D.E. *et al.*, 2002. Genomewide identification of *Pseudomonas syringae* pv. *tomato* DC3000 promoters controlled by the HrpL alternative sigma factor. *P Natl Acad Sci USA*. 99(4), pp.2275–2280

Freigoun, S.O. & Crosse, J.E., 1975. Host relations and distribution of a physiological and pathological variant of *Pseudomonas morsprunorum*. *Ann Appl Biol*. 81(1), pp.317–330

French, E., Kim, B.S. & Iyer-Pascuzzi, A.S., 2016. Mechanisms of quantitative disease resistance in plants. *Semin Cell Dev Biol*. 56(1), pp.201–208

Frost, L.S. *et al.*, 2005. Mobile genetic elements: The agents of open source evolution. *Nat Rev Microbiol*. 3(9), pp.722–732

Fu, J. & Wang, S., 2011. Insights into auxin signaling in plant-pathogen interactions. *Front Plant Sci*, 2(Nov), p.74.

Gardan, L. *et al.*, 1999. DNA relatedness among the pathovars of *Pseudomonas syringae* and description of *Pseudomonas tremae* sp. nov. and *Pseudornonas cannabina* sp. nov. (ex Sutic and Dowson 1959). *Int J Sys Bacteriol*. 49(Apr), pp.469–478

Garrett, C.M.E., 1979. Screening *Prunus* rootstocks for resistance to bacterial canker, caused by *Pseudomonas morsprunorum. J Hort Sci*. 54(3), pp.189–193

Garrett, C.M.E., 1978. Pathogenic races of *Pseudomonas morsprunorum*. *In:* Station de pathologie végétale et phytobactériologie (ed). *Proceedings of the IVth International Conference on Plant pathogenic Bacteria Vol II*. Angers: Gilbert-Clarey, pp.889–890

Garrett, C.M.E., 1986. Influence of rootstock on the susceptibility of sweet cherry scions to bacterial canker, caused by *Pseudomonas syringae* pvs *morsprunorum* and *syringae*. *Plant Pathol*. 35(1), pp.114–119

Geng, X. *et al.*, 2012. The coronatine toxin of *Pseudomonas syringae* is a multifunctional suppressor of *Arabidopsis* defense. *Plant Cell*. 24(11), pp.4763–74

Getaz, M. *et al.*, 2017. A diagnostic tool for improved detection of *Xanthomonas fragariae* using a rapid and highly specific LAMP assay designed with comparative genomics. *Plant Pathol*. DOI: 10.1111/ppa.12665

Ghelardini, L. *et al.*, 2016. Drivers of emerging fungal diseases of forest trees. *Forest Ecol Manage.* 381(1), pp.235–246

Ghorbani, S. & Wilcockson, R., 2007. Reducing copper-based fungicide use in organic crop production systems. *In:* Cooper, J. Leifert, C. & Niggli, U. (eds). *Handbook of Organic Food Safety and Quality*. p.392.

Gilbert, V. *et al.*, 2008. Genetic analyses of *Pseudomonas syringae* isolates from Belgian fruit orchards reveal genetic variability and isolate-host relationships within the pathovar syringae, and help identify both races of the pathovar *morsprunorum*. *Eur J Plant Pathol*. 124(2), pp.199–218

Gilbert, V. *et al.*, 2009. Pathogenicity and aggressiveness in populations of *Pseudomonas syringae* from Belgian fruit orchards. *Eur J Plant Pathol*. 126(2), pp.263–277

Gimenez-Ibanez, S. *et al.*, 2014. The bacterial effector HopX1 targets JAZ transcriptional repressors to activate jasmonate signaling and promote infection in *Arabidopsis*. *PLoS Biol*. 12(2), p.e1001792.

Glazebrook, J., 2005. Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Ann Rev Phytopathol*. 43(1), pp.205–227

Godfrey, S. *et al.*, 2011. The stealth episome: Suppression of gene expression on the excised genomic island PPHGI-1 from *Pseudomonas syringae* pv. *phaseolicola*. *PLoS Pathogens*, 7(3), p.e1002010

Göhre, V. *et al.*, 2008. Plant pattern-recognition receptor FLS2 is directed for degradation by the bacterial ubiquitin ligase AvrPtoB. *Curr Biol*. 18(23), pp.1824–32

Gomez-Cortecero, A. *et al.*, 2016. Variation in host and pathogen in the *Neonectria*/*Malus* interaction; towards an understanding of the genetic basis of resistance to European canker. *Front Plant Sci*. 7(Sept), p.1365

Govrin, E.M. & Levine, A., 2000. The hypersensitive response facilitates plant infection by the necrotrophic pathogen *Botrytis cinerea*. *Curr Biol*. 10(13), pp.751–757

Grant, M. & Jones, J., 2009. Hormone (Dis)harmony moulds plant health and disease. *Science*. 324(5928), pp.750–752

Green, S. *et al.*, 2010. Comparative genome analysis provides insights into the evolution and adaptation of *Pseudomonas syringae* pv. *aesculi* on *Aesculus hippocastanum*. *PloS One*. 5(4), p.e10224

Greenberg, J., 2003. Identifying type III effectors of plant pathogens and analyzing their interaction with plant cells. *Curr Opin in Microbiol*. 6(1), pp.20–28

Gross, D.C. & DeBay, J.E., 1977. Population dynamics and pathogenesis of *Pseudomonas syringae* in maize and cowpea in relation to the *in vitro* production of

syringomycin. *Phytopathology*. 67(1), pp.475–483

Guo, M. *et al.*, 2009. The majority of the type III effector inventory of *Pseudomonas syringae* pv. *tomato* DC3000 can suppress plant immunity. *Mol Plant Microbe In.* 22(9), pp.1069–1080

Guo, M. *et al.*, 2016. A Bacterial effector co-opts calmodulin to target the plant microtubule network. *Cell Host Microbe*. 19(1), pp.67–78

Gurevich, A. *et al.*, 2013. QUAST: Quality assessment tool for genome assemblies. *Bioinformatics*, 29(8). pp.1072–1075

Guttman, D.S., 2009. Bacterial evolution: Dynamic genomes and the power of transformation. *Curr Biol*. 19(18), pp.R857-R859

Guttman, D.S., McHardy, A.C. & Schulze-Lefert, P., 2014. Microbial genome-enabled insights into plant–microorganism interactions. *Nat Rev Genet*. 15(12), pp.797–813

Gyles, C. & Boerlin, P., 2014. Horizontally transferred genetic elements and their role in pathogenesis of bacterial disease. *Vet Pathol*. 51(2), pp.328–340

Ham, J.H. *et al.*, 2010. Multiple activities of the plant pathogen type III effector proteins WtsE and AvrE1 require WxxxE motifs. *Mol Plant Microbe In.* 22(6), pp.703–712

Hinrichs-Berger, J., 2004. Epidemiology of *Pseudomonas syringae* pathovars associated with decline of plum trees in the southwest of Germany. *J Phytopathol*. 152(3), pp.153–160

Hmelo, L.R. *et al.*, 2015. Precision-engineering the *Pseudomonas aeruginosa* genome with two-step allelic exchange. *Nat Protoc*. 10(11), pp.1820–1841

Hockett, K.L. *et al.*, 2013. Interactions between genome architecture and virulence genes in *Pseudomonas syringae*, strain CC1557 as a model. *BioRxiv.* DOI: http://dx.doi.org/10.1101/000869

Hockett, K.L. *et al.*, 2014. *Pseudomonas syringae* CC1557: A highly virulent strain with an unusually small type III effector repertoire that includes a novel effector. *Mol Plant Microbe In.* 27(9), pp.923–32

Hogenhout, S.A. *et al.*, 2009. Emerging concepts in effector biology of plant-associated organisms. *Mol Plant Microbe In.* 22(2), pp.115–122

van der Hoorn, R.A.L. & Kamoun, S., 2008. From guard to decoy: A new model for perception of plant pathogen effectors. *Plant Cell*. 20(8), pp.2009–2017

Huelsenbeck, J.P. & Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogeny.

Bioinformatics. 17, pp.1572–1574

Hummer, K. & Janick, J., 2009. Rosaceae: Taxonomy, economic importance, genomics. In K. Folta & S. Gardiner, eds. *Genetics and Genomics of Rosaceae*. New York: Springer Science, pp. 73–84

Hunt, M. *et al.*, 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol*. 16(1), p.294

Hurley, B. *et al.*, 2014. The *Pseudomonas syringae* type III effector HopF2 suppresses *Arabidopsis* stomatal immunity. *PLoS One*. 9(12), p.e114921

Huson, D.H. *et al.*, 2007. Dendroscope: An interactive viewer for large phylogenetic trees. *BMC Bioinformatics*, 8(Nov), p.460

Iakovidis, M. *et al.*, 2016. Effector triggered immune response in *Arabidopsis thaliana* is a quantitative trait. *Genetics*. 204(1), pp.337-353

Jackson, D. & Looney, N., 2011. Stone fruits. In Jackson, D. Looney, N. & Morley-Bunker, M. (eds). *Temperate and Subtropical Fruit Production*. Cambridge: CABI, pp. 165–170

Jackson, R.W. *et al.*, 1999. Identification of a pathogenicity island, which contains genes for virulence and avirulence, on a large native plasmid in the bean pathogen *Pseudomonas syringae* pathovar *phaseolicola*. *P Natl Acad Sci USA*, 96(19), pp.10875–10880

Jackson, R.W. *et al.*, 2000. Excision from tRNA genes of a large chromosomal region, carrying avrPphB, associated with race change in the bean pathogen, *Pseudomonas syringae* pv. *phaseolicola*. *Mol Microbiol*. 38(2), pp.186–197

Jackson, R.W. *et al.*, 2011. The influence of the accessory genome on bacterial pathogen evolution. *Mob Genet Elements*. 1(1), pp.55–65

Jacques, M.A. *et al.*, 2016. Using Ecology, physiology, and genomics to understand host specificity in *Xanthomonas*. *Ann Rev Phytopath*. 54(1), pp.163-187

Janjusevic, R. *et al.*, 2006. A bacterial inhibitor of host programmed cell death defenses is an E3 ubiquitin ligase. *Science*. 311(5758), pp.222–226

Jehl, M.A., Arnold, R. & Rattei, T., 2011. Effective - A database of predicted secreted bacterial proteins. *Nucleic Acids Res*. 39(S1), pp.591–595

Jelenska, J., van Hal, J.A. & Greenberg, J.T., 2010. *Pseudomonas syringae* hijacks plant stress chaperone machinery for virulence. *Proc Natl Acad Sci USA*. 107(29), pp.13177–13182

Jenner, C. *et al.*, 1991. Gene-for-gene interactions between *Pseudomonas syringae* pv. *phaseolicola* and *Phaseolus*. *Mol Plant Microbe In*. 4(6), pp.553–562

Joardar, V. *et al.*, 2005. Whole-genome sequence analysis of *Pseudomonas syringae* pv. *phaseolicola* 1448A reveals divergence among pathovars in genes involved in virulence and transposition. *J Bacteriol*. 187(18), pp.6488–6498.

Jones, J.D.G. & Dangl, J.L., 2006. The plant immune system. *Nature*. 444(7117), pp.323–9

Kaluzna, M. *et al.*, 2010. Characterisation and genetic diversity of *Pseudomonas syringae* from stone fruits and hazelnut using reptitive-PCR and MLST. *J Plant Pathol.* 92(3), pp.781–787

Kamiunten, H., Nakaol, T. & Oshida, S., 2000. Agent of bacterial gall of cherry tree. *J Gen Plant Pathol.* 66, pp.219–224

Kang, Y. *et al.*, 2014. HopW1 from *Pseudomonas syringae* disrupts the actin cytoskeleton to promote virulence in *Arabidopsis*. *PLoS Pathogens*. 10(6), p.e1004232

Kappel, F. *et al.*, 2012. Cherry. *In:* Badenes, M. & Byrne, D. (eds). *Plant Breeding.* New York: Springer Science, pp.459-504

Karasov, T., Horton, M. & Bergelson, J., 2014. Genomic variability as a driver of plant–pathogen coevolution? *Curr Opin Plant Biol*. 18(1), pp.24–30

Katoh, K. *et al.*, 2002. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*. 30(14), pp.3059–3066.

Kearney, B. & Staskawicz, B.J., 1990. Characterization of IS476 and its role in bacterial spot disease of tomato and pepper. *J Bacteriol*. 172(1), pp.143–148

Kearse, M. *et al.*, 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 28(12), pp.1647–1649

Kelley, D.R., Schatz, M.C. & Salzberg, S.L., 2010. Quake: Quality-aware detection and correction of sequencing errors. *Genome Biol*. 11(11), p.R116

Kim, J.F. *et al.*, 1998. Sequences related to transposable elements and bacteriophages flank avirulence genes of *Pseudomonas syringae*. *Mol Plant Microbe In*. 11(12), pp.1247–1252.

Kim, M. *et al.*, 2014. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J of Syst and Evol Microbiol*. 64(2), pp.346–351

Kim, Y.J., Lin, N.C. & Martin, G.B., 2002. Two distinct *Pseudomonas* effector proteins interact with the Pto kinase and activate plant immunity. *Cell*. 109(5), pp.589–598

Knief, C., Delmotte, N. & Vorholt, J.A., 2011. Bacterial adaptation to life in association with plants - A proteomic perspective from culture to *in situ* conditions. *Proteomics*. 11(15), pp.3086–3105

Kobayashi, D.Y., Tamaki, S.J. & Keen, N.T., 1989. Cloned avirulence genes from the tomato pathogen *Pseudomonas syringae* pv. *tomato* confer cultivar specificity on soybean. *P Natl Acad Sci USA*. 86(1), pp.157–61

Koren, S. *et al.*, 2016. Canu: Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *BioRxiv*. DOI: https://doi.org/10.1101/071282

Krzesinska, E.Z., Nina, A. & Azarenko, M., 1992. Excised twig assay to evaluate cherry rootstocks for tolerance to *Pseudomonas syringae* pv. *syringae*. *Hort Sci.* 27(2), pp.153–155

Krzywinski, M. *et al.*, 2009. Circos: An information aesthetic for comparative genomics. *Genome Res*. 19(9), pp.1639–1645

Kunkeaw, S., Tan, S. & Coaker, G., 2010. Molecular and evolutionary analyses of *Pseudomonas syringae* pv. *tomato* race 1. *Mol Plant Microbe In*. 23(4), pp.415–424

Kus, J.V. *et al.*, 2002. Age-related resistance in *Arabidopsis* is a developmentally regulated defense response to *Pseudomonas syringae*. *Plant Cell*. 14(2), pp.479–90

Kvitko, B.H. *et al.*, 2009. Deletions in the repertoire of *Pseudomonas syringae* pv. *tomato* DC3000 type III secretion effector genes reveal functional overlap among effectors. *PLoS Pathogens*. 5(4), p.e1000388

Lamichhane, J.R. *et al.*, 2014. Chapter Four - Disease and Frost Damage of Woody Plants Caused by Pseudomonas syringae: Seeing the Forest for the Trees. *In:* Boyer, J.S. Alexander, M. and Kamprath, E.J (eds). *Advances in Agronomy Vol 127.* San Diego: Academic Press, pp. 235-295

Landgraf, A. *et al.*, 2006. Different versions of *Pseudomonas syringae* pv. tomato DC3000 exist due to the activity of an effector transposon. *Mol Plant Pathol.* 7(5), pp.355–364

Lang, G.A., 2000. Precocious, dwarfing, and productive - How will new cherry rootstocks impact the sweet cherry industry? *Hort Tech.* 10(4), pp.719–725

Langmead, B. & Salzberg, S., 2013. Fast gapped-read alignment with Bowtie2. *Nat Methods*. 9(4), pp.357–359

Larkin, M.A. *et al.*, 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*. 23(21),

pp.2947–2948

Leben, C., 1981. How plant-pathogenic bacteria survive. *Plant Dis*. 65(8), pp.633–637

Lee, A.H.Y. *et al.*, 2012. A bacterial acetyltransferase destroys plant microtubule networks and blocks secretion. *PLoS Pathogens*. 8(2), p.e1002523

Lee, S., Whitaker, V.M. & Hutton, S.F., 2016. Potential applications of non-host resistance for crop improvement. *Front Plant Sci*. 7(Jul), pp.1–6

Lewis, J.D. *et al.*, 2011. The YopJ superfamily in plant-associated bacteria. *Mol Plant Pathol.* 12(9), pp.928–937

Li, B. *et al.*, 2015. Rapid, automated detection of stem canker symptoms in woody perennials using artificial neural network analysis. *Plant Methods*. 11(1), p.57.

Li, H. *et al.*, 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 25(16), pp.2078–2079

Li, L., Stoeckert, C.J.J. & Roos, D.S., 2003. OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res*. 13(9), pp.2178–2189

Liang, L. *et al.*, 1994. Variation in virulence, plasmid content and genes for coronatine synthesis between *Pseudomonas syringae* pv. *morsprunorum* and *P.s. syringae* from *Prunus*. *Plant Dis*. 78(4), pp.389–392

Lin, N.C. & Martin, G.B., 2005. An avrPto / avrPtoB mutant of *Pseudomonas syringae* pv . *tomato* DC3000 does not elicit Pto-mediated resistance and is less virulent on tomato. *Mol Plant Microbe In.* 18(1), pp.43–51

Lin, N.C. *et al.*, 2006. Diverse AvrPtoB homologs from several *Pseudomonas syringae* pathovars elicit Pto-dependent resistance and have similar virulence activities. *Appl Env Microbiol*. 72(1), pp.702–712

Lin, N.C. & Martin, G.B., 2007. Pto- and Prf-mediated recognition of AvrPto and AvrPtoB restricts the ability of diverse *Pseudomonas syringae* pathovars to infect tomato. *Mol Plant Microbe In.* 20(7), pp.806–815

Lindeberg, M. *et al.*, 2005. Proposed guidelines for a unified nomenclature and phylogenetic analysis of type III Hop effector proteins in the plant pathogen *Pseudomonas syringae*. *Mol Plant Microbe In.* 18(4), pp.275–282.

Lindeberg, M., Cunnac, S. & Collmer, A., 2009. The evolution of *Pseudomonas syringae* host specificity and type III effector repertoires. *Mol Plant Pathol.* 10(6), pp.767–775

Lindeberg, M., Cunnac, S. & Collmer, A., 2012. *Pseudomonas syringae* type III effector repertoires: last words in endless arguments. *Trends Microbiol*. 20(4), pp.199–208

Lindgren, P.B., 1997. The role of hrp genes during plant-bacterial interactions. *Ann Rev Phytopathol*. 35(40), pp.129–52

Lindow, S.E. *et al.*, 1989. Localization of ice nucleation activity and the *iceC* gene product in *Pseudomonas syringae* and *Escherichia coli*. *Mol Plant Microbe In*. 2(5), p.262

Liu, W. *et al.*, 2015. IBS: An illustrator for the presentation and visualization of biological sequences. *Bioinformatics*. 31(20), pp.3359–3361

Lo, T. *et al.*, 2016. The HopF family of *Pseudomonas syringae* type III secreted effectors. *Mol Plant Pathol*. pp.1–12

Loewe, V., González, M. & Balzarini, M., 2013. Wild cherry tree (*Prunus avium* L.) growth in pure and mixed plantations in South America. *Forest Ecol Manage*. 306(Oct), pp.31–41

Loman, N.J. & Quinlan, A.R., 2014. Poretools: A toolkit for analyzing nanopore sequence data. *Bioinformatics*. 30(23), pp.3399–3401

Long, J. & Olsen, L., 2013. Sweet cherry cultivars for brining, freezing, and canning in Oregon. Available at: https://catalog.extension.oregonstate.edu/em9056 [Accessed March 4, 2017]

López-Solanilla, E. *et al.*, 2004. HopPtoN is a *Pseudomonas syringae* Hrp (type III secretion system) cysteine protease effector that suppresses pathogen-induced necrosis associated with both compatible and incompatible plant interactions. *Mol Microbiol*. 54(2), pp.353–65

Lovell, H.C. *et al.*, 2009. Bacterial evolution by genomic island transfer occurs via DNA transformation in planta. *Curr Biol*. 19(18), pp.1586–1590

Lovell, H.C. *et al.*, 2011. *In planta* conditions induce genomic changes in *Pseudomonas syringae* pv. *phaseolicola*. *Mol Plant Pathol*, 12(2), pp.167–176

Lu, H. *et al.*, 2008. Acquisition and evolution of plant pathogenesis-associated gene clusters and candidate determinants of tissue-specificity in *Xanthomonas*. *PLoS One*, 3(11), p.e3828

Ma, Q.S. *et al.*, 1988. Identification of DNA sequences involved in host specificity in the pathogenesis of *Pseudomonas solanacearum* strain T2005. *Mol Plant Microbe In.* 1(4), pp.169–174

Ma, W. *et al.*, 2006. Type III effector diversification via both pathoadaptation and horizontal transfer in response to a coevolutionary arms race. *PLoS Genet*. 2(12), p.e209

Macho, A. *et al.*, 2014. A bacterial tyrosine phosphatase inhibits plant pattern recognition recpetor activation. *Science*. 343(6178), pp.1509–1512

Mackey, D. *et al.*, 2002. RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in *Arabidopsis*. *Cell*. 108(6), pp.743–754

Manoharan, B. *et al.*, 2015. The identification of genes important in *Pseudomonas syringae* pv. *phaseolicola* plant colonisation using in vitro screening of transposon libraries. *PLoS ONE*. 10(9), pp.1–19

Mansfield, J. *et al.*, 2012. Top 10 plant pathogenic bacteria in molecular plant pathology. *Mol Plant Pathol*. 13(6), pp.614–629

Marcelletti, S. & Scortichini, M., 2014. Definition of plant-pathogenic *Pseudomonas* genomospecies of the *Pseudomonas syringae* complex Through multiple comparative Approaches. *Phytopathology*. (58), pp.1274–1282

Martínez-García, P.M. *et al.*, 2015. Bioinformatics analysis of the complete genome sequence of the mango tree pathogen *Pseudomonas syringae* pv. *syringae* UMAF0158 reveals traits relevant to virulence and epiphytic lifestyle. *PLoS One*. 10(8), pp.1–26

Matas, I.M. *et al.*, 2014. Translocation and functional analysis of Pseudomonas savastanoi pv. savastanoi NCPPB 3335 type III secretion system effectors reveals two novel effector families of the *Pseudomonas syringae* complex. *Mol Plant Microbe In.* 27(5), pp.424-436

Matthews, P., 1979. Progress in breeding cherries for resistance to bacterial canker. Institut National de la Recherche Agronomique (ed). *In: Proceedings of the Eucarpia Fruit Section Symposium on Tree Fruit Breeding*. Paris: Eucarpia, pp. 157–174

Mazel, D., 2006. Integrons: Agents of bacterial evolution. *Nat Rev Microbiol*. 4(8), pp.608–620

McCann, H.C. & Guttman, D.S., 2008. Evolution of the type III secretion system and its effectors in plant-microbe interactions. *New Phytol*. 177(1), pp.33–47

McCann, H.C. *et al.*, 2013. Genomic analysis of the Kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae* provides insight into the origins of an emergent plant disease. *PLoS Pathogens*. 9(7), p.e1003503

McCann, H.C. *et al.*, 2016. Origin and evolution of a pandemic lineage of the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae*. *BioRxiv*.
DOI: https://doi.org/10.1101/085613

McDonald, B.A. & Linde, C., 2002. The population genetics of plant pathogens and

breeding strategies for durable resistance. *Euphytica*. 124(2), pp.163–180

Ménard, M. *et al.*, 2003. *Pseudomonas syringae* pv. *avii* (pv. nov.), the causal agent of bacterial canker of wild cherries (*Prunus avium*) in France. *Eur J Plant Pathol*. 109(6), pp.565–576

de Mendiburu, F., 2016. Agricolae: Statistical procedures for agricultural research. Available at: https://cran.r-project.org/package=agricolae [Accessed March 4, 2017]

Michelmore, R.W., Christopoulou, M. & Caldwell, K.S., 2013. Impacts of resistance gene genetics, function, and evolution on a durable future. *Ann Rev of Phytopathol*. 51(1), pp.291–319

Misaghi, I. & Grogan, R., 1969. Nutritional and biochemical comparisons of plant-pathogenic and saprophytic fluorescent pseudomonads. *Phytopathology*. 59(10), pp.1436–1450

Monteil, C.L. *et al.*, 2013. Nonagricultural reservoirs contribute to emergence and evolution of *Pseudomonas syringae* crop pathogens. *New Phytol*. 199(3), pp.800–811

Monteil, C.L. *et al.*, 2016. Population-genomic insights into emergence, crop-adaptation, and dissemination of *Pseudomonas syringae* pathogens. *Microbial Genomics*. DOI:10.1099/mgen.0.000089

Moragrega, C. & Llorente, I., 2003. Susceptibility of European pear cultivars to *Pseudomonas syringae* pv. *syringae* using immature fruit and detached leaf assays. *Eur J Plant Pathol*. 109(4), pp.319–326.

Moriarty Lemmon, E. & Lemmon, A.R., 2013. High-throughput genomic data in systematics and phylogenetics. *Ann Rev Ecol Evol Syst*. 44(1), pp.99–121

Morris, C.E. *et al.*, 2007. Surprising niche for the plant pathogen *Pseudomonas syringae*. *Infect, Genet and Evol*. 7(1), pp.84–92

Morris, C.E., Monteil, C.L. & Berge, O., 2013. The life history of *Pseudomonas syringae*: linking agriculture to earth system processes. *Ann Rev Phytopathol*. 51(1), pp.85–104

Morris, C.E. *et al.*, 2017. Frontiers for research on the ecology of plant pathogenic bacteria: Fundamentals for sustainability. *Mol Plant Pathol*. 18(2), pp.1–20

Moulton, J. *et al.*, 1993. Changes in cultivar-specificity toward pea can result from transfer of plasmid RP4 and other incompatibility group P1 replicons to *Pseudomonas syringae* pv. *pisi*. *J Gen Microbiol*. 139(12), pp.3149–3155

Mur, L.A.J. *et al.*, 2008. The hypersensitive response; The centenary is upon us but how much do we know? *J Exp Bot*. 59(3), pp.501–520

Neale, H.C. *et al.*, 2013. *In planta* induced changes in the native plasmid profile of *Pseudomonas syringae* pathover *phaseolicola* strain 1302A. *Plasmid*. 70(3), pp.420–424

Neale, H.C. *et al.*, 2016. A low frequency persistent reservoir of a genomic island in a pathogen population ensures island survival and improves pathogen fitness in a susceptible host. *Env Microbiol*. 18(11), pp.4144–4152

Nicaise, V. *et al.*, 2013. Pseudomonas HopU1 modulates plant immune receptor levels by blocking the interaction of their mRNAs with GRP7. *EMBO J*. 32(5), pp.701–12

Nomura, K. *et al.*, 2006. A bacterial virulence protein suppresses host innate immunity to cause plant disease. *Science.* 313(5784), pp.220–223

Nowell, R.W. *et al.*, 2014. The extent of genome flux and its role in the differentiation of bacterial lineages. *Genome Biol Evol*. 6(6), pp.1514–1529

Nowell, R.W. *et al.*, 2016. Comparative genomics reveals genes significantly associated with woody hosts in the plant pathogen *Pseudomonas syringae*. *Mol Plant Pathol*. pp.1–16

Nyerges, Á. *et al.*, 2016. A highly precise and portable genome engineering method allows comparison of mutational effects across bacterial species. *Proc Natl Acad Sci USA*. 113(9), pp.2502–2507

O'Brien, H.E., Thakur, S. & Guttman, D.S., 2011. Evolution of plant pathogenesis in *Pseudomonas syringae*: A genomics perspective. *Ann Rev Phytopathol*. 49(1), pp.269–89

O'Brien, H.E. *et al.*, 2012. Extensive remodeling of the *Pseudomonas syringae* pv. *avellanae* type III secretome associated with two independent host shifts onto hazelnut. *BMC Microbiol*. 12(1), p.141

Pagel, M., 2004. Detecting correlated evolution on phylogenies: A general method for the comparative analysis of discrete characters. *Proc R Soc Lond (Biol).* 255(1342), pp.37–45

Pan, X., Yang, Y. & Zhang, J.R., 2014. Molecular basis of host specificity in human pathogenic bacteria. *Emerg Microbes Infect*. 3(3), p.e23

Parkinson, N. *et al.*, 2011. Rapid phylogenetic identification of members of the *Pseudomonas syringae* species complex using the *rpoD* locus. *Plant Pathol*. 60(1), pp.338–344

Pink, D.A.C., 2002. Strategies using genes for non-durable disease resistance. *Euphytica*. 124(2), pp.227–236

Pitman, A.R. *et al.*, 2005. Exposure to host resistance mechanisms drives evolution of bacterial virulence in plants. *Curr Biol.* 15(24), pp.2230–2235

Poland, J.A. *et al.*, 2009. Shades of gray: The world of quantitative disease resistance. *Trends Plant Sci*. 14(1), pp.21–29

Posada, D., 2008. jModelTest: Phylogenetic model averaging. *Mol Biol Evol*. 25(7), pp.1253–1256

Pour, G.N. & Taghavi, S.M., 2011. Comparison of *P. syringae* pv. *syringae* from different hosts based on pathogenicity and BOX-PCR in Iran. *J Agric Sci Tech*. 13(3), pp.431–442

Press, M.O. *et al.*, 2013. Genome-scale co-evolutionary inference identifies functions and clients of bacterial Hsp90. *PLoS Genet.* 9(7), p.e1003631

Preston, G.M., 2000. *Pseudomonas syringae* pv. *tomato*: The right pathogen, of the right plant, at the right time. *Mol Plant Pathol.* 1(5), pp.263–275

Price, M.N., Dehal, P.S. & Arkin, A.P., 2010. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS One*. 5(3), p.e9490

Python, 2017. Python programming language version 2.7. Available at: http://www.python.org [Accessed March 4, 2017]

Quevillon, E. *et al.*, 2005. InterProScan: Protein domains identifier. *Nucleic Acids Res*. 33(S2), pp.116–120

R Core Team, 2012. *R: A language and environment for statistical computing*, Vienna: R Foundation for Statistical Computing.

Reignault, P. & Sancholle, M., 2005. Plant-pathogen interactions: Will the understanding of common mechanisms lead to the unification of concepts? *C R Biol*. 328(9), pp.821–833

Renick, L.J., Cogal, A.G. & Sundin, G.W., 2008. Phenotypic and genetic analysis of epiphytic *Pseudomonas syringae* populations from sweet cherry in michigan. *Plant Dis*. 92(3), pp.372–378

Reuter, S. *et al.*, 2013. Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology. *JAMA intern Med.* 173(15), pp.1397–404

Rezaei, R. & Taghavi, S.M., 2014. Host specificity, pathogenicity and the presence of virulence genes in Iranian strains of *Pseudomonas syringae* pv. *syringae* from different hosts. *Arch Phytopathol Plant Protect*. 47(19), pp.2377–2391

RHS, Bacterial canker. Available at: https://www.rhs.org.uk/advice/profile?PID=86

[Accessed March 4, 2017]

Rietman, H. *et al.*, 2012. Qualitative and quantitative late blight resistance in the potato cultivar Sarpo Mira is determined by the perception of five distinct RXLR effectors. *Mol Plant Microbe In*. 910(7), pp.910–919

Rioux, D., 1996. Compartmentalization in trees: New findings during the study of dutch elm disease. *In:* Nicole, M. & Gianinazzi-Pearson, V. (eds). *Histology, ultrastructure and molecular cytology of plant-microorganism interactions*. Berlin: Springer Netherlands, pp. 211–225

Rivas, L.A. *et al.*, 2005. Changes in race-specific virulence in *Pseudomonas syringae* pv. *phaseolicola* are associated with a chimeric transposable element and rare deletion events in a plasmid-borne pathogenicity island. *Appl Env Microbiol.* 71(7), pp.3778–3785

Roberts, S., 2012. *HNS 179: Management of bacterial canker in Prunus spp.* Available at: http://www.hdc.org.uk [Accessed March 4, 2017]

Robertson, A.E. *et al.*, 2004. Relationship between avirulence gene (*avrA*) diversity in *Ralstonia solanacearum* and bacterial wilt incidence. *Mol Plant Microbe In*. 17(12), pp.1376–1384

Rodríguez-Palenzuela, P. *et al.*, 2010. Annotation and overview of the *Pseudomonas savastanoi* pv. *savastanoi* NCPPB 3335 draft genome reveals the virulence gene complement of a tumour-inducing pathogen of woody hosts. *Env Microbiol*. 12(6), pp.1604–1620

Rohmer, L., Guttman, D.S. & Dangl, J.L., 2004. Diverse evolutionary mechanisms shape the type III effector virulence factor repertoire in the plant pathogen *Pseudomonas syringae*. *Genetics*, 167(3), pp.1341–1360

Ronald, P.C., Salmeron, J.M. & Carland, F.M., 1992. The cloned avirulence gene *avrPto* induces disease resistance in tomato cultivars containing the Pto resistance gene. *J Bacteriol.* 174(5), pp.1604-1611

Rosebrock, T.R. *et al.*, 2007. A bacterial E3 ubiquitin ligase targets a host protein kinase to disrupt plant immunity. *Nature*. 448(7151), pp.370-374

Rufián, J. *et al.*, 2016. *Pseudomonas syringae* differentiates into phenotypically distinct subpopulations during colonization of a plant host. *Env Microbiol*. 18(10), pp.3593–3605

Ruinelli, M. *et al.*, 2016. Comparative genomics-informed design of two LAMP assays for detection of the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae* and discrimination of isolates belonging to the pandemic biovar 3. *Plant Pathol*. pp.140–149

Ryan, R.P. *et al.*, 2011. Pathogenomics of *Xanthomonas*: Understanding bacterium-plant interactions. *Nat Rev Microbiol*. 9(5), pp.344–355

Sacristán, S. & GarcÍa-Arenal, F., 2008. The evolution of virulence and pathogenicity in plant pathogen populations. *Mol Plant Pathol*. 9(3), pp.369–384

Şahin, F., 2001. Severe outbreak of bacterial speck, caused by *Pseudomonas syringae* pv. *tomato*, on field-grown tomatoes in the eastern Anatolia region of Turkey. *Plant Pathol*. 50(6), p.799

Santi, F. *et al.*, 2004. Screening wild cherry (*Prunus avium*) for resistance to bacterial canker by laboratory and field tests. *Forest Pathol*. 34(6), pp.349–362

Sarkar, S.F. & Guttman, D.S., 2004. Evolution of the core genome of *Pseudomonas syringae*, a highly clonal, endemic plant pathogen. *Appl Env Microbiol*. 70(4), pp.1999–2012

Sarkar, S.F. *et al.*, 2006. Comparative genomics of host-specific virulence in *Pseudomonas syringae*. *Genetics*. 174(2), pp.1041–1056

Schellenberg, B., Ramel, C. & Dudler, R., 2010. *Pseudomonas syringae* virulence factor syringolin A counteracts stomatal immunity by proteasome inhibition. *Mol Plant Microbe In.* 23(10), pp.1287–1293

Schenk, S.T. & Schikora, A., 2015. Staining of callose depositions in root and leaf tissues. *Bio Protoc*. 5(6), pp.3–6

van Schie, C.C.N. & Takken, F.L.W., 2014. Susceptibility genes 101: How to be a good host. *Ann Rev Phytopathol.* 52(1), pp.551-581

Scholz-Schroeder, B.K. *et al.*, 2001. The contribution of syringopeptin and syringomycin to virulence of *Pseudomonas syringae* pv. *syringae* strain B301D on the basis of *sypA* and *syrB1* biosynthesis mutant analysis. *Mol Plant Microbe In*. 14(3), pp.336–348

Schulze-Lefert, P. & Panstruga, R., 2011. A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. *Trends Plant Sci*. 16(3), pp.117–25

Scortichini, M., 2010. Epidemiology and predisposing factors of some major bacterial diseases of stone and nut fruit trees species. *J Plant Pathol*. 92(S1), pp.73–78

Selote, D. & Kachroo, A., 2010. RPG1-B-derived resistance to AvrB-expressing *pseudomonas syringae* requires RIN4-like proteins in soybean. *Plant Physiol*. 153(3), pp.1199–1211

Shimono, M. *et al.*, 2016. The *Pseudomonas syringae* Type III effector HopG1 induces actin remodeling to promote symptom development and susceptibility during Infection. *Plant Physiol*. 171(3), pp.2239–2255

Shindo, T. *et al.*, 2016. Screen of non-annotated small secreted proteins of *Pseudomonas syringae* reveals a virulence factor that inhibits tomato immune proteases. *PLoS Pathogens*. 12(9), pp.1–24

Siguier, P. *et al.*, 2006. ISfinder: The reference centre for bacterial insertion sequences. *Nucleic Acids Res*. 34(Database issue), pp.D32-D36

Sohn, K.H., Zhang, Y. & Jones, J.D.G., 2009. The *Pseudomonas syringae* effector protein, AvrRPS4, requires in planta processing and the KRVY domain to function. *Plant J*. 57(6), pp.1079–1091

Spiers, A.J., Buckling, A. & Rainey, P.B., 2000. The causes of *Pseudomonas* diversity. *Microbiol*. 146(2000), pp.2345–2350

Spoel, S.H. & Dong, X., 2012. How do plants achieve immunity? Defence without specialized immune cells. *Nat Rev Immunol*. 12(2), pp.89–100

Spotts, R.A. *et al.*, 2010. Bacterial canker of sweet cherry in Oregon—Infection of horticultural and natural wounds, and resistance of cultivar and rootstock combinations. *Plant Dis*. 94(3), pp.345–350

Stamatakis, A., 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 30(9), pp.1312–1313

Staskawicz, B. *et al.*, 1987. Molecular characterization of cloned avirulence genes from race-0 and race-1 of *Pseudomonas syringae* pv. *glycinea*. *J Bacteriol*. 169(12), pp.5789–5794

Stavrinides, J., Ma, W. & Guttman, D.S., 2006. Terminal reassortment drives the quantum evolution of type III effectors in bacterial pathogens. *PLoS Pathogens*. 2(10), pp.0913–0921

Stevens, C. *et al.*, 1998. Sequence variations in alleles of the avirulence gene *avrPphE*.R2 from *Pseudomonas syringae* pv. *phaseolicola* lead to loss of recognition of the AvrPphE protein within bean cells and a gain in cultivar-specific virulence. *Mol Microbiol*. 29(1), pp.165–77

Sundin, G.W. & Bender, C.L., 1993. Ecological and genetic analysis of copper and streptomycin resistance in *Pseudomonas syringae* pv. *syringae*. *Appl Env Microbiol*. 59(4), pp.1018–1024

Takeuchi, K. *et al.*, 2003. Flagellin glycosylation island in *Pseudomonas syringae* pv . *glycinea* and its role in host specificity. *J Bacteriol.* 185(22), pp.6658-6665

Tao, Y. *et al.*, 2003. Quantitative nature of *Arabidopsis* responses during compatible and incompatible interactions with the bacterial pathogen *Pseudomonas syringae*. *Plant Cell*. 15(2), pp.317–30

Thomma, B.P.H.J., Nürnberger, T. & Joosten, M.H.A.J., 2011. Of PAMPs and effectors: The blurred PTI-ETI dichotomy. *Plant cell*. 23(1), pp.4–15

Tobutt, K., 1985. New approaches to breeding sweet cherry scion varieties at East Malling with particular reference to small tree size. *Acta Hort*. 169, pp.43–50

Torres, M.A., Jones, J.D. & Dangl, J.L., 2005. Pathogen-induced, NADPH oxidase-derived reactive oxygen intermediates suppress spread of cell death in *Arabidopsis thaliana*. *Nat Genet*. 37(10), pp.1130–1134

de Torres, M. *et al.*, 2006. *Pseudomonas syringae* effector AvrPtoB suppresses basal defence in *Arabidopsis*. *Plant J*. 47(3), pp.368–382

Untergasser, A. *et al.*, 2012. Primer3-new capabilities and interfaces. *Nucleic Acids Res*. 40(15), pp.1–12

Üstün, S. *et al.*, 2014. HopZ4 from *Pseudomonas syringae*, a member of the HopZ type III effector family from the YopJ Superfamily, inhibits the proteasome in plants. *Mol Plant Microbe In.* 27(7), pp.611–623

Üstün, S. *et al.*, 2016. The proteasome acts as a hub for plant immunity and is targeted by Pseudomonas type-III effectors. *Plant Physiol.* 172(3), pp.1941-1958

Vencato, M. *et al.*, 2006. Bioinformatics-enabled identification of the HrpL regulon and type III secretion system effector proteins of *Pseudomonas syringae* pv. *phaseolicola* 1448A. *Mol Plant Microbe In*. 19(11), pp.1193–1206

Vicente, J.G. & Roberts, S.J., 2003. Screening wild cherry micropropagated plantlets for resistance to bacterial canker. In: Santa Lacobellis N., Collmer A., Hutcheson S.W., Mansfield J.W., Morris C., Murillo J., Schaad N.W., Stead D.E., Surico G. & Ullrich M.S., (eds). Pseudomonas syringae *and related pathogens*. Netherlands: Springer, pp.1–8.

Vicente, J.G. & Roberts, S.J., 2007. Discrimination of *Pseudomonas syringae* isolates from sweet and wild cherry using rep-PCR. *Eur J Plant Pathol*. 117(4), pp.383–392

Vinatzer, B.A. *et al.*, 2006. The type III effector repertoire of *Pseudomonas syringae* pv. *syringae* B728a and its role in survival and disease on host and non-host plants. *Mol Microbiol*. 62(1), pp.26–44

Vleeshouwers, V.G.A.A. & Oliver, R.P., 2014. Effectors as tools in disease resistance breeding against biotrophic, hemibiotrophic, and necrotrophic plant pathogens. *Mol Plant Microbe In.* 27(3), pp.196–206.

Vos, M. *et al.*, 2015. Rates of lateral gene transfer in prokaryotes: High but why? *Trends Microbiol*. 23(10), pp.598–605

Walker, B.J. *et al.*, 2014. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One*. 9(11), p.e112963

Wang, Y. *et al.*, 2010. A *Pseudomonas syringae* ADP-ribosyltransferase inhibits *Arabidopsis* mitogen-activated protein kinase kinases. *Plant Cell*. 22(6), pp.2033–2044

Warnes, G. *et al.*, 2016. gplots: Various R programming tools for plotting data. Available at: https://cran.r-project.org/package=gplots.

Washington, E.J. *et al.*, 2016. *Pseudomonas syringae* type III effector HopAF1 suppresses plant immunity by targeting methionine recycling to block ethylene induction. *P Natl Acad Sci USA*. 113(25), pp.E3577–E3586

Wei, C.F. *et al.*, 2007. A *Pseudomonas syringae* pv. *tomato* DC3000 mutant lacking the type III effector HopQ1-1 is able to cause disease in the model plant *Nicotiana benthamiana*. *Plant J.* 51(1), pp.32–46

Whalen, M.C. *et al.*, 1991. Identification of *Pseudomonas syringae* pathogens of *Arabidopsis* and a bacterial locus determining avirulence on both *Arabidopsis* and soybean. *Plant Cell*. 3(1), pp.49–59

Whiteside, M.D. *et al.*, 2016. SuperPhy: Predictive genomics for the bacterial pathogen *Escherichia coli*. *BMC Microbiol*. 16(1), p.65

Whitham, S.A. *et al.*, 2016. Molecular soybean-pathogen interactions. *Ann Rev Phytopathol*. 54(1), pp.443–468

Wickham, H., 2009. *ggplot2: Elegant graphics for data analysis*, New York: Springer-Verlag.

Wilson, E., 1933. Bacterial canker of stone-fruit trees in California. *Hilgardia*. 8(1), p.83

Wilson, E., 1939. Factors affecting development of the bacterial canker of stone fruits. *Hilgardia*. 12(1), pp.259–298

Wimalajeewa, D. *et al.*, 1991. Chemical control of bacterial canker (*Pseudomonas syringae* pv. *syringae*) of apricot and cherry in Victoria. *Aust J Expl Agr*. 31(5), pp.705–708

Wood, R., 1972. Introduction: Disease resistance in plants. *Proc R Sco Lond (Biol)*. 181(1064), pp.213–232

Wormald, H., 1932. Bacterial diseases of stone-fruit trees in Britain: IV. The organism

causing bacterial canker of plum trees. *T Brit Mycol Soc*. 17(3), pp.157–169

Wu, S. *et al.*, 2011. Bacterial effector HopF2 suppresses *Arabidopsis* innate immunity at the plasma membrane. *Mol Plant Microbe In.* 24(5), pp.585–593

Xiang, T. *et al.*, 2008. *Pseudomonas syringae* effector AvrPto blocks innate immunity by targeting receptor kinases. *Curr Biol*, 18(1), pp.74–80

Xiang, T. *et al.*, 2011. BAK1 is not a target of the *Pseudomonas syringae* effector AvrPto. *Mol Plant Microbe In.* 24(1), pp.100–107

Xiao, Y. *et al.*, 1994. Identification of a putative alternate sigma factor and characterization of a multicomponent regulatory cascade controlling the expression of *Pseudomonas syringae* pv. *syringae Pss*61 *hrp* and *hrmA* genes. *J Bacteriol*. 176(4), pp.1025–1036

Xin, X.F. & He, S.Y., 2013. *Pseudomonas syringae* pv. *tomato* DC3000: A model pathogen for probing disease susceptibility and hormone signaling in plants. *Ann Rev Phytopathol*. 51(1), pp.473–498

Yan, S. *et al.*, 2008. Role of recombination in the evolution of the model plant pathogen *Pseudomonas syringae* pv. *tomato* DC3000, a very atypical tomato strain. *Appl Env Microbiol*. 74(10), pp.3171–3181

Yang, L. *et al.*, 2017. *Pseudomonas syringae* type III effector HopBB1 promotes host transcriptional repressor degradation to regulate phytohormone responses and virulence. *Cell Host Microbe*. 21(2), pp.156–168

Yanisch-Perron, C., Vieira, J. & Messing, J., 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*. 33(1), pp.103–119

Yessad-Carreau, S., Manceau, C. & Luisetti, J., 1994. Occurrence of specific reactions induced by *Pseudomonas syringae* pv. *syringae* on bean pods, lilac and pear plants. *Plant Pathol.* 43(1), pp.528–536

Yin-Yuan, M. & Gross, D., 1991. Expression *in vitro* and during plant pathogenesis of the *syrB* gene required for syringomycin production by *Pseudomonas syringae* pv. *syringae*. *Mol Plant Microbe In*. 4(1), pp.28–36

Young, J.M. *et al.*, 1978. A proposed nomenclature and classification for plant pathogenic bacteria. *New Zeal J Agr Res*. 21(1), pp.153–177

Young, J.M., 1987. Orchard management and bacterial diseases of stone fruit. *New Zeal J Exp Agr*. 15(2), pp.257–266

Young, J.M., 1991. Pathogenicity and identification of the lilac pathogen,

Pseudomonas syringae pv. *syringae* van Hall 1902. *Ann Appl Biol.* 118(2), pp.283–298

Zhang, J., Li, W. *et al.*, 2010. Receptor-like cytoplasmic kinases integrate signaling from multiple plant immune receptors and are targeted by a *Pseudomonas syringae* effector. *Cell Host Microbe*. 7(4), pp.290–301

Zhang, Z. *et al.*, 2012. Disruption of PAMP-induced MAP kinase cascade by a *Pseudomonas syringae* effector activates plant immunity mediated by the NB-LRR protein SUMM2. *Cell Host Microbe*. 11(3), pp.253–263

Zhou, J. *et al.*, 2014. *Pseudomonas syringae* effector HopF2 suppresses *Arabidopsis* immunity by targeting BAK1. *Plant J*. 77(2), pp.235–245

Zhou, Y. *et al.*, 2011. PHAST: A Fast Phage Search Tool. *Nucleic Acids Res*. 39(S2), pp.1–6

Zvereva, A.S. & Pooggin, M.M., 2012. Silencing and innate immunity in plant defense against viral and non-viral pathogens. *Viruses*. 4(11), pp.2578–2597

Thesis appendix

REML criterion at convergence: 152.9818

Groups	Name	Std.Dev.
tree	(Intercept)	2.70E-12
Residual	4.72E-01	

ANOVA

	Sum Sq	Mean Sq	NumDf	DenDf	F.value	Pr(>F)	
strain	122.8	5.8476	21	88	26.255	<2.20E-16	***

Groups

strain	lsmean	SE	df	lower.CL	upper.CL	.group
Pph	1	0.2110579	88	0.5805667	1.419433	1
Control	1	0.2110579	88	0.5805667	1.419433	1
RMA1	1.2	0.2110579	88	0.7805667	1.619433	12
R1-5300	1.2	0.2110579	88	0.7805667	1.619433	12
R1-9326	1.6	0.2110579	88	1.1805667	2.019433	12
R1-9629	1.6	0.2110579	88	1.1805667	2.019433	12
Ps-9643	1.6	0.2110579	88	1.1805667	2.019433	12
Psav	1.8	0.2110579	88	1.3805667	2.219433	12
R1-9657	2.2	0.2110579	88	1.7805667	2.619433	23
R2-5260	2.2	0.2110579	88	1.7805667	2.619433	23
Pss-9293	3	0.2110579	88	2.5805667	3.419433	34
R1-5244	3	0.2110579	88	2.5805667	3.419433	34
R2-leaf	3	0.2110579	88	2.5805667	3.419433	34
R2-sc214	3	0.2110579	88	2.5805667	3.419433	34
R2-5255	3	0.2110579	88	2.5805667	3.419433	34
Pss-9656	3.6	0.2110579	88	3.1805667	4.019433	4
R1-9646	3.6	0.2110579	88	3.1805667	4.019433	4
Pss-9097	3.8	0.2110579	88	3.3805667	4.219433	4
Pss-9630	3.8	0.2110579	88	3.3805667	4.219433	4
Pss-9659	4	0.2110579	88	3.5805667	4.419433	4
Pss-9644	4	0.2110579	88	3.5805667	4.419433	4
Pss-9654	4	0.2110579	88	3.5805667	4.419433	4

Table A1: ANOVA table and Tukey-HSD groups for glasshouse wound inoculations *of all* P. syringae *strains at day 10 (corresponds to Figure 3.2).*

REML criterion at convergence 1147.276

Groups	Name	Std.Dev.
no.:block	(Intercept)	0.4094
block	(Intercept)	0.2575
Residual	1.4777	

ANOVA

	Sum Sq	Mean Sq	NumDf	DenDf	F.value	Pr(>F)	
cv	28.081	9.36	3	54.647	4.2866	0.008676	**
strain	291.283	36.41	8	257.493	16.6742	<2.20E-16	***
cv:strain	51.668	2.153	24	256.549	0.9859	0.485315	

Groups

Cultivar=Napoleon							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	psaq	0.7115409	0.7688193	287.65	-0.80168399	2.224766	1
	control	0.9650962	0.4040532	252.55	0.16935308	1.760839	1
	p9643	1.1651608	0.6300674	286.81	-0.07498177	2.405303	1
	pph	1.1826647	0.546974	283.54	0.10601973	2.25931	1
	p5300	2.0227021	0.4905474	277.92	1.05704164	2.988363	1
	pleaf	2.5388528	0.4901255	278.54	1.57403223	3.503673	12
	p9293	2.7918794	0.6301668	286.91	1.55154319	4.032216	123
	p5244	4.4378216	0.5164664	280.63	3.42118155	5.454462	23
	p9097	4.8299772	0.546639	283.38	3.7539891	5.905965	3
Cultivar=Roundel							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	psaq	0.7465498	0.6299323	286.39	-0.49333446	1.986434	12
	control	0.7793893	0.4191394	250.14	-0.04610281	1.604881	1
	p9643	1.0279468	0.5470303	282.72	-0.04882233	2.104716	123
	pph	1.2035462	0.6300778	286.68	-0.03661915	2.443712	1234
	p5300	1.4205421	0.5471296	282.67	0.34357662	2.497508	1234
	p9293	2.9215907	0.5469634	283	1.84495765	3.998224	234
	p5244	3.0009181	0.5840827	284.81	1.85125165	4.150585	234
	p9097	3.4873141	0.5468634	283.22	2.41088171	4.563747	4
	pleaf	3.5340581	0.6303581	286.41	2.29333618	4.77478	34
Cultivar=Van							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	p5300	0.3958621	0.4898468	278.68	-0.56840755	1.360132	1
	psaq	0.7752972	0.5832431	285.19	-0.37271031	1.923305	12
	control	0.7944169	0.3706285	235.1	0.06423961	1.524594	1
	pph	0.8761393	0.5463041	283.54	-0.19918711	1.951466	12
	p9643	1.0287082	0.4899536	278.64	0.06422751	1.993189	12
	p9293	1.3070645	0.5463887	283.28	0.23156747	2.382562	12
	pleaf	1.9763537	0.5157726	281.13	0.96108743	2.99162	12
	p5244	2.7353955	0.5158945	281.12	1.71988886	3.750902	2
	p9097	3.0238906	0.5464519	283.22	1.94826816	4.099513	2
Cultivar=Mglory							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	pph	0.6249512	0.4897526	278.69	-0.33913306	1.589035	1
	p9643	0.8503081	0.4898971	278.65	-0.11406101	1.814677	1
	control	1.0259437	0.3706886	234.76	0.29564252	1.756245	1
	psaq	1.0998576	0.4898085	278.68	0.13566323	2.064052	12
	p9293	1.3988725	0.5157931	281.3	0.38356843	2.414177	12
	p5300	1.5680127	0.546373	283.19	0.49254506	2.64348	12
	pleaf	1.805174	0.5157572	280.98	0.78993558	2.820412	12
	p5244	2.159184	0.5463503	283.15	1.08376048	3.234608	12
	p9097	3.0714823	0.5158191	280.99	2.05612227	4.086842	2

Table A2: ANOVA table and Tukey-HSD groups for cherry scar inoculations in the field *(corresponds to Figure 3.3).*

REML criterion at convergence 781.1216

Groups	Name	Std.Dev.
no.:block	(Intercept)	0.2025
block	(Intercept)	0.1348
Residual	0.7584	

ANOVA

	Sum Sq	Mean Sq	NumDf	DenDf	F.value	Pr(>F)	
cv	10.722	3.5741	3	58.35	6.2132	0.0009797	***
strain	107.014	13.3767	8	268.37	23.2541	<2.20E-16	***
cv:strain	19.491	0.8121	24	266.85	1.4118	0.1001379	

Groups

Cultivar=Napoleon							
strain	lsmean	SE	df	lower.CL	upper.CL	.group	
	pph	3.653366	0.2800747	290.71	3.102135	4.204597	1
	control	3.885997	0.2010153	246.7	3.490071	4.281922	1
	psaq	3.985179	0.2800069	290.47	3.434079	4.536279	12
	p9643	4.022781	0.25105	284.9	3.528633	4.516929	12
	p5300	4.07178	0.2510955	284.91	3.577542	4.566018	12
	p9293	4.584441	0.2510556	284.9	4.090281	5.0786	12
	p9097	4.937681	0.2511169	284.9	4.443401	5.431961	23
	pleaf	5.031797	0.2989259	292.59	4.443479	5.620115	23
	p5244	5.826491	0.2643723	287.95	5.306144	6.346839	3
Cultivar=Roundel							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	pph	3.91928	0.3227659	294.46	3.28406	4.5545	1
	p9643	3.96481	0.2992863	292.56	3.375783	4.553837	1
	control	3.986056	0.2307565	270.95	3.531752	4.44036	1
	p5300	4.319487	0.3227346	294.37	3.684327	4.954646	12
	psaq	4.334718	0.2992081	292.89	3.745847	4.923588	12
	p9293	4.747526	0.2992498	292.49	4.15857	5.336482	123
	p9097	5.212044	0.2645068	287.95	4.691432	5.732656	23
	pleaf	5.467978	0.2802705	290.41	4.916359	6.019597	23
	p5244	5.826919	0.2993383	292.54	5.237789	6.416048	3
Cultivar=Van							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	control	3.842611	0.2142396	259.49	3.420741	4.26448	1
	p5300	3.847813	0.2645289	287.83	3.327157	4.36847	12
	psaq	3.955141	0.2991172	292.64	3.366448	4.543835	12
	p9643	4.009597	0.2800894	290.54	3.458335	4.560858	12
	pph	4.138014	0.2800922	290.37	3.586746	4.689283	12
	pleaf	4.444619	0.280047	290.61	3.893441	4.995796	12
	p9293	5.003258	0.2801659	290.44	4.451845	5.554671	23
	p9097	5.619921	0.2645137	287.59	5.099292	6.140549	3
	p5244	6.1149	0.2645705	287.5	5.594159	6.635641	3
Cultivar=Mglory							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	p9643	3.741044	0.2510521	284.89	3.246892	4.235196	1
	control	3.747158	0.1851152	231.18	3.38243	4.111887	1
	psaq	3.779704	0.2642308	287.69	3.259633	4.299775	1
	pph	3.96303	0.2510246	284.88	3.468931	4.457128	1
	p5300	3.997156	0.2510258	284.88	3.503056	4.491257	1
	p9293	4.19852	0.2510654	284.89	3.704342	4.692699	1
	pleaf	4.338137	0.2642709	287.74	3.817987	4.858286	1
	p5244	4.475735	0.2642565	287.71	3.955614	4.995857	1
	p9097	4.618541	0.2509618	284.82	4.124566	5.112516	1

Table A3: ANOVA table and Tukey-HSD groups for cherry wound inoculations in the field *(corresponds to Figure 3.3).*

REML criterion at convergence: 430.7719

Groups	Name	Std.Dev.
no.:block	(Intercept)	0.2346
block	(Intercept)	0.4343
Residual	1.2296	

ANOVA

	Sum Sq	Mean Sq	NumDf	DenDf	F.value	Pr(>F)	
cv	2.652	2.6521	1	19.802	1.7542	0.20043	
strain	38.055	4.7568	8	105.166	3.1464	0.003123	**
cv:strain	15.815	1.9769	8	105.144	1.3076	0.247714	

Cultivar=Mseed

strain	lsmean	SE	df	lower.CL	upper.CL	.group
control	0.1440531	0.4086608	84.93	-0.6684831	0.9565894	1
pph	0.2073209	0.5829317	114.33	-0.9474271	1.3620689	1
p9293	0.6135994	0.5344373	111.24	-0.4453981	1.6725969	1
p9643	0.7677172	0.5343924	111.3	-0.2911852	1.8266197	1
pleaf	0.9588263	0.5344488	111.24	-0.1001948	2.0178473	1
p5244	1.2158417	0.534445	111.24	0.1568286	2.2748548	1
p9097	1.2663599	0.5829711	114.32	0.111533	2.4211867	1
p5300	1.3291135	0.582869	114.38	0.1744958	2.4837313	1
psaq	1.5073158	0.5828574	114.39	0.3527219	2.6619097	1

Cultivar=Victoria

strain	lsmean	SE	df	lower.CL	upper.CL	.group
pleaf	0.4968944	0.4648764	105.68	-0.4248007	1.4185896	1
p9643	0.536564	0.4401931	100.8	-0.3366822	1.4098102	1
control	0.5719486	0.3532792	71.31	-0.1324169	1.2763142	1
psaq	0.7673459	0.4951211	109.3	-0.2139377	1.7486295	1
pph	0.8567131	0.4648958	105.68	-0.0650206	1.7784467	1
p5244	1.2096196	0.439834	101.78	0.3371879	2.0820512	12
p9293	1.6550116	0.4651781	105.09	0.7326591	2.5773642	12
p5300	1.8170385	0.4401987	100.73	0.9437749	2.690302	12
p9097	3.126846	0.4650973	105.13	2.2046569	4.049035	2

Table A4: ANOVA table and Tukey-HSD groups for plum scar inoculations in the field *(corresponds to Figure 3.4).*

REML criterion at convergence: 336.8069

Groups	Name	Std.Dev.
no.:block	(Intercept)	0
block	(Intercept)	0.267
Residual	0.6695	

ANOVA

	Sum Sq	Mean Sq	NumDf	DenDf	F.value	Pr(>F)	
cv	0.3911	0.3911	1	137.49	0.8726	0.3519	
strain	18.5883	2.3235	8	132.48	5.184	1.21E-05	***
cv:strain	2.56	0.32	8	132.51	0.7139	0.6789	

Groups

Cultivar=Mseed

strain	lsmean	SE	df	lower.CL	upper.CL	.group
p9643	3.500483	0.2279265	109.5	3.048763	3.952202	1
psaq	3.614827	0.2684932	126.73	3.083517	4.146137	12
control	3.620368	0.1930004	80.75	3.236339	4.004396	1
p9293	3.844657	0.2523697	121.26	3.345036	4.344278	12
pleaf	3.927855	0.2684408	126.82	3.396652	4.459059	12
pph	4.091558	0.2523846	121.24	3.591905	4.59121	12
p5244	4.31871	0.2523866	121.23	3.819054	4.818366	12
p5300	4.420649	0.2523866	121.23	3.920993	4.920306	12
p9097	4.590057	0.268495	126.73	4.058743	5.121371	2

Cultivar=Victoria

strain	lsmean	SE	df	lower.CL	upper.CL	.group
control	3.577607	0.1880049	76.15	3.203174	3.95204	1
psaq	3.796599	0.2684672	126.77	3.265341	4.327856	12
pph	3.840107	0.2523378	121.32	3.340551	4.339662	12
p9643	3.913038	0.2523169	121.36	3.413525	4.412551	12
p5244	3.975253	0.2523884	121.23	3.475593	4.474913	12
p9293	4.269892	0.2523821	121.24	3.770245	4.769539	12
pleaf	4.294374	0.2685072	126.71	3.763035	4.825713	12
p5300	4.532977	0.2391008	115.44	4.059383	5.006571	2
p9097	4.674532	0.2279265	109.5	4.222812	5.126252	2

Table A5: ANOVA table and Tukey-HSD groups for plum wound inoculations in the field *(corresponds to Figure 3.4).*

ANOVA

	Df	Sum Sq	Mean Sq	F	value	Pr(>F)
strain	8	635.8	79.47	31.755	<2.00E-16	***
host	1	0.5	0.52	0.207	0.649	
block	9	21.3	2.37	0.945	0.485	
strain:host	8	209	26.12	10.437	1.64E-13	***
strain:host:cv	36	232.1	6.45	2.576	3.37E-06	***
Residuals	474	1186.2	2.5			

Groups

Cultivar=Mglory							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	Control	-0.11493957	0.4672835	312	-1.034364916	0.8044858	1
	Psav	0.18066157	0.4672835	312	-0.738763783	1.1000869	1
	R1-5300	0.19871539	0.4672835	312	-0.720709955	1.1181407	12
	Pph	0.4055746	0.4672835	312	-0.513850748	1.325	12
	Pss-9293	0.92468447	0.4672835	312	0.005259124	1.8441098	12
	RMA1	1.30666391	0.4932841	312	0.336079776	2.277248	12
	R1-5244	2.25995511	0.4672835	312	1.340529762	3.1793805	23
	Pss-9097	3.43771949	0.4672835	312	2.518294137	4.3571448	3
	R2-5255	3.66592157	0.4672835	312	2.746496222	4.5853469	3
Cultivar=Napoleon							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	Control	0.02432314	0.4672835	312	-0.895102213	0.9437485	1
	Pph	0.92055407	0.4672835	312	0.00112872	1.8399794	1
	Psav	0.95043792	0.4672835	312	0.031012568	1.8698633	1
	RMA1	1.26796721	0.4672835	312	0.348541857	2.1873926	1
	R1-5300	1.29231273	0.4672835	312	0.372887381	2.2117381	1
	Pss-9293	1.49090147	0.4672835	312	0.571476119	2.4103268	12
	R1-5244	3.39300954	0.4672835	312	2.473584187	4.3124349	23
	R2-5255	3.93432366	0.4672835	312	3.014898306	4.853749	3
	Pss-9097	4.90317771	0.4672835	312	3.983752361	5.8226031	3
Cultivar=Roundel							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	Pph	-0.08248068	0.4672835	312	-1.001906032	0.8369447	1
	Psav	0.52677457	0.4672835	312	-0.392650778	1.4461999	1
	Pss-9293	0.72222511	0.4672835	312	-0.197200238	1.6416505	1
	Control	0.79154722	0.4672835	312	-0.127878131	1.7109726	1
	R1-5300	0.81030135	0.4932637	312	-0.160242618	1.7808453	1
	RMA1	1.22266027	0.4932841	312	0.25207614	2.1932444	12
	R1-5244	3.07172907	0.4672835	312	2.152303725	3.9911544	23
	Pss-9097	3.90424766	0.4672835	312	2.984822312	4.823673	3
	R2-5255	5.10557961	0.4672835	312	4.186154264	6.025005	3
Cultivar=Van							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	Control	0.37151977	0.4672835	312	-0.54790558	1.2909451	1
	Pph	0.46987059	0.4672835	312	-0.449554756	1.3892959	1
	Psav	0.767414	0.4672835	312	-0.152011352	1.6868393	1
	R1-5300	1.51796279	0.4672835	312	0.598537444	2.4373881	12
	Pss-9293	2.06797603	0.4672835	312	1.14855068	2.9874014	123
	RMA1	2.20906898	0.4672835	312	1.289643632	3.1284943	123
	R2-5255	3.07573553	0.4672835	312	2.156310183	3.9951609	234
	R1-5244	4.02845881	0.4672835	312	3.109033457	4.9478842	34
	Pss-9097	4.36985694	0.4672835	312	3.450431593	5.2892823	4
Cultivar=Mseed							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	R1-5244	0.394205	0.5518194	153	-0.69596402	1.484374	1
	Control	0.6479907	0.5518194	153	-0.44217836	1.73816	1
	Pph	0.9034142	0.5518194	153	-0.18675485	1.993583	1
	Pss-9293	0.9490638	0.5518194	153	-0.14110528	2.039233	1
	R2-5255	1.0126181	0.5518194	153	-0.07755099	2.102787	1
	RMA1	1.0253719	0.5518194	153	-0.06479715	2.115541	1
	R1-5300	1.2889424	0.5518194	153	0.19877334	2.379111	1
	Psav	1.5139238	0.5518194	153	0.4237547	2.604093	1
	Pss-9097	1.9707152	0.5518194	153	0.88054614	3.060884	1
Cultivar=Victoria							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	Psav	0.3408936	0.5518194	153	-0.74927547	1.431063	1
	R1-5244	1.3643079	0.5518194	153	0.27413881	2.454477	12
	Control	1.4999825	0.5518194	153	0.40981346	2.590152	12
	Pph	1.9237886	0.5518194	153	0.83361949	3.013958	123
	Pss-9293	2.5322749	0.5518194	153	1.44210583	3.622444	123
	R2-5255	2.6431936	0.5518194	153	1.55302449	3.733363	123
	R1-5300	3.0659037	0.5518194	153	1.9757346	4.156073	23
	Pss-9097	4.1548425	0.5518194	153	3.0646734	5.245012	3
	RMA1	4.3579526	0.5518194	153	3.26778351	5.448122	3

Table A6: ANOVA table and Tukey-HSD groups for cutshoot assays of cherry and plum *(corresponds to Figure 3.5).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	21	71.25	3.393	32.827	<2.00E-16	***
rep	4	1.98	0.495	4.788	0.00158	**
Residuals	84	8.68	0.103			

Groups

trt	means	M
Pss-9097	3.710626	a
Pss-9630	3.5695	ab
Pss-9644	3.172356	abc
Pss-9654	3.035987	abc
Pss-9293	3.019075	abc
Pss-9659	2.927907	bc
Pss-9656	2.65432	cd
R2-5255	1.901356	de
R2-sc214	1.854288	e
R1-5244	1.847817	e
Ps-9643	1.728745	ef
R1-9646	1.689723	ef
R2-leaf	1.598033	ef
R2-5260	1.558717	ef
R1-5300	1.555996	ef
R1-9326	1.550532	ef
R1-9657	1.49933	ef
R1-9629	1.418988	ef
RMA1	1.314859	ef
Psavellanae	1.229821	ef
Pph	1.199682	ef
Control	1.067207	f

Table A7: ANOVA table and Tukey-HSD groups for cherry fruitlet assays of all* P. syringae *strains *(corresponds to Figure 3.6).*

ANOVA

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr(>F)	
cultivar	1.986	0.662	3	107.32	6.25	0.0005985	***
strain	105.345	35.115	3	106.33	331.26	<2.20E-16	***
cultivar:strain	2.38	0.264	9	106.49	2.49	0.0125354	*

Groups

Cultivar=Mglory							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	Control	0.9676333	0.1089471	112.94	0.7517882	1.183478	1
	R2-5255	1.4843107	0.1155512	112.97	1.2553824	1.713239	2
	R1-5244	1.6406268	0.123526	112.99	1.3958991	1.885354	2
	Pss-9097	3.1194942	0.1033594	112.91	2.9147189	3.32427	3
Cultivar=Napoleon							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	Control	0.8629926	0.1155524	112.96	0.6340616	1.091924	1
	R1-5244	1.6978307	0.1089471	112.94	1.4819856	1.913676	2
	R2-5255	1.7083639	0.1155518	112.97	1.4794341	1.937294	2
	Pss-9097	3.0475417	0.1155516	112.97	2.8186125	3.276471	3
Cultivar=Roundel							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	Control	0.8372866	0.1461479	113	0.5477413	1.126832	1
	R2-5255	1.5721559	0.1334187	113	1.3078292	1.836483	2
	R1-5244	1.8314291	0.1089466	112.95	1.6155852	2.047273	2
	Pss-9097	3.6017906	0.115551	112.97	3.3728626	3.830719	3
Cultivar=Van							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	Control	0.8254119	0.1461478	113	0.5358667	1.114957	1
	R2-5255	1.8331431	0.1033594	112.91	1.6283678	2.037918	2
	R1-5244	2.0258618	0.1089466	112.95	1.8100178	2.241706	2
	Pss-9097	3.7853686	0.1033594	112.91	3.5805933	3.990144	3

Table A8: ANOVA table and Tukey-HSD groups for cherry fruitlet assays of reference* P. syringae *strains on different cultivars *(corresponds to Figure 3.7).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	6	1198.1	199.69	97.127	<2.00E-16	***
leaf	2	22.1	11.06	5.381	0.00779	**
leaf:rep	6	1.1	0.18	0.089	0.9971	
Residuals	48	98.7	2.06			

Groups

strain	lsmean	SE	df	lower.CL	upper.CL	.group
pph	14.47373	0.477956	48	13.51274	15.43473	1
psaq	14.80357	0.477956	48	13.84257	15.76456	1
psav	16.1298	0.477956	48	15.1688	17.09079	1
5300	19.15372	0.477956	48	18.19273	20.11472	2
9097	23.31078	0.477956	48	22.34979	24.27178	3
5244	24.91811	0.477956	48	23.95711	25.8791	3
leaf	25.04114	0.477956	48	24.08015	26.00214	3

Table A8: ANOVA table and Tukey-HSD groups for leaf day 10 population counts of reference* P. syringae *strains on cherry *(corresponds to Figure 3.9A).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	6	1371.9	228.65	67.711	<2e-16	***
leaf	2	19.3	9.65	2.857	0.0672	.
leaf:rep	6	0.4	0.07	0.021	1	
Residuals	48	162.1	3.38			

Groups

strain	lsmean	SE	df	lower.CL	upper.CL	.group
pph	12.06416	0.6125429	48	10.83256	13.29576	1
psav	14.64228	0.6125429	48	13.41068	15.87388	1
psaq	21.64243	0.6125429	48	20.41083	22.87403	2
5244	22.51337	0.6125429	48	21.28177	23.74497	23
leaf	22.84738	0.6125429	48	21.61578	24.07898	23
5300	24.41012	0.6125429	48	23.17852	25.64172	3
9097	25.06238	0.6125429	48	23.83078	26.29398	3

Table A9: ANOVA table and Tukey-HSD groups for leaf day 10 population counts of reference* P. syringae *strains on plum *(corresponds to Figure 3.9B).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	20	2465.6	123.28	64.068	<2.00E-16	***
leaf	2	31.8	15.9	8.261	0.000385	***
leaf:rep	6	0.2	0.03	0.013	0.99999	
Residuals	160	307.9	1.92			

Groups

strain	lsmean	SE	df	lower.CL	upper.CL	.group
psaq	14.35272	0.4623813	160	13.43956	15.26588	1
9643	17.8635	0.4623813	160	16.95034	18.77666	2
pph	18.39166	0.4623813	160	17.4785	19.30482	2
psav	21.95057	0.4623813	160	21.03741	22.86372	3
5300	22.06974	0.4623813	160	21.15658	22.98289	3
9629	22.17053	0.4623813	160	21.25737	23.08368	3
9326	22.31059	0.4623813	160	21.39743	23.22374	3
9657	23.11253	0.4623813	160	22.19938	24.02569	3
5260	25.74769	0.4623813	160	24.83453	26.66085	4
9293	26.05132	0.4623813	160	25.13816	26.96448	4
5255	26.16608	0.4623813	160	25.25293	27.07924	4
9644	26.21295	0.4623813	160	25.29979	27.12611	4
9654	26.24051	0.4623813	160	25.32735	27.15367	4
9646	26.34192	0.4623813	160	25.42876	27.25507	4
9097	26.87423	0.4623813	160	25.96107	27.78738	4
leaf	26.9997	0.4623813	160	26.08655	27.91286	4
9659	27.04727	0.4623813	160	26.13411	27.96042	4
9656	27.25017	0.4623813	160	26.33702	28.16333	4
9630	27.27298	0.4623813	160	26.35982	28.18613	4
5244	27.38274	0.4623813	160	26.46959	28.2959	4
sc214	27.39496	0.4623813	160	26.4818	28.30812	4

Table A10: ANOVA table and Tukey-HSD groups for leaf day 10 population counts of all* P. syringae *strains on cherry *(corresponds to Figure 3.10).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	8	804	100.5	56.732	<2.00E-16	***
leaf	2	63.6	31.81	17.957	6.45E-07	***
leaf:rep	6	0.2	0.03	0.018	1	
Residuals	64	113.4	1.77			

Groups

trt	means	M
R2-leaf	25.30739	a
Pss-9097	23.89098	ab
R1-5244	23.86917	ab
Pss-9293	22.50342	b
R1-5300	19.35157	c
Ps-9643	18.8921	c
psavellanae	18.75037	c
pph	18.56295	c
RMA1	15.11659	d

Table A11: ANOVA table and Tukey-HSD groups for leaf day 10 population counts of reference* P. syringae *strains on cherry *(corresponds to Figure 3.12B).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	8	323.2	40.4	30.732	<2e-16	***
leaf	2	18.2	9.1	6.923	0.0019	**
leaf:rep	6	0.3	0.05	0.037	0.9998	
Residuals	64	84.1	1.31			

Groups

trt	means	M
Pss-9097	26.63806	a
R1-5300	25.67461	ab
R2-leaf	25.6439	ab
Pss-9293	24.91062	abc
R1-5244	24.70928	bc
RMA1	23.66233	cd
Ps-9643	22.72922	de
pph	21.37888	ef
psavellanae	20.32219	f

Table A12: ANOVA table and Tukey-HSD groups for leaf day 10 population counts of reference* P. syringae *strains on plum *(corresponds to Figure 3.12B).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	8	7.374	0.9217	43.98	1.74E-12	***
leaf	3	0.754	0.2514	11.99	5.38E-05	***
Residuals	24	0.503	0.021			

Groups

trt	means	M
Ps-9643	-2.090725	a
Pss-9097	-2.090725	a
Pss-9293	-2.090725	a
RMA1	-2.090725	a
R1-5300	-2.652572	b
pph	-2.726436	bc
psavellanae	-2.852061	bc
R2-leaf	-3.070678	cd
R1-5244	-3.321928	d

Table A13: ANOVA table and Tukey-HSD groups for slope of symptom score (0-24 hours) of reference* P. syringae *strains on cherry leaves *(corresponds to Figure 3.12A).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	8	4.851	0.6064	9.934	5.36E-06	***
leaf	3	0.385	0.1283	2.101	0.127	
Residuals	24	1.465	0.061			

Groups

trt	means	M
Pss-9293	-2.225867	a
Pss-9097	-2.299731	a
Ps-9643	-2.652572	ab
RMA1	-2.726436	ab
pph	-2.778197	abc
psavellanae	-2.945053	bc
R1-5300	-3.196303	bc
R2-leaf	-3.196303	bc
R1-5244	-3.321928	c

Table A14: ANOVA table and Tukey-HSD groups for slope of symptom score (0-24 hours) of reference* P. syringae *strains on plum leaves *(corresponds to Figure 3.12A).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	2	157.8	78.9	75.418	<2.00E-16	***
cultivar	3	48.32	16.11	15.396	3.92E-08	***
leaf	2	2.81	1.4	1.343	0.266	
strain:cultivar	6	62.86	10.48	10.014	2.06E-08	***
leaf:rep	6	0.26	0.04	0.042	1	
Residuals	88	92.06	1.05			

Groups

Cultivar=Mglory							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	Pss-9097	21.64551	0.3409395	88	20.96797	22.32306	1
	R1-5244	22.20273	0.3409395	88	21.52519	22.88028	1
	R2-leaf	23.94364	0.3409395	88	23.2661	24.62119	2
Cultivar=Napoleon							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	Pss-9097	20.39085	0.3409395	88	19.71331	21.0684	1
	R1-5244	24.41244	0.3409395	88	23.73489	25.08999	2
	R2-leaf	25.15962	0.3409395	88	24.48207	25.83716	2
Cultivar=Roundel							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	Pss-9097	20.25446	0.3409395	88	19.57692	20.93201	1
	R1-5244	20.75198	0.3409395	88	20.07443	21.42952	1
	R2-leaf	23.90887	0.3409395	88	23.23132	24.58642	2
Cultivar=Van							
	strain	lsmean	SE	df	lower.CL	upper.CL	.group
	Pss-9097	22.63402	0.3409395	88	21.95648	23.31157	1
	R1-5244	23.26184	0.3409395	88	22.58429	23.93938	1
	R2-leaf	23.75344	0.3409395	88	23.0759	24.43099	1

Table A15: ANOVA table and Tukey-HSD groups for leaf day 10 population counts of reference* P. syringae *strains on different cherry cultivars *(corresponds to Figure 3.14).*

A: ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)
strain	6	21	3.499	1.401	0.234
leaf	2	1.57	0.783	0.313	0.732
leaf:rep	6	5.44	0.907	0.363	0.899
Residuals	48	119.89	2.498		

B: ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	6	117.1	19.516	13.524	6.50E-09	***
leaf	2	42.97	21.486	14.889	9.31E-06	***
leaf:rep	6	5.3	0.884	0.612	0.719	
Residuals	48	69.27	1.443			

B: Groups

trt	means	M
9097	2.8036494	a
pph	2.4192321	a
psaq	1.9930192	a
psav	1.9930192	a
5300	1.3096882	ab
5244	-0.2240372	bc
leaf	-1.1927137	c

Table A16: ANOVA table and Tukey-HSD groups for leaf day 0 population counts of reference* P. syringae *strains on cherry (A) and plum (B) *(corresponds to Figure S3.3).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	20	869.5	43.48	4.854	4.03E-09	***
leaf	2	24.9	12.46	1.391	0.252	
leaf:rep	6	16.6	2.77	0.31	0.931	
Residuals	160	1433	8.96			

Groups

trt	means	M
Pss-9097	10.30274	a
R1-5300	10.053562	a
R1-9646	9.95293	a
R1-9657	9.682486	a
R1-9326	9.643946	a
R1-9629	9.627344	a
Pss-9644	9.130461	a
R1-5244	9.118061	a
Pss-9659	9.099709	a
psavellanae	8.903034	a
Pss-9654	8.723586	a
Pss-9630	8.591124	a
Pss-9656	8.112656	a
pph	8.049394	a
RMA1	7.506392	ab
Pss-9293	7.49977	ab
R2-5255	6.878981	ab
R2-sc214	6.556108	ab
Ps-9643	5.337621	ab
R2-leaf	2.613482	b
R2-5260	2.548513	b

Table A17: ANOVA table and Tukey-HSD groups for leaf day 0 population counts of all* P. syringae *strains on cherry *(corresponds to Figure S3.4).*

A: ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	8	50.4	6.298	0.316	0.957	
leaf	2	4.1	2.054	0.103	0.902	
leaf:rep	6	96.9	16.143	0.811	0.565	
Residuals	64	1274.4	19.912			

B: ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	8	169.5	21.18	1.209	0.308	
leaf	2	8.6	4.3	0.245	0.7833	
leaf:rep	6	269.6	44.93	2.565	0.0274	*
Residuals	64	1121.1	17.52			

Table A18: ANOVA tables for leaf day 0 population counts of reference* P. syringae *strains on cherry (A) and plum (B) *(corresponds to Figure S3.5).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)
strain	2	76.5	38.26	1.852	0.163
cultivar	3	92.6	30.86	1.493	0.222
leaf	2	21.5	10.75	0.52	0.596
strain:cultivar	6	74.6	12.43	0.601	0.728
leaf:rep	6	163.8	27.3	1.321	0.256
Residuals	88	1818.2	20.66		

Table A19: ANOVA tables for leaf day 0 population counts of reference* P. syringae *strains on different cherry cultivars *(corresponds to Figure S3.6).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	11	2311.5	210.14	73.765	<2e-16	***
leaf	2	20.1	10.06	3.531	0.0335	*
leaf:rep	6	0.2	0.04	0.014	1	
Residuals	88	250.7	2.85			

Groups

hopAO1	24.52347	a
WT	24.46874	a
avrRps4	24.28181	a
hopG1	23.80866	a
pBBR1MCS5	23.78394	a
hopAA1	23.70971	a
avrPto1	22.71873	a
avrA	22.67135	a
hopAW1	19.56831	b
hopC1	13.59922	c
hopAB1	13.44748	c
ΔhrpA	12.05762	c

Table A20: ANOVA table and Tukey-HSD groups for leaf day 10 population counts of* Psm *R2 leaf transconjugants on cherry
(corresponds to Figures 5.4 and 5.5).

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	11	2509.8	228.16	141.438	<2e-16	***
leaf	2	3.5	1.73	1.071	0.347	
leaf:rep	6	0.9	0.16	0.097	0.997	
Residuals	88	142	1.61			

Groups

WT	24.86676	a
hopAW1	24.72974	ab
pBBR1MCS5	24.31672	ab
hopAO1	23.43526	abc
hopAA1	23.2862	abc
avrPto1	22.92216	abcd
hopG1	22.72825	bcd
avrRps4	21.83345	cd
avrA	20.98589	d
hopAB1	17.39312	e
hopC1	11.53217	f
ΔhrpA	10.08799	f

Table A21: ANOVA table and Tukey-HSD groups for leaf day 10 population counts of* Psm *R1 5244 transconjugants on cherry
(corresponds to Figure 5.5).

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	10	4.83E+15	4.83E+14	17.079	4.33E-16	***
leaf	2	2.01E+14	1.01E+14	3.55	0.0333	*
leaf:rep	6	3.54E+13	5.90E+12	0.208	0.9732	
Residuals	80	2.26E+15	2.83E+13			

Groups

hopAW1	22333333	a
avrPto1	20333333	ab
hopAO1	20111111	ab
avrRps4	19666667	ab
WT	19666667	ab
hopAA1	18444444	abc
pBBR1MCS5	16555556	abc
avrA	13088889	bc
hopG1	10733333	c
hopC1	1877778	d
hopAB1	1700000	d

Table A22: ANOVA table and Tukey-HSD groups for leaf day 10 population counts of* Pss *9644 transconjugants on cherry
(corresponds to Figure 5.5).

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	4	1931.4	482.9	112.066	<2e-16	***
exp	1	800.6	800.6	185.803	<2e-16	***
exp:leaf	4	10.9	2.7	0.63	0.641	
exp:leaf:rep	12	2.7	0.2	0.053	1	
Residuals	248	1068.5	4.3			

Groups

trt	means	M
pBBR1MCS5	25.15762	a
hopAB2	20.96854	b
hopAB3	20.69406	b
hopC1	17.87883	c
hopAB1	17.8644	c

Table A23: ANOVA table and Tukey-HSD groups for leaf day 10 population counts of* Psm *R2 leaf transconjugants expressing candidate* avr *genes on cherry *(corresponds to Figure 5.6).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	4	806.3	201.6	97.975	<2e-16	***
exp	1	388.1	388.1	188.622	<2e-16	***
exp:leaf	4	21.9	5.5	2.664	0.0397	*
exp:leaf:rep	12	1.3	0.1	0.055	1	
Residuals	68	139.9	2.1			

Groups

trt	means	M
pBBR1MCS5	25.53471	a
hopAB2	20.18772	b
hopAB3	20.02443	b
hopAB1	17.48129	c
hopC1	17.20359	c

Table A24: ANOVA table and Tukey-HSD groups for leaf day 10 population counts of* Psm *R1 5244 transconjugants expressing candidate* avr *genes on cherry *(corresponds to Figure 5.6).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	4	396.2	99.06	56.055	<2.00E-16	***
exp	1	75	75	42.438	1.05E-08	***
exp:leaf	4	20.8	5.19	2.937	0.0266	*
exp:leaf:rep	12	0.4	0.04	0.021	1	
Residuals	68	120.2	1.77			

Groups

trt	means	M
pBBR1MCS5	24.74083	a
hopAB2	23.65924	ab
hopAB3	22.43657	b
hopAB1	19.66143	c
hopC1	19.53238	c

Table A25: ANOVA table and Tukey-HSD groups for leaf day 10 population counts of* Pss *9644 transconjugants expressing candidate* avr *genes on cherry *(corresponds to Figure 5.6).*

Global ANOVA symptom score at different concentrations

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	9	9.666	1.074	61.115	<2e-16	***
conc	3	5.256	1.7519	99.692	<2e-16	***
leaf	3	0.022	0.0073	0.415	0.742	
strain:conc	27	4.164	0.1542	8.777	<2e-16	***
Residuals	117	2.056	0.0176			

Table A26: ANOVA table for slope of symptom score development (0-24 hours) for different transconjugant strains at varying concentrations on cherry leaves *(corresponds to Figure 5.8).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	9	0.00625	0.0006944	7.04E+30	<2e-16	***
leaf	3	0	0	4.92E-01	0.691	
Residuals	27	0	0			

Groups

trt	means	M
R2-leaf(hopC1)	0.04166667	a
R1-5244(hopAB1)	0	b
R1-5244(hopC1)	0	b
R1-5244(hrpA-)	0	b
R1-5244(pbbr)	0	b
R1-5300	0	b
R2-leaf(hopAB1)	0	b
R2-leaf(hrpA-)	0	b
R2-leaf(pbbr)	0	b
RMA1	0	b

Table A27: ANOVA table and Tukey-HSD groups for slope of symptom score development (0-24 hours) for different transconjugant strains inoculated at 2.5x10^7 CFU/ml on cherry leaves *(corresponds to Figure 5.8).*

Concentration 2

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	9	0.011979	0.001331	1.59E+01	1.21E-08	***
leaf	3	0.000347	0.0001157	1.39E+00	0.269	
Residuals	27	0.002257	0.0000836			

Groups

trt	means	M
R2-leaf(hopC1)	0.05208333	a
R1-5244(hopC1)	0.03125	a
R1-5244(hopAB1)	0	b
R1-5244(hrpA-)	0	b
R1-5244(pbbr)	0	b
R1-5300	0	b
R2-leaf(hopAB1)	0	b
R2-leaf(hrpA-)	0	b
R2-leaf(pbbr)	0	b
RMA1	0	b

Table A28: ANOVA table and Tukey-HSD groups for slope of symptom score development (0-24 hours) for different transconjugant strains inoculated at 5.0x10^7 on cherry leaves *(corresponds to Figure 5.8).*

Concentration 3

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	9	0.0365	0.004056	1.63E+01	9.49E-09	***
leaf	3	0.00152	0.000506	2.03E+00	0.133	
Residuals	27	0.00673	0.000249			

Groups

trt	means	M
R2-leaf(hopC1)	0.10416667	a
R1-5244(hopC1)	0.04166667	b
R2-leaf(hopAB1)	0.03125	bc
R1-5244(hopAB1)	0.02083333	bc
R1-5300	0.02083333	bc
RMA1	0.02083333	bc
R1-5244(hrpA-)	0	c
R1-5244(pbbr)	0	c
R2-leaf(hrpA-)	0	c
R2-leaf(pbbr)	0	c

Table A29: ANOVA table and Tukey-HSD groups for slope of symptom score development (0-24 hours) for different transconjugant strains inoculated at 1.0x10^8 on cherry leaves *(corresponds to Figure 5.8).*

Concentration 4

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	9	0.07708	0.008565	3.51E+01	1.29E-12	***
leaf	3	0.00382	0.001273	5.21E+00	0.00572	**
Residuals	27	0.0066	0.000244			

Groups

trt	means	M
R2-leaf(hopC1)	0.125	a
RMA1	0.09375	ab
R1-5244(hopAB1)	0.07291667	b
R1-5244(hopC1)	0.07291667	b
R1-5300	0.07291667	b
R2-leaf(hopAB1)	0.0625	b
R1-5244(hrpA-)	0	c
R1-5244(pbbr)	0	c
R2-leaf(hrpA-)	0	c
R2-leaf(pbbr)	0	c

Table A30: ANOVA table and Tukey-HSD groups for slope of symptom score development (0-24 hours) for different transconjugant strains inoculated at 2.0x10^8 on cherry leaves *(corresponds to Figure 5.8).*

Global ANOVA candidates avr expression in different pathogens

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
clade	2	0.06512	0.03256	66.011	2.52E-11	***
leaf	2	0.00239	0.0012	2.425	0.107	
clade:strain	12	0.04352	0.00363	7.352	7.25E-06	***
Residuals	28	0.01381	0.00049			

Table A31: ANOVA table for slope of symptom score development (0-24 hours) for transconjugant strains expressing candidate avr genes on cherry leaves *(corresponds to Figure 5.9).*

Pss

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	4	0.017361	0.00434	25	0.000142	***
leaf	2	0.000926	0.000463	2.667	0.1296	
Residuals	8	0.001389	0.000174			

Groups

trt	means	M
hopAB2	0.16666667	a
hopC1	0.16666667	a
pBBR1MCS5	0.16666667	a
hopAB	0.09722222	b
hopAB3	0.09722222	b

Table A32: ANOVA table and Tukey-HSD groups for slope of symptom score development (0-24 hours) for different **Pss** ***9644 transconjugants expressing candidate*** **avr** ***genes*** *(corresponds to Figure 5.9).*

Psm R1

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	4	0.013426	0.003356	8.286	0.00604	**
leaf	2	0.000231	0.000116	0.286	0.75883	
Residuals	8	0.003241	0.000405			

Groups

trt	means	M
hopC1	0.09722222	a
hopAB	0.08333333	a
hopAB2	0.06944444	ab
hopAB3	0.04166667	ab
pBBR1MCS5	0.01388889	b

Table A33: ANOVA table and Tukey-HSD groups for slope of symptom score development (0-24 hours) for different **Psm** ***R1 5244 transconjugants expressing candidate*** **avr** ***genes*** *(corresponds to Figure 5.9).*

Psm **R2**

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	4	0.012731	0.003183	6.471	0.0126	*
leaf	2	0.006481	0.003241	6.588	0.0204	*
Residuals	8	0.003935	0.000492			

Groups

trt	means	M
hopC1	0.08333333	a
hopAB2	0.06944444	a
hopAB	0.06944444	a
hopAB3	0.05555556	ab
pBBR1MCS5	0	b

Table A34: ANOVA table and Tukey-HSD groups for slope of symptom score development (0-24 hours) for different **Psm** ***R2 leaf transconjugants expressing candidate*** **avr** ***genes*** *(corresponds to Figure 5.9).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)
genotype	1	0.55	0.5477	0.188	0.667
leaf	2	3.04	1.5188	0.521	0.598
leaf:rep	6	0.28	0.0473	0.016	1
Residuals	35	101.99	2.914		

Table A35: ANOVA table for leaf day 10 population counts of* Psm *R1 5300* hopAB1 *deletion mutants *(corresponds to Figure 5.10A).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)
genotype	1	6.00E+13	6.00E+13	2.643	0.122
leaf	2	2.68E+13	1.34E+13	0.59	0.565
leaf:rep	6	8.16E+12	1.36E+12	0.06	0.999
Residuals	17	3.86E+14	2.27E+13		

Table A36: ANOVA table for leaf day 10 population counts of* Psm *R1 9657* hopAB1 *deletion mutants *(corresponds to Figure 5.10B).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)
genotype	1	1.48E+14	1.48E+14	4.258	0.0547
leaf	2	1.35E+14	6.76E+13	1.945	0.1735
leaf:rep	6	8.59E+12	1.43E+12	0.041	0.9996
Residuals	17	5.91E+14	3.48E+13		

Table A37: ANOVA table for leaf day 10 population counts of avellBP631* hopAB2 *deletion mutants *(corresponds to Figure 5.10C).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
genotype	1	1.75E-01	1.75E-01	0.223	0.6426	
leaf	2	1.32E+01	6.61E+00	8.417	0.00288	**
leaf:rep	6	4.68E-01	7.80E-02	0.099	0.99546	
Residuals	17	1.34E+01	7.86E-01			

Table A38: ANOVA table for leaf day 10 population counts of* Psm *R2 leaf* hopAB3 *deletion mutants *(corresponds to Figure 5.11).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	3	15.074	5.025	5.661	0.00444	**
leaf	2	4.266	2.133	2.403	0.11186	
leaf:rep	6	1.013	0.169	0.19	0.97665	
Residuals	24	21.301	0.888			
Groups						
hopAR	26.20112	a				
Wildtype	25.32703	ab				
hopBB	24.66926	b				
hopAZ	24.5877	b				

Table A39: ANOVA table for leaf day 10 population counts of* Psm *R2 leaf single* vir *gene deletion mutants *(corresponds to Figure 5.13).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)
strain	1	9.34E+13	9.34E+13	0.021	0.889
leaf	2	1.95E+16	9.76E+15	2.184	0.175
leaf:rep	6	8.82E+15	1.47E+15	0.329	0.904
Residuals	8	3.58E+16	4.47E+15		

Table A40: ANOVA table for leaf day 10 population counts of the* Pss *9644 hopAR1 deletion mutant *(corresponds to Figure 5.13).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
genotype	7	1.04E+01	1.48E+00	1.237	0.289	
leaf	2	9.62E+00	4.81E+00	4.013	0.0208	*
leaf:rep	6	1.70E+00	2.83E-01	0.236	0.964	
Residuals	110	1.32E+02	1.20E+00			

Table A41: ANOVA table for leaf day 10 population counts of the* Psm *R2 leaf single, double and triple* vir *gene deletion mutant *(corresponds to Figure 5.14).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	11	514.7	46.79	2.451	0.0101	*
leaf	2	6.1	3.07	0.161	0.8516	
leaf:rep	6	130.2	21.7	1.137	0.3478	
Residuals	88	1679.8	19.09			

Groups

trt	means	M
avrPto1	8.7692298	a
hopAB1	8.1538307	a
hopC1	7.4414225	ab
hopAO1	7.1640707	ab
WT	7.0608408	ab
hrpA	6.2878911	ab
hopAW1	5.5954183	ab
hopAA1	5.0505633	ab
avrRps4	4.6316582	ab
pBBR1MCS5	4.2950673	ab
avrA	3.6099206	ab
hopG1	0.5557158	b

Table A42: ANOVA table and Tukey-HSD groups for leaf day 0 population counts of* Psm *R1 5244 transconjugants on cherry *(corresponds to Figure S5.6).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	11	1749.9	159.08	11.29	9.12E-13	***
leaf	2	8.3	4.13	0.293	0.747	
leaf:rep	6	146.2	24.36	1.729	0.124	
Residuals	88	1239.9	14.09			

Groups

trt	means	M
avrA	8.213567	a
hopAO1	6.932982	ab
WT	6.832839	ab
hopG1	2.56742	abc
hrpA	1.10795	bc
avrPto1	-2.103427	c
hopC1	-2.103427	c
avrRps4	-2.214459	c
hopAA1	-2.214459	c
hopAB1	-2.214459	c
hopAW1	-2.214459	c
pBBR1MCS5	-2.214459	c

Table A43: ANOVA table and Tukey-HSD groups for leaf day 0 population counts of* Psm *R2 leaf transconjugants on cherry *(corresponds to Figure S5.6).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	10	505.9	50.59	2.189	2.67E-02	*
leaf	2	39.6	19.8	0.857	0.4285	
leaf:rep	6	258.8	43.13	1.867	0.0968	.
Residuals	80	1848.7	23.11			

Groups

trt	means	M
hopC1	9.621062	a
hopAO1	9.055034	ab
pBBR1MCS5	8.316009	ab
avrRps4	7.80869	ab
WT	7.42978	ab
avrPto1	6.885251	ab
avrA	6.52766	ab
hopAW1	6.274728	ab
hopAA1	3.862394	ab
hopAB1	3.832023	ab
hopG1	2.051012	b

Table A44: ANOVA table and Tukey-HSD groups for leaf day 0 population counts of* Pss *9644 transconjugants on cherry *(corresponds to Figure S5.6).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	4	75.3	18.83	1.384	2.60E-01	
leaf	2	108.4	54.18	3.981	0.0277	*
leaf:rep	3	123.4	41.12	3.022	0.0426	*
Residuals	35	476.3	13.61			

Table A45: ANOVA table for leaf day 0 population counts of* Psm *R1 5244 transconjugants expressing candidate* avr *genes on cherry *(corresponds to Figure S5.7).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)
strain	4	45.7	11.43	1.027	4.07E-01
leaf	2	36	18.01	1.618	0.213
leaf:rep	3	51.5	17.17	1.542	0.221
Residuals	35	389.6	11.13		

Table A46: ANOVA table for leaf day 0 population counts of* Psm *R2 leaf transconjugants expressing candidate* avr *genes on cherry *(corresponds to Figure S5.7).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	4	127.26	31.82	3.709	1.28E-02	*
leaf	2	75.22	37.61	4.385	0.02	*
leaf:rep	3	11.04	3.68	0.429	0.7333	
Residuals	35	300.2	8.58			

Groups

trt	means	M
pBBR1MCS5	8.469772	a
hopAB2	8.257993	a
hopC1	7.774596	ab
hopAB3	6.931598	ab
hopAB1	3.867784	b

Table A47: ANOVA table for leaf day 0 population counts of* Pss *9644 transconjugants expressing candidate* avr *genes on cherry *(corresponds to Figure S5.7).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)
strain	4	33	8.244	0.833	0.514
leaf	2	10.1	5.074	0.513	0.604
leaf:rep	6	46.9	7.809	0.789	0.585
Residuals	32	316.7	9.896		

Table A48: ANOVA table for leaf day 0 population counts of* Psm *R1 5300* hopAB1 *gene deletion mutants on cherry *(corresponds to Figure S5.8).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)
strain	2	33.83	16.917	0.969	0.401
leaf	2	24.74	12.372	0.709	0.507
leaf:rep	6	46.18	7.697	0.441	0.841
Residuals	16	279.29	17.456		

Table A49: ANOVA table for leaf day 0 population counts of* Psm *R1 9657* hopAB1 *gene deletion mutants on cherry *(corresponds to Figure S5.8).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)
strain	2	11.13	5.56	0.332	0.7226
leaf	2	103.51	51.76	3.085	0.0736.
leaf:rep	6	169.26	28.21	1.681	0.1897
Residuals	16	268.44	16.78		

Table A50: ANOVA table for leaf day 0 population counts of* avellBPIC631 hopAB2 *gene deletion mutants on cherry *(corresponds to Figure S5.8).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)
strain	2	25.76	12.88	0.805	0.464
leaf	2	42.89	21.44	1.34	0.29
leaf:rep	6	61.08	10.18	0.636	0.7
Residuals	16	255.96	16		

Table A51: ANOVA table for leaf day 0 population counts of* Psm *R2 leaf* hopAB3 *gene deletion mutants on cherry *(corresponds to Figure S5.9).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)	
strain	3	266.2	88.72	4.046	0.0184	*
leaf	2	8.1	4.05	0.185	0.8327	
leaf:rep	6	136.6	22.76	1.038	0.4257	
Residuals	24	526.3	21.93			

Groups

trt	mean	M
hopAR	8.097502	a
Wildtype	6.832839	ab
hopBB	2.473284	ab
hopAZ	1.774236	b

Table A52: ANOVA table for leaf day 0 population counts of* Psm *R2 leaf single* vir *gene deletion mutants on cherry *(corresponds to Figure S5.10A).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)
strain	1	3.11	3.106	0.304	0.597
leaf	2	9.06	4.53	0.443	0.657
leaf:rep	6	36.8	6.133	0.6	0.725
Residuals	8	81.79	10.224		

Table A53: ANOVA table for leaf day 0 population counts of a* Pss *9644* hopAR1 *gene deletion mutant on cherry *(corresponds to Figure S5.10B).*

ANOVA

	Df	Sum Sq	Mean Sq	F.value	Pr(>F)
strain	13	239.8	18.44	1.173	0.31
leaf	2	53.3	26.64	1.694	0.189
leaf:rep	6	72.3	12.04	0.766	0.598
Residuals	104	1635.2	15.72		

Table A54: ANOVA table for leaf day 0 population counts of* Psm *R2 leaf single, double and triple* vir *gene deletion mutants on cherry *(corresponds to Figure S5.11).*

The following paper was published in collaboration with Dr Bo Li, where M. Hulin and B. Li are joint first authors. It describes an automated image analysis method to rapidly quantify the percentage area of infection from images of inoculated cut shoots. The same data is presented in Chapter 3. As the paper is mainly concerned with the development of computational software it is presented here.

Li *et al. Plant Methods (2015) 11:57*
DOI 10.1186/s13007-015-0100-8

Plant Methods

RESEARCH Open Access

Rapid, automated detection of stem canker symptoms in woody perennials using artificial neural network analysis

Bo Li[1†], Michelle T. Hulin[1,3†], Philip Brain[1], John W. Mansfield[2], Robert W. Jackson[3] and Richard J. Harrison[1,3*]

Abstract

Background: *Pseudomonas syringae* can cause stem necrosis and canker in a wide range of woody species including cherry, plum, peach, horse chestnut and ash. The detection and quantification of lesion progression over time in woody tissues is a key trait for breeders to select upon for resistance.

Results: In this study a general, rapid and reliable approach to lesion quantification using image recognition and an artificial neural network model was developed. This was applied to screen both the virulence of a range of *P. syringae* pathovars and the resistance of a set of cherry and plum accessions to bacterial canker. The method developed was more objective than scoring by eye and allowed the detection of putatively resistant plant material for further study.

Conclusions: Automated image analysis will facilitate rapid screening of material for resistance to bacterial and other phytopathogens, allowing more efficient selection and quantification of resistance responses.

Keywords: Stem canker, Artificial neural network, Image analysis

Background

The bacterial phytopathogen *Pseudomonas syringae* encompasses pathovars that infect over 180 plant species. Three distinct clades of *P. syringae* (pv *morsprunorum* race 1, pv. *morsprunorum* race 2 and pv. *syringae*) are the major causal agents of bacterial canker of *Prunus* species grown worldwide [1]. This genus of stone fruit trees includes economically important species such as cherry and plum. The bacteria are able to infect all aerial plant organs, including leaves, blossom and fruit. Severe damage to the tree occurs when bacteria infect woody tissues via wounds or leaf scars to produce necrotic cankers that are often associated with extensive gummosis [2]. These cankers cause girdling of branches and may result in dieback or eventual death of the tree when affecting the main trunk [3]. The disease commonly results in tree losses of approximately 20 %, however, in severe cases, losses of up to 75 % have been reported in the US [4, 5].

Current control methods for this disease are limited. They include good hygiene when pruning, to reduce the likelihood of infection and the use of copper-based spays to control epiphytic bacterial populations [6]. The breeding of resistant cultivars, complemented with excellent sanitation methods, would be the most effective control of this disease [7]. At present, no cultivars have been shown to exhibit complete resistance; however there is variation in disease susceptibility [2], meaning breeding approaches could be successful. Therefore, a rapid disease screening method would be highly beneficial in *Prunus* breeding programmes, to allow the identification of resistant genotypes.

Susceptibility to bacterial canker is usually determined by visually assessing natural infection in the field over several years [8]. This approach is time consuming and different environmental conditions between fields may lead to misleading results [9]. Several rapid laboratory-based assays have been proposed, including the use of cut shoots [3, 8, 10], immature fruits [11, 12] and

*Correspondence: richard.harrison@emr.ac.uk
†Bo Li and Michelle T. Hulin equally contributed to this work
[3] School of Biological Sciences, University of Reading, Reading RG6 6AJ, UK
Full list of author information is available at the end of the article

BioMed Central

© 2015 Li et al. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

micro-propagated plantlets [9] to examine disease susceptibility. In this study we assessed the use of the cut shoot assay to screen *Prunus* cultivars for susceptibility to bacterial canker. The assay involves inoculating first-year dormant shoots with *P. syringae* and estimating disease severity based on the extent of necrosis. This approach, although more rapid than field-based observations, was found to be variable between assessors, being based on a subjective appraisal of lesion development and therefore lacked reproducibility, as has been shown in other similar studies [13]. A more rapid and high-throughput alternative to visual assessment involves the use of automated image analysis software [14, 15].

Automated image analysis is becoming a popular tool for plant disease assessment as it potentially provides greater speed, accuracy and reliability [16]. Nilsson [17] was the first to report the utility of remote sensing and image analysis for plant pathology. After Nilsson, various studies successfully applied image analysis in the visible region for disease severity assessment [18–22], with such techniques excellently reviewed in [23]. Digital image analysis has been compared with visual disease assessment for several diseases such as coffee rust [24], powdery mildew [25], yellow rust [26] and citrus canker [27]. These studies indicated that colour or monochrome image analysis provided more accurate measurement, whilst drastically reducing the time required for examination [16, 28].

Among the different image analysis algorithms used to measure disease severity, the conversion from RGB (Red Green Blue) to HSI (Hue, Saturation and Intensity) colour space is commonly used and the hue value has been considered to be an effective channel to discriminate healthy and diseased areas on colour images [16]. The hue channel threshold can be set manually or automatically to segment diseased from healthy areas using software such as Adobe Photoshop [29], ASSESS© [30], Scion image software (Scion Corporation, Frederick, MD) [21], ImageJ [31] or other custom developed software programs [32, 33].

Other more sophisticated algorithms have been proposed for the automatic classification of plant diseases using colour images. Naikwadi [34] converted RGB images to HSI format and applied Spatial Gray-level Dependence Matrices (SGDM) as the colour co-occurrence texture analysis method for only H (hue) and S (saturation) images. Grey-level co-occurrence methodology was used to calculate the features, which were inputted into neural networks for recognition. Apart from HSI colour space, colour images have also been converted to the L1 L2 L3 colour model for disease area measurement [18, 35]. Schikora [19] utilised this method for the image-based analysis of plant infection with human pathogens. The L2 and L3 values plus the information of the surrounding pixels were classified via supervised learning techniques such as neural networks or support vector machines.

The use of Artificial Neural Networks (ANN) has recently become a popular tool of pattern recognition in image analysis [36] and disease quantification [37]. ANN is an efficient computational model inspired by the parallel nervous systems of animals [38]. It is widely implemented in machine learning and has been applied to the food and agricultural industry [39, 40]. The use of ANN has also been trialed for detection and quantification of various plant diseases [41–44]. The whole system is based upon an interconnection of neurons, which computes the output from the input variables. Besides input and output layers, ANN systems always have one or more hidden layers between them. A training dataset is used to update the adaptive weights of all the neurons in order to minimize the mean square error between the output and ideal values below a certain criteria [38].

This paper reports the development of an automated image analysis software which utilises ANN to analyse images of cherry and plum shoots exhibiting necrosis due to bacterial canker, with the goal of improving the accuracy of disease resistance screening. The software developed reduces the time and subjectivity involved in disease assessment and has the potential to be applied during screening of other important tree diseases.

Results and discussion

Quantification based on automated image analysis

A feed-forward artificial neural network (ANN), which is also known as multi-layer perceptrons (MLP), was implemented for the classification of diseased and healthy shoot tissue (see "Methods" section for full details). The recognition of diseased area is based on the colour, and only R, G and B values were used as the input variables of the ANN model. The training samples consisted of pixels labelled as healthy and diseased, and in total 75,155 pixels were manually labeled from 13 images, covering all the variation in colour due to disease. All the images were taken under the same illumination, and the colours of the diseased region showed little variation. The image analysis was applied to 420 images of inoculated shoots, producing estimates of percentage area and length of necrosis to determine disease severity.

To determine the utility of our image analysis software we compared results with both a current method of disease image analysis and expert measurements made by eye. 84 images (block 1 and 2) (e.g. Fig 1a) were analysed by our software to determine percentage area of necrosis, which was then correlated with the output for the same images produced using ImageJ manual thresholding

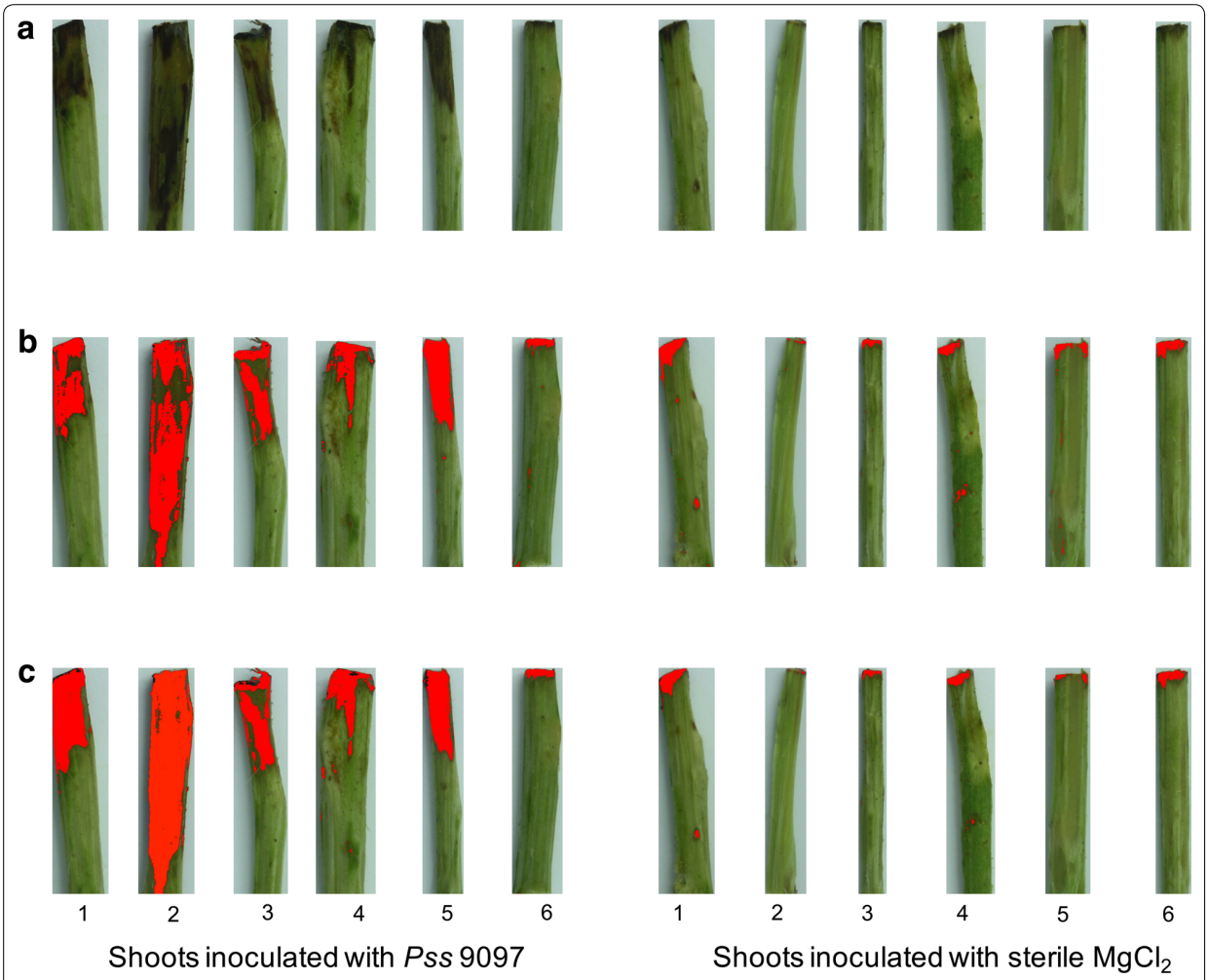


Fig. 1 Images of cut shoots with thresholding of disease using ImageJ and the automated software. Shoots were inoculated with either the virulent strain *Pss* 9097 or with sterile 10 mM $MgCl_2$ as a control. **a** Original image, **b** thresholding with Image J, **c** thresholding with automated software. *1*: Cherry cv. Van, *2*: Cherry cv. Napoleon, *3*: Cherry cv. Roundel, *4*: Cherry cv. Merton Glory, *5*: Plum cv. Victoria, *6*: Plum cv. Marjorie's seedling

(Fig. 1b). We also used the software to determine the length of necrosis on each shoot (at the longest point), which was correlated with data for the same images, measured manually using a caliper. The software produced both a pictorial output (e.g. Fig 1c) and raw data (available on github).

Correlation analysis and linear regression indicated results were highly similar using the image analysis software and the other methods of assessment. Figure 2 shows the correlation of percentage area of necrosis whilst Fig. 3 shows the same for necrosis length. A linear regression produced r^2 values of 0.87 and 0.81 for percentage area and length respectively. In both Figs. 2 and 3, there was deviation between the linear regression line and the ideal calibration line. This difference between methods likely resulted from using an arbitrarily threshold in ImageJ and subjective labelling of diseased pixels in training images. To further test this, Lin's concordance coefficient [45] was calculated with rhoC values of 0.9 (moderate correlation) and 0.89 (poor correlation as <0.9) for the area and length data respectively. Due to the lower score for the length data when comparing manual measurement and the new software, this data was not used in further analysis of the experiment. This poor rhoC value for the length dataset could be due to manual assessment of length being more subjective. It was sometimes difficult to measure length of necrosis accurately due to natural blemishes on the sample shoots. The automated software could provide a more objective method than classification by eye, however this would need

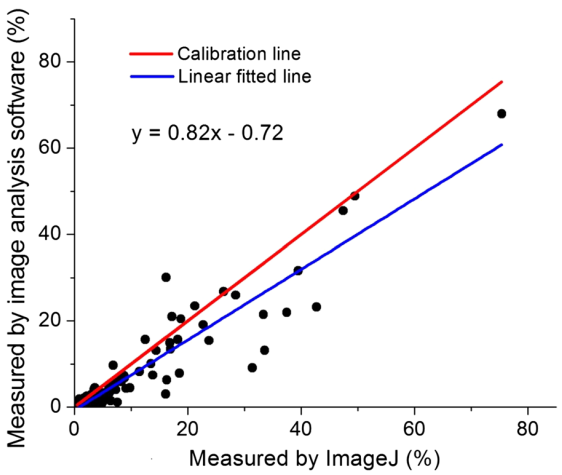


Fig. 2 Comparison between the measurements by automated image analysis software and assessment by ImageJ on diseased area. The *linear fitted line* indicated the deviation from the ideal calibration line

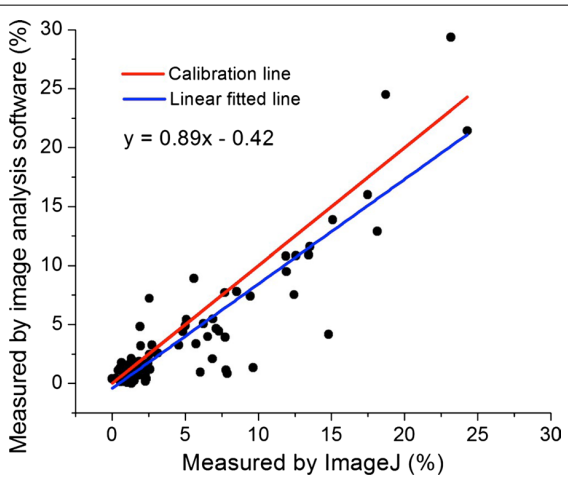


Fig. 3 Comparison between the measurements by automated image analysis software and manual measurements by experts on the length of the diseased area. The *linear fitted line* indicated the deviation from the ideal calibration line

further testing to validate. Overall, the correlation analysis indicated that the automated software could produce results comparable to currently used manual assessment.

The accuracy of the automated measurements relied on an expert's selection of diseased areas on the images used as the training data. This was necessary to ensure all the typical colours of both diseased and healthy areas were included, reducing the potential for misclassification. The criteria used during prediction of the percentage disease were selected empirically. To our knowledge this is the first time that image analysis and machine-learning algorithms have been applied to disease quantification on plant shoots. Compared with assessment by eye/use of ImageJ manual thresholding, the image analysis software only needs to be trained once by an experienced expert. Many images captured under the same lighting condition can therefore be processed using the same model, which could reduce the subjectivity. The time taken to process all 420 images was approximately 42 s (0.1 s per image) with current hardware and ANN model, so the image analysis software was much faster than traditional methods (ImageJ 60–100 s per sample). The results were also compared with other common image thresholding methods such as fixed thresholding and Otsu's method. It was found that the fixed thresholding produced a comparable correlation with manual assessment ($r^2 = 0.86$) but Otsu's thresholding methods showed poor results (see Additional file 1: Figure S6 and S7).

With a proper training dataset, the chosen method provided a fast, automated and objective method for disease quantification on cherry shoots. It could be utilised for general disease quantification during other biological experiments with different illumination condition. ANN is a more flexible approach than other thresholding methods, since biologists only need to label regions as diseased or healthy rather than arbitrarily determining a threshold for disease. Further development of the software could involve more input parameters such as texture information, so ANN is more extendable to other input variables.

Development of automated image analysis software and a graphical user interface

In order to make the software user-friendly, a graphical user interface was developed. The GUI can be used to select the training data on a series of images from a particular folder (see Additional file 1: Figure S1). This selection is semi-automatic as user interaction is necessary to drag the mouse and draw a rectangle within healthy and diseased regions. The colour information of all the pixels inside the rectangle is recorded as healthy or diseased to train the ANN model.

The trained ANN model can subsequently be applied to calculate the percentage area of necrosis. The pixels labelled as diseased are coloured as red (Additional file 1: Figure S2). The resulting image with false colour can be further analysed to estimate the length of disease by measuring the height of the fitted rectangles (Additional file 1: Figure S3). The source code of the software is available on Github (https://github.com/eastmallingresearch/Cherry_shoots).

Results of pathogenicity assays on cherry and plum

Following training, the automated image analysis software was used for a resistance screen to produce percentage area necrosis data for six strains of *P. syringae* inoculated onto four cultivars of cherry and two cultivars of plum. The strains included *P. syringae* pv. *morsprunorum* race 1 isolated from cherry (5244) and plum (5300), *P. syringae* pv. *morsprunorum* race 2 isolated from cherry (5255) and *P. syringae* pv. *syringae* isolated from cherry (9097) and plum (9293). A strain isolated from hazelnut (*P. syringae* pv. *avellanae*) was also used for comparison as a non-pathogen of *Prunus*.

The plant cultivars (cvs) were chosen as they have a range of susceptibility to the different races of *P. syringae* that infect *Prunus*. The cherry cv Van is reported to be universally susceptible, whilst cv Merton Glory is tolerant/has a lower susceptibility to the pathogen [46, 47]. The cultivars Napoleon and Roundel are reported to show differential susceptibility to the different races of *P. syringae* pv. *morsprunorum* [47], with cv Napoleon being resistant to R2 but susceptible to R1 and vice versa for cv Roundel. For plum, the cv Victoria is highly susceptible, while cv Marjorie's Seedling is reportedly resistant/tolerant [48].

The different strains of *P. syringae* caused variable levels of necrosis on shoots of cherry (Fig. 4) and plum (Fig. 5). An analysis of variance (ANOVA) was performed using the log transformed percentage data (Additional file 1: Figure S4). The ANOVA revealed that there was a significant effect of *Pseudomonas* strain on percentage area of necrosis ($p < 0.001$, $df = 6$), likely due to variation in the virulence of the different strains. There was no significant difference in percentage area of necrosis between the two *Prunus* species ($p = 0.06$, $df = 1$) indicating both species exhibit similar levels of susceptibly to the disease. However there was a significant interaction between *Prunus* species and *P. syringae* strain ($p < 0.001$, $df = 6$). This indicates that the different *P. syringae* strains show differential virulence on cherry and plum (Figs. 4, 5).

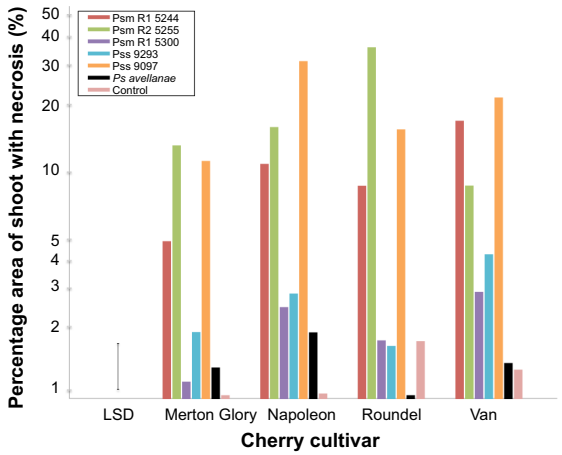


Fig. 4 Percentage area of necrosis on cherry shoots inoculated with *P. syringae* for four cherry cultivars (plotted on a log scale with back-transformed values as the scale). Cv Van is universally susceptible, whilst cv Merton Glory has tolerance/lower susceptibility. Cv Napoleon is resistant to *Psm* R2 but susceptible to R1 and vice versa for cv Roundel. The control was sterile 10 mM $MgCl_2$. The mean values were calculated using ANOVA. *LSD* Least Significant Difference

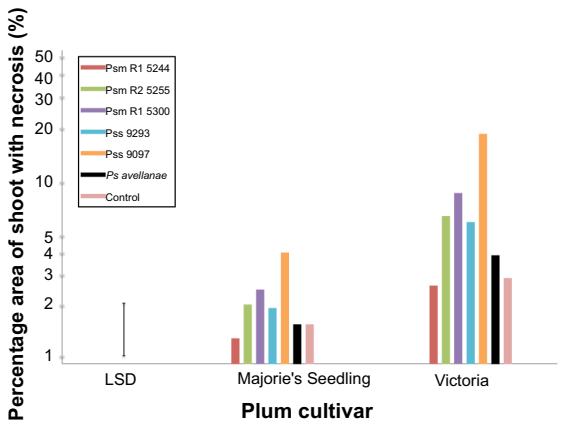


Fig. 5 Percentage area of necrosis on plum shoots inoculated with *P. syringae* for two plum cultivars (plotted on a log scale with back-transformed values as the scale). Cv Victoria is highly susceptible and cv Marjorie's Seedling is resistant/tolerant. The control was sterile 10 mM $MgCl_2$. Mean values were calculated using ANOVA. *LSD* Least Significant Difference

On cherry, the three strains isolated from cherry (*Psm* R1 5244, *Psm* R2 5255 and *Pss* 9097) were generally associated with severe necrosis (>5 % of total shoot area), whilst necrosis caused by other strains failed to exceed 5 % shoot area. *Pss* 9097 caused significant symptom development on all cultivars, whereas necrosis caused by the two races of *Psm* isolated from cherry, varied considerably between cultivars. This supports previous hypotheses that cherry cultivars exhibit differential susceptibility towards the two races of *Psm* [49]. In the global ANOVA (Table S1) there was no overall interaction between strain, cultivar and species. However, when the comparison was restricted to Van and Roundel, a highly significant interaction ($p = 0.004$) was detected between the two cultivars and the strains, which is driven by the differences between *Psm* R1 and *Psm* R2. The cultivars Roundel and Van showed differential susceptibility to the two *Psm* races. On Van, *Psm* R1 caused more severe necrosis than *Psm* R2, whilst on Roundel this response was reversed. One reason for this could be that plant immunity responses to the different races vary between

cultivars. Overall the results indicated that no single cultivar of cherry was tolerant to all strains. The symptoms on Merton Glory never exceeded 25 % of the shoot area, indicative of partial tolerance. Therefore, a cross between Merton Glory and a more susceptible cultivar could be used to further investigate the genes involved in tolerance/resistance.

On plum (Fig. 5), the level of necrosis was generally higher on cv Victoria compared to Marjorie's Seedling. Interestingly, the two strains originally isolated from plum (*Psm* R1 5300 and *Pss* 9293) caused a higher level of necrosis on plum than on cherry. Also, when inoculated on plum they generally caused more severe necrosis than strains isolated from cherry and hazelnut (*Psm* R1 5244, *Psm* R2 5255 and *Ps. avellanae*). The virulence of these plum strains on plum could be due to host-specific factors, which allow the pathogens to survive longer and cause more necrosis in their natural (homologous) host.

The plum cultivar Marjorie's Seedling showed some resistance to most strains, with the severity of necrosis being similar to the control (inoculation with sterile $MgCl_2$). It was also more tolerant to the virulent *Pss* strain 9097. This supports previous reports that this cultivar is tolerant to bacterial canker. Therefore, Marjorie's Seedling could be a target for further investigations of the genetics of resistance.

Conclusion

In this study a method for automated image analysis to measure the severity of disease symptoms was developed using a machine learning approach. To validate the reliability of our automated software, cherry and plum shoot images were analysed to measure necrosis using the free program ImageJ [31]. The ImageJ analysis was based on the hue value of the colour images and the threshold between the diseased and healthy area was determined arbitrarily, resulting in a loss of the colour information from the other two channels. The 3D shape of cherry shoots resulted in shadows, leading to a colour similar to the diseased area in grayscale images or the hue channel of HSV space, but still distinguishable by the naked eye. Furthermore, manual image analysis using ImageJ can only process one image at a time and the images need to be loaded manually before applying the thresholding technique, which is extremely time consuming.

Due to the variation in the colour of diseased and healthy areas, it is difficult to set arbitrary thresholds for all three channels of colour space. The new image analysis method employed artificial neural networks (ANN) for the training and classification of a colour dataset. With the expert's selection of training data featured by the RGB values and ANN as the classification algorithm, the quantification of disease was highly correlated with a subjective quantification method implemented in ImageJ. The software greatly reduced the time requirements for disease assessment when compared to manual thresholding with imageJ. This assisted in the objective identification of differences in cultivar susceptibility to the various strains that cause bacterial canker. This software therefore provides opportunities to shorten time taken for disease assessment dramatically. The software would facilitate the use of the cut shoot test for high-throughput screening during breeding programmes. This would enable the selection of putatively resistant material from mapping populations, which often contain hundreds of individuals. Finally, this software is highly adaptable and could be implemented during the screening of other tree diseases.

Methods

Bacterial strains

Strains of *Pseudomonas syringae* were grown on King's B agar (Sigma) at 25 °C. For liquid culture, strains were grown in Luria Broth (Melford) at 25 °C, 150 rpm. Strains were obtained from various sources (Table 1) and included representatives of the three major clades that infect *Prunus* (*Psm* Race 1, *Psm* Race 2 and *Pss*) as well as an out-group strain belonging to pv. *avellanae* which was isolated from hazelnut (*Corylus avellana*).

Plant material

Dormant first-year shoots were collected from mature cherry and plum trees in December 2014 at East Malling Research, Kent.

Pathogenicity assay on cut shoots

The cut shoot pathogenicity assay was performed as in previous studies [3, 8]. Each cultivar x strain treatment

Table 1 List of bacterial strains used in pathogenicity assays, with source host and reference

Strain	Species	Pathovar	Race	Source of isolation	Isolate curator
5244	*P. syringae*	*morsprunorum*	1	*Prunus avium*	SJ Roberts
5300	*P. syringae*	*morsprunorum*	1	*Prunus domestica*	SJ Roberts
5255	*P. syringae*	*morsprunorum*	2	*Prunus avium*	SJ Roberts
9097	*P. syringae*	*syringae*	–	*Prunus avium*	SJ Roberts
9293	*P. syringae*	*syringae*	–	*Prunus domestica*	SJ Roberts
BPIC631	*P. syringae*	*avellanae*	–	*Corylus avellana*	DS Guttman

was replicated 10 times, resulting in 420 inoculations. To prepare the bacteria, single colonies were inoculated in LB and shaken overnight. These cultures were spun down using a centrifuge (4000 rpm, 10 min) and re-suspended in 10 mM $MgCl_2$. The concentration was adjusted to 1×10^7 CFU/ml (confirmed by dilution plating) and sterile 10 mM $MgCl_2$ was used for the control. For the plant material, dormant first-year shoots of similar diameter (5 mm) were collected from cherry and plum trees in December and cut into 10 cm sections using secateurs. These were surface-sterilised in 0.5 % hypochlorite for 5 min and rinsed with tap water. The shoot sections were air-dried overnight.

To inoculate, the top 5 mm of each shoot tip was removed with a scalpel and dipped for 5 min in the bacterial suspension. The wound was covered with parafilm (Fisher Scientific, UK) and the shoot bases were freshly cut (approx. 5 mm) and placed in transparent-boxes immersed in water to a depth of 20 mm. The shoots were incubated in the closed boxes at 15 °C with 16-hour light, 8-hour dark cycle for 1 week. Separate boxes were used for each bacterial isolate to prevent cross-contamination. Next, the shoots were transferred to −2 °C for one week to simulate frost damage. Finally, the basal 10 mm of each shoot was removed and they were placed in a completely randomised design (generated using Genstat [50]) in water-soaked Oasis Foam (Oasis Floral, UK) in trays containing 30 mm of water. These were incubated for a further 4 weeks at 15 °C with the same light conditions as previously described. The trays were covered with clingfilm to maintain a high humidity.

The shoots were assessed for severity of stem canker by peeling back the uppermost layer of bark from the top 30 mm of the shoot to expose the symptoms, which were photographed digitally. The length of necrosis was also manually measured with a caliper.

Imaging system

All the images were captured using a SLR camera (Canon EOS 1000D) with 53 mm focal length and 1/15 s exposure time. Two 60 W incandescent light bulbs were used to illuminate the samples from each side. The distance between the lens and the samples was 35 cm. Due to the high resolution of the imagery device (3888 × 2592 pixels), three shoots were placed on a spectralon white platform (SphereOptics) and imaged together in order to enhance the contrast between the foreground and background. The images were captured using EOS utility software (Canon) and saved as JPG files. Individual shoots were cropped from each image and due to small variations in the size of shoots, the resolution of the images varied from 43 × 754 to 282 × 839 pixels. All the images were saved and processed on a Dell desktop computer (Intel® Xeon(R) CPU X5560 @ 2.80 GHz × 16). The automated image analysis software was written in C++ [51] utilising the OpenCV Library [52] on an Ubuntu 14.04 operating system.

Statistics

Genstat [50] was used to perform the statistical analysis using a nested ANOVA (nesting cultivar by species), whilst Excel [53] was used to produce bar charts (Fig. 4 and 5). The residuals of ANOVA tests were assessed for normality using qqnorm (residuals). If the residuals were not normally distributed the data was log transformed (with the addition of 0.1 to area prior to log transformation) and the ANOVA repeated. Log transformation was selected rather than the more conventional square root transformation, as the full spectrum of percentage disease was not used and the relationship between residuals and fitted values was less biased. Furthermore, a log transformation is more appropriate to study multiplicative interactions between factors. Full ANOVA tables and residual plots can be found in the supplementary information (Additional file 1: Tables S1, S2; Figures S4, S5). A complete randomised design for the positioning of cherry shoots in trays after inoculation was produced using Genstat [50].

Image analysis with ImageJ

ImageJ [31] was used to manually measure the disease severity on an image-by-image basis. Firstly, the three cherry shoots were cropped from the original image and converted from RGB to HSI colour space. A threshold was manually chosen to determine the total number of pixels in the shoot (compared to the total in the whole image containing the background). The total number of pixels in the shoot was named R1. The second threshold on the hue channel was used to segment the diseased and healthy areas. As the diseased area always showed darker intensity than the healthy area, the background could be easily separated. The total number of pixels in the diseased area was called R2. The proportion of the diseased area was calculated using the ratio of the diseased area (R2) to the total shoot area (R1).

Automated image analysis software

The automated image analysis software was developed in C++ with open computer vision library (OpenCV 2.4.9), and the interface was designed by Qt designer. The software is programmed to load all images in a single folder and process them in a batch with the prediction parameters included and output the percentage area of necrosis and the necrosis length.

The original images were converted to grayscale and the pixels belonging to the three shoots were segmented

from the background by setting an arbitrary threshold. All the contours were detected, those <500 pixels were considered noise and were discarded leaving only the three shoots. Rectangles were fitted to all three contours, which were cropped from background and saved as three individual images for further processing.

A feed-forward artificial neural network (ANN) was implemented for the imaging classification. The ANN model consists of one input layer with three neurons, one binary output layer and one hidden layer with 16 neurons (Additional file 1: Figure S8). The input layer was the same size as the sample feature variables (Red Green and Blue) in this experiment. The input is passed to each neuron of the hidden layer and summed up with certain weights. A symmetrical sigmoid function was applied to the sum for each neuron and the output of each neuron on the hidden layer was further summed up with weights to the output. The model is trained with a training dataset to adjust the weights iteratively in order to minimize the error between ideal and real output.

Thirteen images were selected for extraction of the training dataset. The expert labelled pixels as diseased by drawing squares of different sizes by pressing the left mouse button to the diseased region. Similarly, the right button was used to label pixels as healthy. The original images were kept in RGB format and the R, G and B values were used as the three variables for the training phase.

In the prediction phase, the segmentation was applied first with an arbitrary threshold to separate the pixels belonging to the shoot from the background and input to the classification model, which reduces the computation cost. The R, G and B values for each pixel were taken as feature variables, classified by ANN and labelled as diseased or healthy. The pixels labelled as diseased were also false coloured as red for visualization. Ratios between the number of pixels in the diseased area and total area were calculated automatically and saved in text files. The length of necrosis measurement was based on the false colour image. Any red regions with less than 10 pixels were regarded as noise so were removed, whilst all other red regions were fitted with rectangles. If the area of fitted rectangles were less than 10,000 pixels, the correspondent red regions were further removed unless the regions were near the top of the shoots (at the point of infection). This was required to remove any blemishes that were not due to the disease. The final length was calculated by measuring the difference between the top and bottom of the rectangle.

The software is available from the East Malling github repository, (www.github.com/organizations/eastmallingresearch/).

Additional file

Additional file 1. Figure S1. A standard view of the GUI during manual selection of training data. **Figure S2.** A standard view of the GUI during percentage area of necrosis estimation. **Figure S3.** A standard view of the GUI during necrosis length estimation. **Figure S4.** Visual output of ANOVA from Genstat including the histogram of residuals, fitted-value plot and normality plots. **Figure S5.** Interaction plot showing percentage necrosis for all *P. syringae* strains on the different *Prunus* species cherry and plum. The mean values from the ANOVA were used and Least Significant Difference bars calculated during the ANOVA. The means are plotted on a log scale with back-transformed values as the scale. Strains isolated from cherry are highlighted in red whilst those isolated from plum are highlighted in purple. **Figure S6.** The relationship between the percentage area of necrosis measured using ImageJ or automated image analysis software using fixed thresholding. The linear fitted line indicated the deviation from the ideal calibration line. **Figure S7.** The relationship between the percentage area of necrosis measured using ImageJ or automated image analysis software using adaptive thresholding. The linear fitted line indicated the deviation from the ideal calibration line. **Figure S8.** Graphical depiction of an artificial neural network (ANN) model containing one input layer with three neurons, one binary output layer and one hidden layer with sixteen neurons. **Table S1.** List of bacterial strains used in pathogenicity assays, with source host and reference. **Table S2.** ANOVA table generated using Genstat. The ANOVA was performed on the log transformed (+0.1) raw data values. The formula used was aov(strain*(species/cv)+box).

Authors' contributions
RJH devised the study in collaboration with MH and BL. MH and BL carried out experimental work and software development respectively. JM and RWJ assisted with the development of the pathogenicity test. All authors read and approved the final manuscript.

Author details
[1] East Malling Research, New Road, East Malling, ME19 6BJ Kent, UK. [2] Faculty of Natural Sciences, Imperial College London, SW7 2AZ London, UK. [3] School of Biological Sciences, University of Reading, Reading RG6 6AJ, UK.

Acknowledgements
We thank Steve Roberts and David Guttman for generously providing bacterial strains. We also thank Karen Russell and Connie Garrett for valuable advice about the pathogenicity test development and its potential deployment in breeding programmes.

Competing interests
The authors declare that they have no competing interests.

Received: 8 September 2015 Accepted: 9 December 2015
Published online: 24 December 2015

References

1. Bultreys A, Kaluzna M. Bacterial cankers caused by *Pseudomonas syringae* on stone fruit species with special emphasis on the pathovars *syringae* and *morsprunorum* race 1 and race 2. J Plant Pathol. 2010;92:S1.21–33.
2. Roberts SJ. HNS179 Final Report 2013. http://tinyurl.com/hgkrkje.
3. Santi F, Russell K, Menard M, Dufour J. Screening wild cherry (*Prunus avium*) for resistance to bacterial canker by laboratory and field tests. For Pathol. 2004;34:349–62.
4. Wenneker M, Janse JD, De Bruine JA. Bacterial canker of plum trees, caused by *Pseudomonas syringae* pathovars, as a serious threat for plum production in the Netherlands. Commun Agric Appl Biol Sci. 2011;76:575–8.

5. Spotts RA, Wallis KM, Serdani M, Azarenko AN. Bacterial canker of sweet cherry in Oregon—infection of horticultural and natural wounds, and resistance of cultivar and rootstock combinations. Plant Dis. 2010;94:345–50.
6. Wimalajeewa DLS, Cahill R, Hepworth G, Schneider HG, Washbourne JW. Chemical control of bacterial canker (*Pseudomonas syringae* pv. *syringae*) of apricot and cherry in Victoria. Aust J Exp Agric. 1991;31:705–8.
7. Thomidis T, Exadaktylou E. Susceptibility of 30 cherry (*Prunus avium*) genotypes to the bacterium *Pseudomonas syringae* pv. *syringae*. New Zeal J Crop Hortic Sci. 2008;36(October 2013):215–220.
8. Krzesinska EZ, Nina A, Azarenko M. Excised twig assay to evaluate cherry rootstocks for tolerance to *Pseudomonas syringae* pv. *syringae*. HortScience. 1992;27:153–5.
9. Vicente JG, Roberts SJ. Screening wild cherry micropropagated plantlets for resistance to bacterial canker. In: Santa Lacobellis N, Collmer A, Hutcheson SW, Mansfield JW, Morris C, Murillo J, Schaad NW, Stead DE, Surico G, Ullrich MS, editors. *Pseudomonas syringae* and related pathogens. Netherlands: Springer; 2003. p. 1–8.
10. Gilbert V, Planchon V, Legros F, Maraite H, Bultreys A. Pathogenicity and aggressiveness in populations of *Pseudomonas syringae* from Belgian fruit orchards. Eur J Plant Pathol. 2010;126:263–77.
11. Latorre BA, Jones AL. *Pseudomonas morsprunorum*, the cause of bacterial canker of sour cherry in Michigan, and its epiphytic association with *P. syringae*. Phytopathology. 1979;69:335–9.
12. Renick LJ, Cogal AG, Sundin GW. Phenotypic and genetic analysis of epiphytic *Pseudomonas syringae* populations from sweet cherry in Michigan. Plant Dis. 2008;92:372–8.
13. Rousseau C, Belin E, Bove E, Rousseau D, Fabre F, Berruyer R, Guillaumès J, Manceau C, Jacques M-A, Boureau T. High throughput quantitative phenotyping of plant resistance using chlorophyll fluorescence image analysis. Plant Methods. 2013;9:17.
14. Fahlgren N, Gehan MA, Baxter I. Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. Curr Opin Plant Biol. 2015;24:93–9.
15. Patil JK, Kumar R. Advances in image processing for detection of plant diseases. J Adv Bioinform Appl Res. 2011;2:135–41.
16. Bock CH, Poole GH, Parker PE, Gottwald TR. Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. CRC Crit Rev Plant Sci. 2010;29:59–107.
17. Nilsson HE. Remote sensing and image processing for disease. Prot Ecol. 1980;2:271–4.
18. Camargo A, Smith JS. An image-processing based algorithm to automatically identify plant disease visual symptoms. Biosyst Eng. 2009;102:9–21.
19. Schikora M, Neupane B, Madhogaria S, Koch W, Cremers D, Hirt H, Kogel KH, Schikora A. An image classification approach to analyze the suppression of plant immunity by the human pathogen *Salmonella Typhimurium*. BMC Bioinform. 2012;13:171.
20. Kim Khiook IL, Schneider C, Heloir MC, Bois B, Dair X, Adrian M, Trouvelot S. Image analysis methods for assessment of H_2O_2 production and *Plasmopara viticola* development in grapevine leaves: application to the evaluation of resistance to downy mildew. J Microbiol Methods. 2013;95:235–44.
21. Wijekoon CP, Goodwin PH, Hsiang T. Quantifying fungal infection of plant leaves by digital image analysis using Scion Image software. J Microbiol Methods. 2008;74:94–101.
22. Kokko EG, Conner RL, Lee B, Kuzyk AD, Kozu GC. Quantification of common root rot symptoms in resistant and susceptible barley by image analysis. Can J Plant Pathol. 2000;22:38–43.
23. Barbedo JGA. Digital image processing techniques for detecting, quantifying and classifying plant diseases. SpringerPlus. 2013;2:660.
24. Price TV, Gross R, Ho WJ, Osborne CF. A comparison of visual and digital image-processing methods in quantifying the severity of coffee leaf rust (*Hemileia vastatrix*). Aust J Exp Agric. 1993;33:97–101.
25. Olmstead JW, Lang GA, Grove GG. Assessment of severity of powdery mildew infection of sweet cherry leaves by digital image analysis. HortScience. 2001;36:107–11.
26. Huang W, Lamb DW, Niu Z, Zhang Y, Liu L, Wang J. Identification of yellow rust in wheat using in situ spectral reflectance measurements and airborne hyperspectral imaging. Precis Agric. 2007;8:187–97.
27. Bock CH, Cook AZ, Parker PE, Gottwald TR. Automated image analysis of the severity of foliar citrus canker symptoms. Plant Dis. 2009;93:660–5.
28. Lemein T, Cox D, Albert D, Mori N. Accuracy of optical image analysis compared to conventional vegetation measurements for estimating morphological features of emergent vegetation. Estuar Coast Shelf Sci. 2015;155:66–74.
29. Iyer-Pascuzzi AS, Symonova O, Mileyko Y, Hao Y, Belcher H, Harer J, Weitz JS, Benfey PN. Imaging and analysis platform for automatic phenotyping and trait ranking of plant root systems. Plant Physiol. 2010;152:1148–57.
30. Jackson EW, Obert DE, Menz M, Hu G, Avant JB, Chong J, Bonman JM. Characterization and mapping of oat crown rust resistance genes using three assessment methods. Phytopathology. 2007;97:1063–70.
31. Abràmoff MD, Magalhães PJ, Ram SJ. Image processing with ImageJ. Biophotonics Int. 2004;11:36–41.
32. Martin DP, Rybicki EP. Microcomputer-based quantification of maize streak virus symptoms in *Zea mays*. Phytopathology. 1998;88:422–7.
33. Ahmad IS, Reid JF, Paulsen MR, Sinclair JB. Color classifier for symptomatic soybean seeds using image processing. Plant Dis. 1999;83:320–7.
34. Naikwadi S, Niket A. Advances in image processing for detection of plant diseases. Int J Appl or Innov Eng Manag. 2013;2:168–75.
35. Schikora M, Schikora A. Image-based analysis to study plant infection with human pathogens. Comput Struct Biotechnol J. 2014;12:1–6.
36. Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. Psychol Rev. 1958;65:386–408.
37. Al-Hiary H, Bani-Ahmad S, Reyalat M, Braik M, ALRahamneh Z. Fast and accurate detection and classification of plant diseases. Int J Comp Appl. 2011;17(1):31–38.
38. Krogh A. What are artificial neural networks? Nat Biotechnol. 2008;26:195–7.
39. Wu D, Sun D-W. Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: a review—Part I: Fundamentals. Innov Food Sci Emerg Technol. 2013;19:1–14.
40. Wu D, Sun D-W. Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: a review—Part II: Applications. Innov Food Sci Emerg Technol. 2013;19:15–28.
41. Hetzroni A, Miles GE, Engel BA, Hammer PA, Latin RX. Machine vision monitoring of plant health. Adv Space Res. 1994;14:203–12.
42. Pydipati R, Burks TF, Lee WS. Statistical and neural network classifiers for citrus disease detection using machine vision. Trans ASAE. 2005;48:2007–14.
43. Huang KY. Application of artificial neural network for detecting *Phalaenopsis* seedling diseases using color and texture features. Comput Electron Agric. 2007;57:3–11.
44. Wang H, Li G, Ma Z, Li X. Application of neural networks to image recognition of plant diseases. In Proceedings of the 2012 International Conference on Systems and Informatics (ICSAI). 2012:2159-2164.
45. Nita M, Ellis MA, Madden LV. Reliability and accuracy of visual estimation of phomopsis leaf blight of strawberry. Phytopathology. 2003;93:995–1005.
46. Long L, Olsen J Sweet cherry cultivars for brining, freezing, and canning in Oregon. 2013. https://catalog.extension.oregonstate.edu/files/project/pdf/em9056.pdf. Accessed 12 May 2015.
47. APS. Merton cherries from England. J Fruit Var Hortic Dig. 1966; 20:46.
48. RHS. Bacterial canker. [https://www.rhs.org.uk/advice/profile?PID=86].
49. Garrett CME. Pathogenic races of *Pseudomonas morsprunorum*. In Proceedings of the IVth International Conference on Plant pathogenic Bacteria Vol II; 1978:889–890.
50. VSN International. Genstat for Windows 14th Edition. 2011. www.GenStat.co.uk. Accessed 12 May 2015.
51. Oualline S. Practical C++ Programming. 2nd ed. CA: O'Reilly; 2003.
52. Laganière R. OpenCV 2 Computer Vision Application Programming Cookbook. 2011.
53. Microsoft. Microsoft Excel [computer software]. 2011. Redmond: Microsoft.