Exploring the phenotypic and genetic diversity of a bacterial plant pathogen, the *Ralstonia solanacearum* species complex (RSSC)

Evelyn May Farnham

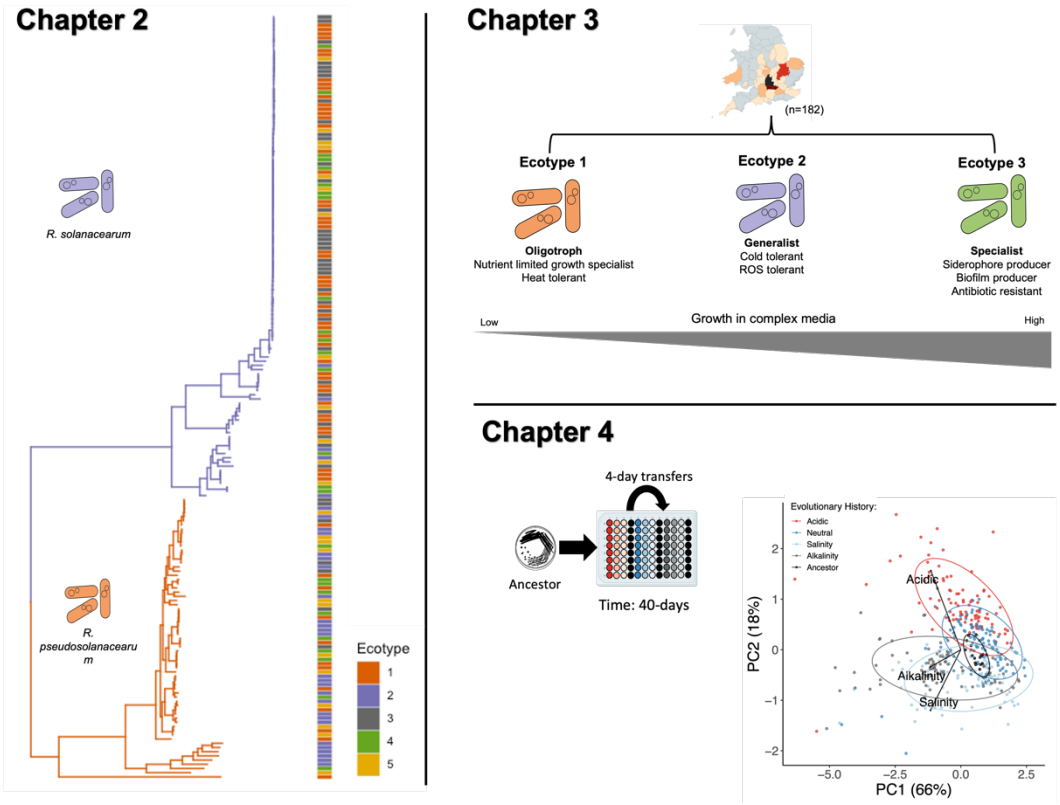
Doctor of Philosophy

University of York
Biology

December 2022

Abstract

The *Ralstonia solanacearum* species complex (RSSC) is a globally distributed bacterial plant pathogen and the causative agent of bacterial wilt disease. This pathogen can infect over 175 plant species, including many important crops, having huge impacts on agriculture. While the genetic and phenotypic diversity of this pathogen species has been explored, especially when concerning pathogenicity, knowledge of the phenotypic diversity using a broad range of ecologically relevant phenotypes is lacking. This thesis uses a combination of comparative analysis, high throughput phenotyping, and experimental evolution to explore the ecological diversity among the RSSC, progressing our understanding at the global (chapter 2), population (chapter 3), and isolate level (chapter 4). Potential causes of this trait variation were also explored, and phenotypic variation was linked with genomic data to reveal genetic mechanisms underpinning traits. Collecting 46 ecologically relevant traits across a collection of RSSC strains revealed that the local environment selects for similar ecological differences within two RSSC species (*R. pseudosolanacearum* and *R. solanacearum*) in chapter 2. In contrast, a lack of genetic variation was observed among a UK population of *R. solanacearum* despite significant phenotypic differences between isolates in chapter 3. Evolving one UK *R. solanacearum* isolate to different abiotic environmental stress conditions within the laboratory (chapter 4) revealed that exposure to stresses increases RSSC diversity, driven by the negative trait correlations observed between the different stresses and metabolic capacity. Furthermore, insertion sequence (IS) movement was found to cause adaptation to environmental stress conditions within the lab, potentially explaining the lack of genetic variation observed within the UK population. A genome-wide association study (GWAS) also revealed genes associated with cold tolerance and rifampicin resistance within the RSSC. Overall, this thesis provides a better understanding of RSSC ecological diversity which can improve our knowledge of the epidemiology of this plant pathogen.



Graphical Abstract: Chapter 2 reveals that the five identified phenotypically distinct 'ecotypes' are spread across the phylogeny, suggesting that the local environment selects for similar ecological differences within two *Ralstonia solanacearum* species complex (RSSC) species, *R. pseudosolanacearum* and *R. solanacearum*. Chapter 3 found significant phenotypic differences, with three phenotypically distinct 'ecotypes', among a UK population of *R. solanacearum* despite large genetic similarities between isolates. Chapter 4 discovered that exposure to stresses (acidity, alkalinity, and salinity) increases RSSC diversity in an evolutionary experiment within the laboratory.

Table of Contents

List of Tables

Chapter 1

List of Figures

Chapter 4

Chapter 5

Acknowledgements

This project, like all good research, has been highly collaborative and therefore, there are many people whom I wish to thank. Firstly, I would like to thank my supervisor Ville for all his guidance and support over the course of this PhD. His advice and encouragement has made me a better scientist, and I couldn't have asked for a better supervisor! He has made the Friman lab an inclusive and enjoyable environment to work in, which is a credit to his hard work and caring personality.

I would also like to mention a special thank you to Martina, who has undertaken the other half of this research project. Your bioinformatic knowledge has been invaluable, and I can't thank you enough for teaching me the world of coding. I am so lucky that we got to start and finish our PhD journeys together, with a small pandemic in between, and I am so grateful that you chose to come to York. I couldn't have asked for a better colleague and friend, and I'll miss our endless chats in the quarantine room!

A huge thanks also goes to all the members of the Friman lab, past and present, for creating such a friendly environment to work in. The lab has both given me great feedback on my work, as well as endless sources of procrastination. There is always someone up for chatting in the office or heading to the pub (mini-Saturday). Thank you to Sophie for the office hot water bottles, Lauri for introducing me to beer, and to Ezra for all the game nights. There are too many to name you all but thank you so much!

Thank you to John, my co-supervisor, for giving me numerous feedback on the world of *Ralstonia,* despite being supposedly retired*.* His endless knowledge has been invaluable to this PhD. Along with John, I would also like to thank the rest of the bacteriology team at FERA for welcoming me into the labs and for all their training. Also, thank you to my co-supervisor Dan, and the Jeffares lab, for giving me the opportunity to present at countless lab meetings and for their advice and help on anything bioinformatics related.

I would also like to thank the PhD community at the Department of Biology for their friendship and fond memories over the past four years. Especially thank you to Katherine for being the best housemate and friend. Thanks so much for looking after our lockdown buddy Ember and for all the movie nights and cocktails.

Furthermore, I'd like to thank my friends and family outside of the university for always believing in me and giving me a much-needed escape from science. Thanks to my grandparents for all their love and support, and to my brothers for always being there, no matter what. I would especially like to thank my parents for their limitless love and encouragement - I couldn't have done this without you!

The White Rose Mechanistic Biology DTP funded this thesis in partnership with an iCASE funded by FERA science ltd.

Declaration by Author

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

1 Chapter 1: General Introduction

1.1 Diversity of organisms

Living things differ in their characteristics or phenotypes, a collection of traits encoded by the genome, known as biological diversity. As well as between species diversity, there is also diversity within the same species, which can lead to splitting into different species if geographical or temporal barriers separate populations for a long enough time. High within species diversity can rise through several adaptive processes. These include niche complementation, which decreases the competition among individuals of the same species by allowing different populations to occupy new environments/niches—for example, altering nutrient choice to partition resources among species, thereby preventing direct competition (Bajic and Sanchez 2020). Secondly, high diversity can be favoured because it improves species survival during natural fluctuations within an environment, such as temporal changes in temperature, nutrient availability and biotic stresses (Horner-Devine et al. 2004). Finally, diversity can increase the evolvability of an organism by increasing the chances of finding and fixing advantageous genotypes (Cordero and Polz 2014).

There are four main genetic mechanisms influencing the diversity of a species. These are: mutation, gene flow, selection and genetic drift (Hanson et al. 2012). Mutations (SNPs, indels) and gene flow (via horizontal gene transfer etc.) produce new genetic variation upon which selection and genetic drift can act. This genetic information can then encode traits, or phenotypes, which are influenced by the environment via interactions between external factors and the genome (Houle et al. 2010). Selection can remove maladaptive traits, and therefore any genetic information, that are detrimental by reducing the number of offspring with these characteristics. While genetic drift acts by chance, randomly increasing and decreasing the number of genetic variants and therefore has a larger impact on smaller populations. In environments where a new genotype encodes a fitter phenotype, this genotype will increase in number while the old genotype will decrease in frequency. Diversity can originate from changes within the genome,

such as mutations and horizontal gene transfer, or can even be regulated by expression networks and feedback loops through phenotypic plasticity. Phenotypic plasticity is common within bacteria, such as within the *Xanthomonas* bacterial species (Timilsina et al. 2020), and various traits are known to be expressed differently depending on the external environment, helping bacteria to utilise different niches and increase the overall fitness of a species (Bull 1987, Smits et al. 2006, Rainey et al. 2011). Evolution can also generate an increase in diversity by increasing the mutation rate, as seen with hyper-mutators (Sniegowski et al. 1997) or by facilitating the incorporation of novel DNA, such as increasing horizontal gene transfer rate (Cordero and Polz 2014).

At the turn of the 21st century, a new era of exploring diversity began. Unlike the 19th and 20th century, where the focus was on exploring the diversity of visible organisms, the diversity of single-cell organisms, made possible through technological advancements like high-throughput sequencing, began (Gibbons and Gilbert 2015). The largest survey on microbial diversity, the Earth Microbiome Project (https://earthmicrobiome.org/), started in 2010 where 5.6 million operational taxonomic units (OTUs) in the first 15,000 samples was discovered, going far beyond diversity estimates of multicellular organisms and highlights the huge diversity of microorganisms on Earth (Gibbons and Gilbert 2015). However, microbial diversity is not as straightforward, as microorganisms undergo rapid evolution/speciation and molecular phylogenies can become confounded due to the horizontal swapping of genes across microbes from distant taxa (Gibbons and Gilbert 2015). Therefore, scientists often rely on highly conserved genes, such as the 16S rRNA, to explore the genetic diversity of bacteria and other microorganisms (Gibbons and Gilbert 2015). However, finer scale distinction between individuals, along with contextualised information regarding the environment (e.g. ecotypes), is needed to observe interesting patterns among microbial species diversity (Gibbons and Gilbert 2015) and to understand better how natural environments shape and maintain microbial diversity over time and space. Traits are often used to characterise and distinguish bacteria and are likely a driving factor in microbial community composition, yet little is known about the trait diversity of most microbes (Weimann et al. 2016).

It is generally accepted that a balance between niche, such as pH/temperature driving microbial dominance (Zeglin 2015), and neutral (e.g. stochastic dispersal) processes will influence the diversity of microbial ecosystems (Gibbons and Gilbert 2015). A stable environment allows organisms to fine-tune their phenotypes to suit a specific condition, becoming specialists to this environment. However, environments are rarely constant, instead being extremely heterogeneous, constantly fluctuating in time and space (Horner-Devine et al. 2004, Hanson et al. 2012). In these environments, diversity can be selected and maintained through frequency-dependent selection, or alternatively generalists can be favoured (Bull 1987, Smits et al. 2006, Rainey et al. 2011). However, adaptation to one environment is typically associated with fitness loss in another (Elena and Lenski 2003), due to negative trait correlations called trade-offs. Such trade-offs can constrain the range of phenotypes open to organisms (Ferenci 2016) and can therefore limit the diversity of an organism. Therefore, in the absence of other forces, environmental heterogeneity should lead to locally adapted populations, where the local population is the most fit in its local environment (Croll and McDonald 2017). However, different specialists can also coexist within the same environment, fluctuating in frequency within the environment. Research has found that for some species of bacteria the presence of different niche specialists, or ecotypes, maintains high species diversity rather than generalist adaptation to a variety of environmental conditions (Jezbera et al. 2011, 2013, Larkin and Martiny 2017). Other studies have also found that environmental differences, such as geographical distribution and host range, help maintain this high diversity of microbial species (Horner-Devine et al. 2004). Furthermore, high genetic differences and plasticity can also result in high phenotypic diversity, such is the case for *Xanthomonas* bacterial species (Timilsina et al. 2020). However, more information on a broad range of traits is needed to build our understanding of how high diversity is maintained among microorganisms.

1.2 Trait Correlations

Trait correlations can be recognised when two or more traits have a positive or negative relationship among individuals within a population (Saltz et al. 2017).

For example, a relationship between two traits can be linear, convex or concave, to name a few possibilities (Stearns 1989), and the strength of correlation between two traits can be statistically measured for example by using Pearson's correlation coefficient (Saltz et al. 2017). Natural selection acts on traits (otherwise known as phenotypes). However, phenotypic correlations can determine the pattern of variation present for natural selection to act on (Stearns 1989), either increasing the evolvability of an organism in the case of positive trait correlations, or limiting adaptive potential due to negative trait correlations (also known as trade-offs). Trait correlations are widespread among traits and organisms. Therefore, understanding how they evolve is essential to predicting their evolutionary effects on the phenotype (Saltz et al. 2017). Negative correlations can constrain adaptation of an organism, especially if the other trait cannot be discarded. Furthermore, they can even result in the loss of the other trait, becoming specialist to a certain environment. This trade-off will then only become apparent when this organism moves into a new environment, or when selective pressures within the environment changes or are removed (Stearns 1989). A multitude of factors can cause trait correlations, and the response of an organism to selection depends on its genetic variation (Stearns 1989). Therefore, genetic interactions can determine whether and in what direction a response to selection will occur (Stearns 1989).

Most traits show polygenic, not mendelian, inheritance, with genome wide association studies (GWAS) frequently finding no or very few large effect variants, even for highly heritable traits (Saltz et al. 2017). Multiple genetic variants can interact to produce a single phenotype, which indicates that the impact of new genetic variation on an organism's phenotype will depend on the existing genetic background, a phenomenon known as epistasis. If multiple genes influence most traits, many genetic variants must therefore influence more than one trait as there are many quantitative traits and a finite number of genetic variants (Saltz et al. 2017). This notion of single genes affecting multiple traits is known as pleiotropy (Foster et al. 2004). Trait correlations can also occur due to linkage disequilibrium (LD), where genes are physically close and therefore linked and subsequently commonly inherited together. However, while trait correlations through pleiotropy will not break down over time, correlations caused by linkage disequilibrium (LD)

will (Saltz et al. 2017). That being said, if LD arises randomly, all outcomes of linkage should be equally common, producing no trait correlation at the phenotypic level (Saltz et al. 2017), suggesting that LD can still be an important factor to take into account for trait correlations.

Pleiotropy can be beneficial by increasing the fitness of an organism through positive trait correlations, such as when beneficial mutations in *Escherichia coli* selected in a glucose-limited environment were also beneficial in the presence of other sugars (Schenk et al. 2015). However, they can also cause negative trait correlations, also known as antagonistic pleiotropy or simply trade-offs, where the different phenotypic effects from the single genetic locus have opposite effects on the organism's fitness. One example is *E. coli* adapting to the enzyme TEM-1 b-lactamase, causing an acquired cefotaxime antibiotic resistance but also reducing resistance to another antibiotic (ceftazidime) (Schenk et al. 2015). If there were no trade-offs, then selection would drive all traits correlated with fitness to limits imposed by history and design. However, we find that many traits are maintained well within those limits, and therefore trade-offs must exist (Stearns 1989).

Trade-offs prevent organisms from achieving maximum fitness due to finite resources (energy, time, molecules etc.) allocated for one trait reducing investment in the other trait (Saltz et al. 2017). An example is sharing RNA polymerase and other transcriptional machinery, which are regulated by transcription factors such as the sigma factor in *E. coli* (Ferenci 2016). Another way in which a trade-off can occur is if a protein has multiple functions, such as a porin that allows nutrients and other molecules to enter a cell, which then adapts preferentially for a particular function. For example, reducing porin size to reduce antibiotic uptake also causes a reduction in its ability to uptake different metabolites (Ferenci 2016). Pleiotropy is a major constraint on evolution because adaptive change in one trait may be prevented because it would compromise other traits affected by the same genes (Foster et al. 2004). However, there are benefits to pleiotropy, such as facilitating cooperation in *Dictyostelium discoideum* amoeba. Here differentiation of some amoeba into dead pre-stalk cells is required to hold reproductive cells, however cheaters that lack the *dimA* gene*,* which is required to receive the signalling molecule to differentiate, are also excluded from spores due to genetic linkage

(Foster et al. 2004). Many trait correlations have a genetic basis, however, predictions about the evolutionary dynamics of trait correlations go beyond heritability. Scientists are often interested in why traits are correlated, if they evolved under selection, and if selection or drift can change the magnitude or direction of trait correlations rather than the exact genetic basis behind them (Saltz et al. 2017). Trait correlations also may not directly reflect trade-offs or true pleiotropy. Trait correlations can be due to genetic linkage, such as linkage disequilibrium (which usually has little effect on long-term evolution), and can also be due to the correlation of both traits of interest with a third unmeasured trait (Sgrò and Hoffmann 2004).

Many species in nature experience varying environmental conditions, which play a role in shaping evolution by altering the genetic architecture of traits (Sgrò and Hoffmann 2004). Therefore, trait correlations can be highly dynamic and flexible, with the relationship between two traits easily changing from positive to negative over the commonly encountered environmental range (Stearns 1989). Phenotypic plasticity is a mechanism used to modulate the expression of genetic variants across environments, which could also play a part in this dynamic aspect of trait correlations (Stearns 1989). For example, genetic correlations among traits will often be positive when resources are abundant, while negative correlations reflecting trade-offs may only be apparent when fitness is measured in resource-poor environments (Sgrò and Hoffmann 2004). It is also thought that genotype-by-environment interactions may be stronger under extreme environmental conditions (Lannou 2012).

Furthermore, within nature, multiple traits are usually under selection. Fisher's geometric model for trait correlations (Fisher 1930) assumes that the rate of adaptation is inversely related to the number of traits under selection (McGee et al. 2016). This is because under complex environments, with multiple selection pressures acting on various aspects of the phenotype, trade-offs are more likely to emerge due to competing functions (McGee et al. 2016). This can be seen in lab experiments, such as when bacteriophage were evolved in simple (one trait) and complex (two traits) selection pressures which found smaller improvement rates for mutations fixed in complex conditions (McGee et al. 2016).

1.3 Introducing the *Ralstonia solanacearum* species complex (RSSC)

The *Ralstonia solanacearum* species complex (RSSC) is a diverse group of gram-negative betaproteobacteria that are the causative agent of bacterial wilt disease in a wide range of plants. Some members are also able to cause host-specific diseases, including potato brown rot, Moko disease and blood disease of banana, and Sumatra disease of cloves (Hayward 1991, Safni et al. 2018, Bragard et al. 2019). A species complex is a group of closely related species that are extremely similar, so much so that the boundaries between them can become unclear (Fegan and Prior 2005). DNA-DNA homology of RSSC strains has revealed that the relatedness between members is often less than 70%, commonly considered as the threshold between one species and another, despite similarities in phenotype and disease symptoms (Fegan and Prior 2005). Plant pathologists often relate plant pathogenic species to the disease symptoms they cause and to their host range. All members of the RSSC share the ability to grow in xylem vessels and stem apoplasts of host plants and cause classical wilting symptoms in their host plants (Lowe-Power et al. 2018). The lack of consistency in pathogen host range and behaviour between closely related species in the RSSC, and similarities between distantly related members, justifies the need to call *Ralstonia solanacearum* a species complex as opposed to multiple separate species (Sharma et al. 2022). High amounts of genetic exchange have also been observed between members of the RSSC (Wicker et al. 2012), further supporting the concept that *Ralstonia solanacearum* is a species complex. Gillings and Fahy first used the term “species complex” to refer to *R. solanacearum* to reflect the large phenotypic and genetic variation apparent within this species in 1994 (Fegan and Prior 2005). Since then, Taghavi *et al.* added two other closely related organisms to the complex, the banana blood disease bacterium and the then-called *Pseudomonas syzygii* bacterium, since 16S rDNA sequence analysis revealed they fell within the diversity of the RSSC (Taghavi et al. 1996).

RSSC is one of the most destructive bacterial plant pathogens worldwide with an extensive host range, including at least 175 different plant host species (EPPO 2022), containing many important food crops such as potato (*Solanum*

tuberosum), tomato (*Solanum lycopersicum*) and banana (*Musa* spp.) (Hayward 1991, Mansfield et al. 2012, Bragard et al. 2019). Furthermore, the RSSC can infect various wild hosts (for example, *Solanum dulcamara* and *Urtica dioica* in Europe) (Wenneker et al. 1999), many ornamental plants (including *Pelargonium* spp.), and some trees (including *Eucalyptus* spp.) (Hayward 1991, Bragard et al. 2019). The host range of RSSC is still growing, with new hosts being identified constantly (EPPO 2022). For example, in 2017 roses in the Netherlands were reported as infected with RSSC for the first time (Stevens et al. 2018), blueberries in Florida were discovered to be infected in 2016 (Bocsanczy et al. 2019) and in Korea peanuts were first found to be infected in 2021 (Choi et al. 2022). The *Ralstonia solanacearum* species complex has one of the most diverse host ranges and, unsurprisingly, has one of the broadest geographic distributions of any plant pathogenic bacterium (Safni et al. 2018), currently present within 123 countries (EPPO 2022). This pathogen can be found in countries all around the world from tropical countries, like Indonesia, to temperate climates, such as Sweden and the UK (Hayward 1991, Bragard et al. 2019, EPPO 2022).

1.4 History and classification of the *Ralstonia solanacearum* species complex (RSSC)

The *Ralstonia solanacearum* species complex was initially known as *Bacillus solanacearum* (Smith 1896) due to its elliptical cell morphology. Since then, RSSC has undergone many different classification changes; being placed in the genus *Bacterium* in 1898, then *Pseudomonas* in 1914, then placed under the *Phytomonas* and *Xanthomonas* genus before returning to the *Pseudomonas* genus in 1948. RSSC then remained in the *Pseudomonas* genus until 1992 where it was classified under *Burkholderia* (Paudel et al. 2020). Finally, it was decided that a new genus called *Ralstonia* was to be created to describe *Burkholderia solanacearum* and *Burkholderia picketti* due to their similarity to one another and differences with other *Burkholderia* species based on DNA homology (Yabuuchi et al. 1995).

Plant bacteriologists have been trying to come up with a way to subclassify members of the RSSC due to its huge diversity, both phenotypically and genetically.

Historically (from 1960s) RSSC strains were divided into races and biovars. First they were divided into five races based on their host range (Hayward 1991, Paudel et al. 2020), and then further subdivided into five biovars according to their ability to oxidise three disaccharides (sucrose, maltose, lactose) and three hexose alcohols (mannitol, sorbitol, dulcitol) (Hayward 1991). Eventually, a new system of subclassifying the RSSC arose in 2005 when sequencing technologies advanced and Fegan and Prior 2005 proposed a classification scheme based on sequence analysis of the internal transcribed spacer region, the endoglucanase (*egl*) gene, and the *hrpB* gene (Fegan and Prior 2005). This resulted in four phylotypes corresponding to the clustering of strains into 4 groups based on their sequences (Wicker et al. 2007). Phylotype II, the most diverse phylotype, also had two recognisable distinct clusters, referred to as phylotypes IIA and IIB (Genin and Denny 2012). These four phylotypes roughly correspond to their geographical origin with Phylotypes I, II and III predominantly originating from Asia, Americas and Africa, respectively, and phylotype IV primarily originating from Indonesia (Safni et al. 2018). Each phylotype can also be further subdivided into over 20 sequevars based on differences in the sequence of a portion of the endoglucanase (*egl*) gene (Wicker et al. 2007). Since then, sequencing technology has become more advanced which prompted Safni *et al.,* in 2014 to propose a regrouping of the *R. solanacearum* species complex into three separate species, *R. pseudosolanacearum*, *R. solanacearum* and *R. syzygii*, based on significant variations in the whole genome (Safni et al. 2014, Prior et al. 2016, Kumar et al. 2018, Stevens et al. 2018). These 3 species correlate to the old phylotype groupings, with *R. solanacearum* comprised of phylotype II strains, *R. pseudosolanacearum* of phylotype I and III and *R. syzygii* comprised of phylotype IV (Safni et al. 2014) (see table 1.1 for an overview on RSSC classification). Separation of these three species is thought to date back to the geological separation of the continents as members of the RSSC can be found in virgin jungle soils in both Asia and the Americas (Hayward 1991), with the three separate species reflecting the true evolutionary lineage that arose when ancestors became geographically separated (Genin and Denny 2012). It is also hypothesised that *R. syzygii* was the first to diverge, followed by *R. solanacearum and R. pseudosolanacearum* separating from one another, leading to theories that the tropical RSSC pathogen

originated in Indonesia (where *R. syzygii* is predominantly found) (Wicker et al. 2012). Throughout this thesis, when referring to each species separately they will be named by their assigned species name (e.g. *Ralstonia pseudosolanacearum*), while the term RSSC will refer to all three species as a collective.

Table 1.1: *Ralstonia solanacearum* species complex (RSSC) classification. *Adapted from Garćia et al. 2019,and Paudel et al. 2020.*

Species	Phylotype	Sequevar	Race	Biovar	Geographical Origin
Ralstonia pseudosolanacearum	I	12, 14, 16, 18	1, 4, 5	3, 4, 5	Asia
	III	19, 20, 21, 22, 23	1	1, 2-T	Africa and surrounding islands
Ralstonia solanacearum	IIA	1, 2, 3, 4, 5, 6, 7	1, 2, 3	1, 2, 2-T	America
	IIB				
Ralstonia syzygii	IV	8, 9, 10, 11	1	1, 2, 2-T	Indonesia

Ralstonia syzygii is thought to have the highest genetic diversity out of all three RSSC species (Wicker et al. 2012) and can also be further divided into 3 subspecies, *R. syzygii* subsp. *syzygii*, *R. syzygii* subsp. *celebesensis* and *R. syzygii* subsp. *indonesiensis* (Safni et al. 2014, 2018). *Ralstonia syzygii* subsp. *syzygii* is a pathogen that causes Sumatra disease of clove trees, as well as some species of *Myrtaceae* found in forests in Indonesia and has only been reported in Indonesia. *Ralstonia syzygii* subsp. *celebesensis* causes Banana Blood Disease (BBD), a disease that is very similar to Moko disease of bananas caused by *R. solanacearum* strains. *R. syzygii* subsp. *celebesensis* has a larger host range than subspecies *syzygii,* but not as large as *R. solanacearum,* and is not pathogenic on *Solanum lycopersicum* (tomato) and *Solanum melongena* (aubergine) seedlings (Safni et al. 2018). Finally, *R. syzygii* subsp. *indonesiensis* causes diseases in several solanaceous plants in Asia,

including *Solanum tuberosum* (potato), *Solanum lycopersicum* (tomato) and *Capsicum annuum* (chilli and bell peppers), with bacterial wilt disease symptoms identical to wilting caused by *R. solanacearum* species (Safni et al. 2018, Bragard et al. 2019), however *R. syzygii* subsp. *indonesiensis* cannot infect potatoes in temperate conditions unlike some *R. solanacearum* strains.

Two strains of *R. syzygii* are thought to be transmitted by insects*. Ralstonia syzygii* subsp. *syzygii* is an xylem-restricted bacteria whose main mode of dispersal is by insect vectors that feed on xylem sap (Safni et al. 2018) and *R. syzygii* subsp. *celebesensis* is also thought to be transmitted by insects as this bacterium has been found within some insects. However, the transmigration of people from Java to less populated islands in the country appears to be the main cause of the spread of *R. syzygii* subsp. *celebesensis* and it is thought to have originated on Selayar Island near Sulawesi and spread to Java in the late 1980's .

Ralstonia pseudosolanacearum can infect a wide range of hosts, with the most common hosts being *Solanum* species, including *Solanum tuberosum* (potato) and *Solanum Lycopersicum* (tomato)*, Casuarina equisetifolia* (she-oak), and *Nicotiana* species (Bragard et al. 2019). Uniquely, the *R. pseudosolanacearum* species also comprises strains that can infect ginger (*Zingiber officinale*) and mulberry (*Morus* spp.) (EPPO 2018). There is also a clear geographical separation among *R. pseudosolanacearum* strains with phylotype I strains originating from Asia and phylotype III strains originating from Africa (Safni et al. 2014).

Ralstonia solanacearum also has a wide host range, with the main host plants being *Solanum* species or other members of the family *Solanaceae* (including the cold tolerant *S. tuberosum* brown rot causing strain), *Anthurium, Heliconia* and *Musa* species (banana) (Bragard et al. 2019). The *R. solanacearum* species complex also includes causative strains of Moko disease, infecting both banana (*Musa* spp.) and *Heliconia* species (EPPO 2018). It also includes a causative agent of potato brown rot disease, previously known as race 3 biovar 2 or phylotype IIB sequevar 1 (Fegan and Prior 2005, Safni et al. 2014). These are strains in the RSSC, a tropical pathogen, that is adapted to lower temperatures, therefore causing a serious threat to temperate agriculture (Williamson et al. 2002, EPPO 2018). Before the 1970s, RSSC was thought to be a solely tropical or subtropical pathogen until a series of

brown rot cases across Europe was reported; Sweden in 1972, Belgium in 1989, and the Netherlands, UK, and many other European countries since 1992 (Janse 1996). The origin of *R. solanacearum* is thought to be from the Americas, where it is speculated that this cold adapted strain evolved at higher altitudes in south America before being spread across the world through the potato trade (Hayward 1991, Wicker et al. 2012, Safni et al. 2014). The clonal cold tolerant strain of *R. solanacearum,* a potato brown rot strain, originally described as pathogenic on potato and tomato, but not other solanaceous crops (Genin, 2010), was later discovered to have retained their ability to be virulent on a wider range of plant hosts at higher temperatures, including non-solanaceous *Pelargonium* species (Cellier and Prior, 2010).

1.5 Controlling the spread of *Ralstonia solanacearum* species complex (RSSC)

The RSSC can infect many different species of crops and therefore can cause huge economic loses, with an estimated amount of $1 billion worldwide lost annually due to the infection of potato crops alone because of RSSC (Mansfield et al. 2012). Furthermore, the world's population is expected to increase from 7.5 billion (2017) to 9.8 billion by 2050 (United Nations 2015), increasing global demand for food. To meet these goals a large increase in crop yields is needed, and factors, such as global warming where crop yields are predicted to decline with increasing temperatures (Hijmans 2003, Peng et al. 2004) will make this harder. On top of climate change, pests, and diseases such as RSSC also decrease crop yields even further. It is thought that combined they are responsible for the loss of around 11% to 59% of crops worldwide, with pathogens alone causing up to 24% of crops to be destroyed (Oerke 2006). This means that the control of pathogens, like RSSC, is vital to keep up with global food demands. Due to the danger RSSC poses on our food supply, it has been classed as a quarantine pest in many countries, including the EU (EPPO 2022) and a bioterrorism select agent in the US (USDA 2020) (Cellier and Prior 2010).

There are currently a variety of methods available to attempt to control the spread of RSSC, such as using resistant cultivars of crops. One example of this is the expression of the *Arabidopsis thaliana* PRR elongation factor Tu receptor gene, which significantly reduced tomato susceptibility to RSSC (Kunwar et al. 2018). However, natural host resistance is very limited in modern crop cultivars and RSSC is extremely diverse, meaning resistance to all strains in the RSSC is unrealistic (Champoiseau et al. 2009, Bragard et al. 2019). Another control method is the use of crop rotations with non-host plants, such as wheat, sweet potato, maize, millet, sorghum, or carrots. This can reduce the incidence of bacterial wilt disease by 20-26% compared to monocultured host crops such as tomato (Adhikari and Basnyat 1998).

There are also a number of soil treatments that can be used to control RSSC numbers, such as calcium carbonate, which can significantly reduce RSSC numbers by increasing the pH of the soil (Jiang et al. 2017). A number of chemical pesticides like algicides, fumigants and plant resistance activators (e.g. vilidamycin A) (Yuliar et al. 2015) are also used to reduce bacterial wilt disease symptoms. However, there has been a large reduction in pesticides allowed on the EU market due to public health concerns (Kleter et al. 2009) and a significant increase in resistance towards pesticides has been observed, rendering them useless (Zhou et al. 2012).

One alternative to chemicals are biocontrol agents. These can include the addition of antagonistic bacteria to the field, like *Bacillus amyloliquefaciens* (Singh and Kumar Yadav 2016) or *Pseudomonas brassicacearum* J12 strain, that produces antimicrobial compound 4-diacetylphloroglucinol (2,4-DAPG) (Zhou et al. 2012). Both significantly reduce bacterial wilt disease symptoms *in planta* (Zhou et al. 2012, Singh and Kumar Yadav 2016). Another potential biocontrol agent is using a combination of bacteriophage predators as a treatment to control RSSC, which have also been shown to be effective at reducing bacterial wilt disease symptoms (Wang et al. 2019, Doan et al. 2022).

Despite all this, pathogen eradication of agricultural fields remains challenging due to the large host range of RSSC, accredited to its high diversity, including many weed species, such as *Solanum dulcamara* (woody nightshade) in Europe (Wenneker et al. 1999). Another reason why RSSC is so difficult to control is

that infected hosts can remain asymptomatic, particularly in temperate regions (Bragard et al. 2019) and RSSC can persist in a non-host environments, such as within water sources (and potential irrigation sources), within soil communities, and on farm equipment for a debated (from a few years to 14 days) period of time (Bragard et al. 2019). These environments can act as natural reservoirs for the pathogen to reside in, particularly when wild host plants (e.g. *Solanum dulcamara* in Europe) are present, remaining unaffected by field treatments and can later spread from these reservoirs onto crop fields causing further outbreaks. Due to this, it is unlikely that RSSC can be eradicated from countries, as treatment of river sources and complete removal of wild hosts is unrealistic (Williamson et al. 2002, Elphinstone and Matthews-Berry 2017). Therefore, current control methods include preventing the spread of RSSC even further as the pathogen is mainly dispersed through human trade of contaminated crops and seeds (van der Gaag et al. 2019, Bragard et al. 2019). This is done through regular testing of irrigation water, seed potatoes and other plant material, to prevent use of contaminated water sources and transmission of contaminated crops to other countries (Elphinstone and Matthews-Berry 2017).

1.6 *Ralstonia solanacearum* species complex (RSSC) genetics

The first *Ralstonia solanacearum* species complex (RSSC) strain to be sequenced was the *R. pseudosolanacearum* strain GMI1000 from French Guyana in 2002 (Salanoubat et al. 2002). This revealed that the RSSC genome is bipartite, containing a chromosome (3.7 Mbp) and a separate 'megaplasmid', which is a similar size to the chromosome (2.1 Mbp), with a combined overall length of around 5.8 Mbp (Genin and Boucher 2002). This 'megaplasmid' appears to carry genes associated with strain specific lifestyles , as well as includes genes needed for the overall fitness of the pathogen (Genin and Boucher 2002, Genin and Denny 2012). The 'megaplasmid' is also thought to still be in the process of acquiring new functions through duplications, genomic rearrangements and translocation of genes from the chromosome (Genin and Boucher 2002, Remenant et al. 2010). There only seems to be minor differences in gene expression and regulation of gene expression between both replicons, suggesting that they have a shared evolutionary history

which indicates that the 'megaplasmid' has not been recently acquired (Coenye and Vandamme 2003). Although it has been coined as a 'megaplasmid' in RSSC research, this evidence points towards the fact the second RSSC replicon acts more like another chromosome rather than a plasmid.

Since the first genome of RSSC was sequenced, many other RSSC strains have since been sequenced, with 316 RSSC sequences available on NCBI in December 2022 (National Center for Biotechnology Information (NCBI) 2022). This data has revealed that RSSC strains share a conserved core genome that is presumably essential for their common biology, such as colonising plant xylem vessels and causing wilting symptoms (Ailloud et al. 2011). The number of strain specific genes is variable (Remenant et al. 2010) but each strain appears to exhibit highly diverse genetic content (Ailloud et al. 2011), which is one of the reasons why *R. solanacearum* is considered as a "species -complex" (Kumar et al. 2018). Small plasmids have also been detected but so far in only three strains, CMR15, PS107 and T78 (Remenant et al. 2010, Genin and Denny 2012, Cho et al. 2019). Megaplasmids of *R. syzygii,* especially from subgroup *celebesensis,* are significantly smaller than those strains classified as *R. solanacearum* and *R. pseudosolanacearum.* This could reflect their narrow host range compared to other strains (Remenant et al. 2011). It is also thought that the *R. syzygii* subsp. *celebesensis* genome carries much more unique genetic information from phages than other RSSC strains sequenced (Remenant et al. 2011).

Genome analysis has also revealed that recombination has played a major role in RSSC genome evolution as a large number of genomic islands surrounded by mobile elements suggests horizontal gene acquisition (Remenant et al. 2010, Peeters et al. 2013, Geng et al. 2022). It is also thought that horizontal gene transfer has enhanced the bacteria's aggressiveness on tomatoes and is responsible for the large diversity among RSSC strains required to evade plant immune responses and infect new hosts (Peeters et al. 2013). Research has shown that DNA blocks, up to 30 kb and 33 genes, can be transferred between RSSC strains and that multiple DNA acquisitions along the genome can occur in a single recombinant strain during a single horizontal gene transfer (HGT) event (Guidot et al. 2009). This suggests that

RSSC is capable of rapid adaptation to novel ecological niches via transfer of novel genes from the environment or other organisms.

1.7 Introducing UK *Ralstonia solanacearum:* a potato brown rot pathogen

The first recorded case of RSSC being present in the UK was in Oxfordshire in 1992 (Parkinson et al. 2013, Elphinstone and Matthews-Berry 2017). This outbreak was shown to be caused by the cold tolerant strain of *R. solanacearum*, also known as race 3 biovar 2 or the phylotype IIB sequevar 1 of *R. solanacearum*, also found to cause potato brown rot across many countries in Europe in recent decades (Fegan and Prior 2005, Safni et al. 2014). This original UK outbreak has since been linked to contaminated river water that was used as an irrigation source for the field (Parkinson et al. 2013). Since 1992 there have been seven other outbreaks of *R. solanacearum* in the UK; in 1995 (infected potato), 1997 (infected tomato), 1998 (infected tomato), 1999 (infected potato), 2000 (infected potato), 2005 (infected potato) and 2010 (infected potato) (Elphinstone and Matthews-Berry 2017). All except the 2010 outbreak, which can be linked to infected imported seed potatoes, were associated with contaminated water sources, suggesting establishment of *R. solanacearum* in some watercourses across the UK (Elphinstone and Matthews-Berry 2017).

River water can become contaminated through run-offs from infected potato crops at processing plants (Janse 1996), and the prevalence of *R. solanacearum* in the UK rivers was linked with infection of a wild host, *Solanum dulcamara*, which lives along riverbanks (Parkinson et al. 2013). *Solanum dulcamara,* otherwise known as woody nightshade, is an important wild perennial host for *R. solanacearum* in not just the UK but other countries in Europe as well, and its presence along rivers is constantly correlated with contaminated water sources (Wenneker et al. 1999, Bragard et al. 2019). Wilting symptoms in *S. dulcamara* are extremely rare, unless temperatures exceeds 25°C or if inoculum levels are high, and infected plants normally remain asymptomatic (van der Gaag et al. 2019). *Solanum dulcamara* grows along riverbanks with their roots in the water acting as a reservoir for bacteria to disseminate into the water, opening new opportunities for infecting primary hosts if this river is then used for irrigation of

crops (van der Wolf et al. 1997, Wenneker et al. 1999). It is thought that *R. solanacearum's* ability to survive the cool winter temperatures in rivers is primarily due to the presence of this weed and the bacteria colonising the roots over winter (Genin and Boucher 2002, Champoiseau et al. 2009). It is also thought that *R. solanacearum* may enter a viable but non-culturable (VBNC) state in rivers at low temperatures helping its survival throughout the winter (Genin and Boucher 2002), however evidence for this is limited.

RSSC is detected in river water by culturing the bacterium on semi-selective SMSA media, then conducting either plant infection assays, indirect immunofluorescent-antibody staining (IFAS), or PCR detection of RSSC DNA (Elphinstone et al. 1996, EPPO 2018). *Ralstonia solanacearum* can be detected in low numbers, as few as two viable pathogen cells per ml of river water, however population numbers vary in UK rivers depending on the time of year, being the highest when temperatures are above 15°C, usually between June and September, and falling to undetectable levels in the winter (Elphinstone and Matthews-Berry 2017). In contrast, *R. solanacearum* can be detected all year round in the stems of *S. dulcamara* along riverbanks of known contaminated watercourses (Elphinstone and Matthews-Berry 2017), further supporting this plant's role in helping *R. solanacearum* persist in UK rivers despite cold winters. Currently in the UK, regular testing of rivers is conducted and irrigation of host crops with contaminated water is prohibited to try and control the spread of the pathogen and prevent further outbreaks (Elphinstone and Matthews-Berry 2017).

A few studies have been conducted focussing on European *Ralstonia solanacearum* diversity (van der Wolf et al. 1997, Timms-Wilson et al. 2001, Stevens and Van Elsas 2010, Cruz et al. 2012, Parkinson et al. 2013, Caruso et al. 2017), of which one focusses on the UK (Parkinson et al. 2013). These studies mainly conclude that the cold tolerant European strains are genetically very clonal compared to other strains in the RSSC, consistent with a recent invasion of the temperate region agreeing with the reports of *R. solanacearum* outbreaks in Europe, the first identified one being in Sweden in 1972 (Caruso et al. 2017).

1.8 *Ralstonia solanacearum* species complex (RSSC) life cycle

There are 2 key stages in RSSC's life cycle: As a pathogen the first stage is the infection of host plants, including production of virulence factors, evading plant immune responses and general survival within the host. The second stage is survival/persistence within the environment, such as in the soil and river water, during transmission from host to host. RSSC will face very different selection pressures depending on the stage of the life-cycle and what primary host they are in (Genin 2010).

RSSC bacterium locates plant hosts by sensing and chemotaxis, moving attracted by root exudates using flagella motility (Yao and Allen 2007). Attachment (reversible and irreversible) then occurs to plant surfaces via pili leading to the formation of microcolonies at root elongation zones (Kang et al. 2002). The pathogen often enters plants through wounds and natural openings in the root, and once inside the plant, moves towards the xylem (Hayward 1991). *Ralstonia solanacearum* species complex is well adapted to living in the xylem which contains sugars, amino acids and organic acids that supports bacterial growth, while the xylem flow constantly introduces fresh nutrients and removes waste products (Lowe-Power et al. 2018). In the xylem, cells grow to aggregates in a biofilm matrix and fill vessels obstructing water flow. This creates the classic wilting symptoms linked to RSSC and ultimately kills the plant (Lowe-Power et al. 2018). Bacterial wilt symptoms tend to appear after pathogen populations reach 10^9 colony forming units per ml, as pathogenicity factors are activated using density-dependent quorum sensing molecules, which activates the main virulence gene network via the *PhcA* gene (Genin and Boucher 2002, Mansfield et al. 2012, Lowe-Power et al. 2018). Pathogenic RSSC strains can also secrete exopolysaccharide and pectin degrading enzymes that aid colonisation of the xylem vessels and can also lead to plant death via degradation of the cell walls (Hayward 1991, Lowe-Power et al. 2018). After the primary host dies, the pathogen returns to the soil or water where they may persist until infecting another plant continuing the infection cycle (Genin 2010). RSSC can survive long periods within perennial hosts, such as RSSC tolerant plant *Solanum dulcamara,* at moderately high levels without causing symptoms

(van der Gaag et al. 2019, Bragard et al. 2019) (see figure 1.1 for RSSC life cycle schematic).

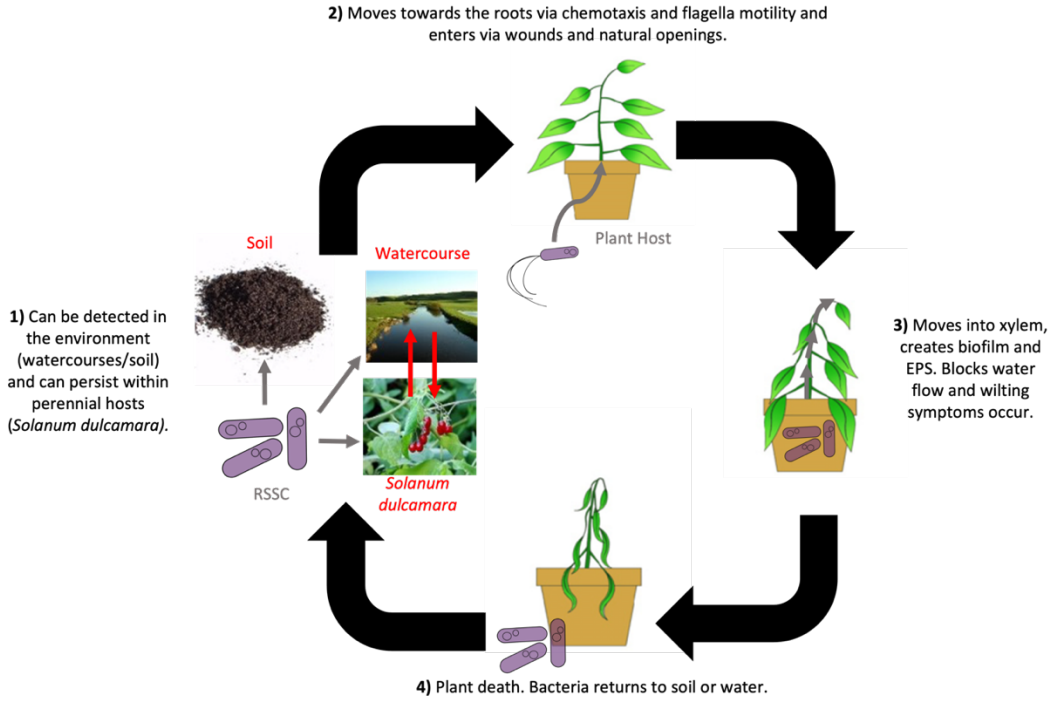


Figure 1.1: *Ralstonia solanacearum* species complex (RSSC) life cycle schematic. (1) *Ralstonia solanacearum* can be detected within environments (soil or watercourses), and mainly persists among asymptomatic perennial hosts, such as *Solanum dulcamara* over winter in the UK. (2) The bacterium can move towards host plans via flagella motility and chemotaxis and is able to enter root tissues through wounds and natural openings to invade the plant vascular system. (3) Once within the xylem, wilting symptoms are associated with bacterial multiplication and production of exopolysaccharides. (4) Disease can lead to plant death, depending on host (i.e. degree of plant resistance), environmental factors and aggressiveness of the strain, where the pathogen can then return to the environment. *Adapted from Genin 2010.*

1.8.1 RSSC virulence factors

There are many traits that are specific towards RSSC's lifestyle as a pathogenic bacterium, known as virulence traits. Virulence factors in RSSC bacterium are controlled by a complex quorum sensing regulatory system, where cell density is sensed, and genes are switched on or off depending on high or low density. *PhcA* is a gene involved in quorum sensing in RSSC, which is in turn controlled by an autoinducer, quorum-sensing signalling molecule, 3-hydroxypalmitic acid methyl ester (3-OH PAME). When extracellular 3-OH PAME is above 5nM, which correlates to high cell densities, active *PhcA* gene expression is

triggered, which then activates several virulence genes, exopolysaccharide (EPS) biosynthesis, *Pme* and *Egl* exoprotein production, as well as represses other genes such as those involved in motility, polygalacturonase and siderophore production (Genin and Boucher 2002). This implies a difference in genes that are active during low cell density during early infection and high cell density later in the infection of the plant when *PhcA* is active. Other complex systems also regulate other genes in RSSC which highlights the complexity of *R. solanacearum* virulence and host specificity (Genin and Boucher 2002).

Members of the RSSC need to enter the host and move into the xylem using flagellar motility which is often lost within the host potentially due to immune invasion (Álvarez et al. 2010, Deslandes and Rivas 2012, Meng 2013). However, twitching motility (i.e., movement via flagella-independent methods such as the type 4 pili) is normally maintained. Twitching motility is known to increase virulence on host plants and promotes auto-aggregation and biofilm production along the host xylem vessel wall (Genin and Denny 2012).

Once inside the plant RSSC needs to overcome host immune defences. To avoid these immune defences RSSC has a large variety and number of effector proteins secreted by the type 3 secretion system (T3SS). Plants' immune response relies on recognising PAMPs/MAMPs (pathogen/microbe-associated molecular patterns) resulting in PTI (PAMP-triggered immunity) (Deslandes and Rivas 2012). Pathogenic bacteria can adapt in response to PTI, changing structure and therefore recognition by plants. Additionally, pathogens use T3SS to evade plant immune response by secreting type 3 effector (T3E) proteins inside plant cells to supress immune response (Deslandes and Rivas 2012, Genin and Denny 2012). However, this can then trigger ETI (effector-triggered immunity) as some T3Es are recognised by plants R (resistance) genes (Deslandes and Rivas 2012). Ailloud *et al.*, also found that compared to other virulence factors the type 3 effectors (T3E) family of RSSC exhibits extremely high plasticity, with 113 T3Es in the pan genome and only 14 are present in every sequenced strain (*RipG5, RipB, RipW, RipAC, RipAB, RipR, RipE1, RipAM, RipAN, RipAY, RipAJ, RipF1, Rip A1* and a *PopC*-like effector), which form the core 'effectome' of RSSC (Ailloud et al. 2011). However, this study was mainly conducted on *R. solanacearum* and not the other 2 species in the complex so we

may find an even smaller core genome and larger amount of strain specific T3Es with further studies. The large range of T3SSs that RSSC have is thought to be due to its coevolutionary arms race against its numerous hosts (Genin 2010). Several agricultural pathogens show strong signatures of diversifying selection acting on their effector repertoire to avoid recognition by the prevalent host genotypes (Croll and McDonald 2017), which we see among specialist species within the RSSC (Landry et al. 2020). However, large unexplained diversity among the T3SS repertoire remains among RSSC members.

Reactive oxygen species (ROS) are also made by plants after infection by pathogens as a defence system, and therefore RSSC must overcome this oxidative environment during plant infection (Flores-Cruz and Allen 2009). RSSC are known to produce lots of mechanisms, such as producing catalases and peroxidases (e.g. the *Bcp* gene), to protect themselves against these ROS (Genin 2010, Genin and Denny 2012). Furthermore, biofilms have been suggested to filter out nutrients from the flow of xylem fluid, helping bacteria to survive in this environment, as well as protect them from host plant immune defences (Álvarez et al. 2010, Genin and Denny 2012, Meng 2013). Another trait that is thought to be linked with bacterial pathogenicity is siderophore production (Bhatt and Denny 2004). Siderophores can be used to increase resource utilisation, however secreted siderophores have also been seen to be supress RSSC populations in the rhizosphere (Gu et al. 2020). Finally, arguably the most important virulence trait of RSSC is its production of exopolysaccharide (EPS). EPS is produced in large quantities and accumulation causes vascular degradation and the classic wilt symptoms seen in susceptible hosts (Genin and Denny 2012).

1.8.2 Environmental selection factors affecting RSSC infections

In order to cause infections RSSC species must first invade bacterial communities that already exist in the rhizosphere and reach a threshold density to activate virulence genes (Genin and Boucher 2004, Mansfield et al. 2012, Wei et al. 2015). Research has shown that diverse communities with clear niche overlaps with RSSC reduce invasion success and can lead to lower levels of bacterial wilt disease (Wei et al. 2015, Li et al. 2019). Negative interactions towards the invading bacteria

can be due to direct inhibition, such as antibiotic production, or resource competition (Wei et al. 2015, Li et al. 2019). Competition for resources, such as carbon, reduces pathogen establishment and growth within the rhizosphere, therefore reducing infection of host plants (Wei et al. 2015). In a diverse rhizosphere where all food resources are utilised, RSSC establishment is reduced massively due to increased competition (Wei et al. 2015). Production of siderophores is one method bacteria can selfishly improve its resource utilisation. These are iron-chelating compounds that can be secreted by microorganisms to help bacteria to collect essential iron from the environment for their survival and growth (Hider and Kong 2010). Production of polycarboxylate siderophores is common in *Ralstonia* species and this trait may also be linked with the strains survival within the soil (Bhatt and Denny 2004). Siderophore production can also affect RSSC's ability to invade microbial communities within the rhizosphere and determine bacterial wilt disease incidence (Gu et al. 2020). *Ralstonia solanacearum* species complex is versatile when it comes to resource utilisation and is known for its extreme strain variation of metabolite utilisation, which was exploited to separate the diverse species into separate biovars before sequencing was available (Hayward 1991). Resource utilisation can indicate the life-history and ecology of RSSC strains, for example, utilisation of different metabolites differs within the same strain depending on its past environmental history as well as its ability to infect plants (Zuluaga et al. 2013, Peyraud et al. 2016). This is because resource utilisation is also an important factor in a pathogen's ability to exploit host resources and can determine how virulent they are (Peyraud et al. 2016). Sugar and amino acid content of plants differ before and after RSSC infection (Zuluaga et al. 2013) and the bacterium can even manipulate nutrients within the host plant, producing effector proteins to synthesise GABA (γ-aminobutyric acid) from other molecules (Xian et al. 2019). This suggests that to infect plants RSSC has a preference on what resources the pathogen use within the environment as well as within the host, even altering their environment to achieve successful infection.

Between host plants, RSSC must survive transmission through the environment, such as in soil or river water, where the pathogen can encounter many different environmental stressors. Adapting to these environmental stressors,

both abiotic and biotic, is hence extremely important for bacteria to survive and persist in the environment during transmission to a new host. Not only do fluctuations in stress occur naturally but global change means that changes in temperature, pH, light levels, carbon dioxide and oxygen concentrations, nutrient availability, salinity, and other environmental variables can occur together creating a complex picture of abiotic stressors in the future (Brennan and Collins, 2015). Adapting in the presence of multiple different stressors in the environment can have different effects on evolution of the microorganism (Brennan and Collins, 2015) and it has been suggested that adapting to these stressful conditions can affect pathogen virulence and plant-pathogen interactions (Zarattini *et al.*, 2021).

Within these environmental reservoirs RSSC would have to cope with the stress of oligotrophic habitats with little nutrients. Survival studies in sterile water have reported persistence of the bacterium for variable periods, and occasional capacity to wilt susceptible hosts (Álvarez et al. 2008), suggesting that the ability to grow in these nutrient limited conditions is extremely beneficial for RSSC. Also, high soil moisture, accumulating from either a high water table or heavy rainfall, usually favour bacterial wilt disease incidence (Hayward 1991). Studies have shown that survival of the pathogen is greatest in wet but well-drained soils, whereas survival is affected adversely by soil desiccation and by flooding (Hayward 1991). Furthermore, bacterial wilt incidence can be predicted based on soil moisture properties (Jiang et al. 2021). Survival in waterways is also a key part in RSSC's lifecycle and can lead to infectious outbreaks in crops (Parkinson et al. 2013). Therefore, water potential stress will be a key stress factor in *R. solanacearum's* lifecycle. When water potential decreases, the metabolic activity of most microbial species is decreased, resulting in lowered respiration and nutrient mineralisation (Manzoni et al. 2012), tolerance to this stress may help pathogen survival and increase bacterial wilt disease incidence.

RSSC can encounter many different abiotic stressors such as changes in the pH, temperature, and salinity. Changes in pH can occur within the soil, river, as well as within the plant xylem (Bahrun *et al.*, 2002; Secchi and Zwieniecki, 2016; Li *et al.*, 2017). The acidity of the soil has a strong influence on the composition of bacterial communities (Rousk *et al.*, 2010) reducing antagonistic microorganism densities,

such as *P. fluorescens* and *B. cereus* (Li *et al.*, 2017; Wang, Liu and Ding, 2020). Research has shown that bacterial wilt disease incidence is a lot higher in these acidic soils (pH 4.5 - 5.5), favouring growth of *Ralstonia solanacearum,* indicating that *R. solanacearum* is well-adapted to an acidic soil environment (Li *et al.*, 2017). RSSC also finds high salinity unfavourable, growing in 1% NaCl media, but struggling to grow in media with 2% salt content (Álvarez et al. 2010). Ions released from weathering minerals into the soil and rivers can cause high natural salinity levels in these environments where *R. solanacearum* resides (Parkinson *et al.*, 2013; Shrivastava and Kumar, 2015). Salt may also be applied to rhizospheres through irrigation water or as fertilisers and when precipitation is insufficient and can accumulate in the soil resulting in high soil salinity (Shrivastava and Kumar, 2015).

Cold tolerant *R. solanacearum* strains (race 3 biovar 2 or phylotype IIB sequevar 1) have been proven to infect plants successfully at lower temperatures compared to other RSSC strains (Milling et al. 2009, Meng et al. 2015). However, low temperatures correlate with lower survival rate within the soil and water (Álvarez et al. 2010), suggesting that lower temperatures are stressful for RSSC. A study on *Ralstonia solanacearum* strains (race 3 biovar 2 or phylotype IIB sequevar 1) from Egyptian canals found that temperature, along with competition with other microorganisms, had a marked effect on survival of this cold tolerant *R. solanacearum* strain (Tomlinson *et al.*, 2009). The ability of RSSC to grow and persist in these extreme temperature environments is crucial for pathogen success and their ability to tolerate them can indicate what environments these strains have encountered in its history.

There are not only abiotic factors in the environment that can cause stress towards RSSC species. RSSC is constantly surrounded by a community of other microorganisms, whether it's in the water or rhizosphere, which impose many different selection pressures on the pathogen depending on the environment and what organisms are present (Fukui 2003). For example, some bacteria in the rhizosphere community, such as *Bacillus* and *Pseudomonas* species, are known to produce a variety of antibiotics that kill RSSC species to reduce competition within the environment (Allen et al. 2010, Yuliar et al. 2015, Wei et al. 2019). Therefore, antibiotic resistance may be a trait the RSSC has evolved to be able to persist in this

community. Many enemies which can kill the bacterial wilt pathogen can be found in these communities, such as a large variety of bacteriophage (viruses that infect bacteria), and predatory protists (Jousset et al. 2006, Yamada et al. 2007, Bhunchoth et al. 2015, Xiong et al. 2019). Being resistant and able to survive in these communities would be a vital trait for RSSC to persist, reproduce and eventually infect plants within their environment.

While many attempts have been made to better understand the genetic diversity among this *Ralstonia solanacearum* species complex (van der Wolf et al. 1997, Castillo and Greenberg 2007, Cellier and Prior 2010, Stevens and Van Elsas 2010, Cho et al. 2018, Sharma et al. 2022), there is still a lack of understanding of the phenotypic diversity among ecologically relevant phenotypes within this pathogen. This information is needed to better understand the life history of this pathogen as well as discover the underlying mechanisms of certain genes, by linking the two (genotypes and phenotypes) together.

1.9 Linking phenotypes with genetic variants using a genome wide association study (GWAS)

Genome wide association studies (GWAS) are used to identify parts of the genome that are associated with a trait, identifying the causative gene, genes or gene regions that create a certain phenotype. A large sample of individuals of a certain species are first collected and a GWAS is performed using these four steps: (1) collecting the genotypic variation among these individuals via microarrays or sequencing. Genotypic variation can be measured using SNP (single nucleotide polymorphisms), the presence or absence of genes, indels (insertions or deletions) or K-mers (small segments of DNA usually 30 bp long) (Read and Massey 2014). (2) Grouping individuals based on the phenotypic trait of interest, which can be continuous or discrete. (3) Testing the association of a genotype with a certain trait (phenotype). For example, if one genetic variant is more frequently associated with a certain trait then this variant is said to be associated with this trait. (4) Finally, the last step is calculating the significance of this association, where correction for multiple testing is also conducted (Read and Massey 2014, Power et al. 2016).

Human GWAS commonly uses SNP microchip arrays to collect SNP variants across the genome. However, creating SNP typing panels are generally associated with high fixed costs and therefore few platforms were designed for bacterial species as each species will need a specific SNP typing panel. This meant that GWAS on bacteria was slow to emerge (Read and Massey 2014). Now with whole genome sequencing costs dramatically decreasing, a recent increase of GWAS has been employed to identify causative genes in bacteria has occurred (Read and Massey, 2014). The first bacteria GWAS using sequenced data was performed in 2013, where vitamin B5 biosynthesis was found to be associated with host specificity in *Campylobacter* (Sheppard et al. 2013). Whole genome sequencing also makes it possible to perform GWAS using other variants other than SNPs such as DNA fragments known as k-mers or unitigs. Using SNPs can cause an ascertainment bias where the strains that are more genetically similar to the reference tend to have more accurate SNP calls. However, using k-mers or unitigs can solve this issue as they are short fragments of DNA that are created with 'reference-free' multiple alignment methods (Read and Massey 2014). Also, with bacteria, validation experiments are relatively easy to conduct, which can be used to prove that the associated gene is causally causing the phenotype of interest. This means bacteria GWAS is less reliant on replication and reduces the chances of false positive results (Read and Massey 2014, Power et al. 2016).

However, compared to humans, bacteria have 3 confounding factors, that need to be taken into consideration when conducting a GWAS: (1) they often face stronger selection pressures, (2) they have a higher linkage disequilibrium within the genome and (3) larger population stratification due to lack of recombination (Read and Massey 2014, Chen and Shapiro 2015). Firstly, strong selection pressures in bacterial communities often occurs due to strong host association, antibiotic resistance etc. This causes a higher linkage between genes via population sweeps, or bottlenecks, which removes much of the genetic diversity within a population. The second confounding factor is the extent that any two alleles within a population are on the same ancestral 'haplotype block' of DNA, termed as their linkage disequilibrium (LD). This usually decreases with genetic distance on the chromosome because of sexual recombination, however bacteria are asexual, with

only one or two chromosomes, and therefore lack this ability to reduce LD as quickly (Read and Massey 2014). The separation of haplotype blocks is important for distinguishing causal loci from passively linked mutations in a GWAS, therefore an important first step before performing a bacterial GWAS is to characterise LD (Read and Massey 2014, Chen and Shapiro 2015) and causative genes should ideally be acquired multiple times by different lineages, as this improves the evidence for the gene causing the trait (Read and Massey 2014). Finally, population structure confounds GWAS as the non-random distribution of alleles within subpopulations (Read and Massey 2014) can cause significant associations between genotypes and phenotypes to be detected which are due to bacteria being related rather than the genome region causing the phenotype (Chen and Shapiro 2015). Problems relating to structured genetic variation would be expected to be worse in bacterial strains as they are both haploid and clonal. The problem of population stratification is particularly acute in highly clonal (rarely recombining) bacteria, and in those with separate geographic or host-associated subpopulations (Chen and Shapiro 2015). New refinements on GWAS approaches for microbes are constantly being developed, each using different methods to control for these confounding factors. Some software use clustering methods to control for population structure while others use phylogenetic trees (Collins and Didelot 2018) to take into account genetic relatedness. However, this correction usually compromises the amount of power the software has, which could potentially be overcome by using larger sample sizes (Chen and Shapiro 2015).

1.10 Project aims and thesis chapter outline

The overall aim of this project was to explore and improve our understanding on the phenotypic and genetic diversity of the *Ralstonia solanacearum* species complex (RSSC). This project was part of a larger one in collaboration with another PhD student, Martina Stoycheva, who has conducted genomic analysis on RSSC strains to discover the genetic diversity among this bacterial plant pathogen group. Therefore, this thesis primarily focusses on the phenotypic diversity among the RSSC bacterial plant pathogen and attempts to link

this trait diversity with their genomic information. Within this research I aimed to answer:

1. **How phenotypically diverse is the *Ralstonia solanacearum* species complex (RSSC)?**
   This was investigated by conducting high-throughput phenotyping, collecting 46 ecologically relevant traits, across two collections of RSSC bacterial isolates together comprising 379 isolates. One collection consisted of 194 isolates spanning the global distribution of the RSSC (chapter 2), while the other consisted of 186 isolates from a single population of *R. solanacearum* within the UK (chapter 3).

2. **What is driving trait diversity among the *Ralstonia solanacearum* species complex (RSSC)?**
   Life-history data, such as the location, time, and host each isolate was sampled from (chapter 2 and 3), along with trait correlations (chapter 3 and 4), were both explored as potential drivers of phenotypic trait diversity.

3. **What genetic mechanisms are underpinning trait variation within the *Ralstonia solanacearum* species complex (RSSC)?**
   A combination of comparative genomics (chapter 2 and 3), genome wide association studies (chapter 2 and 3), and experimental evolution combined with whole genome sequencing (chapter 4) were used to determine genetic mechanisms behind phenotypic trait differences within the RSSC.

To answer these three research questions this thesis includes the following chapters presented in the style of research papers:

Chapter 2: Two species within the *Ralstonia solanacearum* species complex (RSSC) have diverse but similar ecologies

In this first data chapter I explore the phenotypic diversity among the RSSC, using a global collection (n=194) spanning eight decades and two species within the RSSC (*R. pseudosolanacearum* and *R. solanacearum*). Specific comparisons between the two most widely distributed species within the RSSC, *Ralstonia solanacearum* and *R. pseudosolanacearum,* were conducted to better understand the evolutionary history of these two species. Life-history characteristics, sampling location and host, was also linked to phenotypic diversity to determine if they can explain trait differences among the RSSC. Finally, genomic comparisons and a genome wide association study (GWAS) was conducted to link phenotypic variation with genetic mechanisms.

This chapter reveals that a large overlap in trait diversity exists between the two RSSC species. However, small differences were identified, with *R. pseudosolanacearum* being metabolically more efficient and more diverse, while *R. solanacearum* was more oligotrophic. Across all RSSC strains, five distinct phenotypic groups, or 'ecotypes' were also identified each differing in their trait specificity, either being oligotrophic, heat tolerant, cold tolerant, antibiotic resistant/biofilm producer, and metabolically efficient. Members of both species were found in all five ecotypes, highlighting that these two species do belong to a species complex and suggests that ecotype separation drives this large diversity among RSSC strains. To test if life-history characteristics, host and location sampled from, drives diversification each strain's metadata was linked to their phenotypic diversity. This revealed that host and continent can explain small amounts of trait variation among the world collection of RSSC strains. Linking this trait variation with genomic information revealed that all ecotypes were present across the RSSC phylogeny with no clear ecotype clustering. However, large differences in accessory genome variation were observed, potentially suggesting that horizontal gene transfer drives RSSC ecotype differentiation. Finally, a genome wide association study (GWAS) revealed a type II secretion system associated with cold tolerance and three novel genome regions associated with rifampicin resistance. This bridges the gap between genomic information and physical characterisation, which is especially important during this current omics era where vast quantities of genomic data are available with little known links with functional traits. Overall, this chapter

outlines the phenotypic differences between two species within the RSSC, linking some of this trait diversity with life-history data and genetic mechanisms.

Chapter 3: The UK *Ralstonia solanacearum* bacterial plant pathogen population has diversified into three 'ecotypes'

In this second data chapter, I explored the phenotypic diversity among a single population of *Ralstonia solanacearum* within the UK (n=186), from both environmental and plant host samples. The UK *R. solanacearum* belongs to the clonal cold tolerant strain (phylotype IIB sequevar 1) and has thought to have been introduced to the country relatively recently (around 30 years ago). This collection contains isolates spanning from the first recorded outbreak in 1992 to 2019, and therefore this chapter aimed to determine the phenotypic diversity of a recently introduced RSSC population to the UK environment, providing insights on how initial adaptation occurs. Again, like chapter 2, life-history characteristics, the time, location and host isolated from, were linked to phenotypic diversity to determine if they help explain trait diversity among this population. Furthermore, trait correlations can aid (positive trait correlations) or limit adaptation (negative trait correlations) due to genetic linkages. Therefore, the impact of trait correlations on UK *R. solanacearum's* phenotypic diversity was also explored. Finally, to study the genetic mechanisms behind trait diversity of UK *R. solanacearum* comparative genomics and genome wide association studies (GWAS) were also conducted.

This chapter reveals that the UK population of *R. solanacearum* separates into three distinct 'ecotype' groups that differ in their growth in nutrient limited conditions, and overall trait generalism and specialism. This presence of separate ecotypes in both the UK *R. solanacearum* population and the global RSSC collection (chapter 2) suggests that niche specialisation is important in driving high diversity of this plant pathogen both at the species and population level. Exploring trait correlations revealed that they differed between each ecotype, while linking metadata with trait variation also revealed that location and isolation source cannot explain trait variation. However, the isolation date correlated with increased diversification between UK *R. solanacearum* isolates over time. This highlights the recent introduction of this bacterial species to the UK as diversification occurs when

new niche opportunities open. Temporal trait variation was also driven by an increase in antibiotic resistance and biofilm production, revealing that these traits are important in driving *R. solanacearum* adaptation within the UK. Linking this phenotypic variation with genetic data revealed extremely high genetic similarity between isolates with 98% shared genes across all three ecotypes. This suggests a role of other genetic mechanisms, such as epigenetics, in driving RSSC adaptation and diversification in the UK. Overall, this chapter reveals that a recently introduced *R. solanacearum* population shows relatively little genetic variation, while also harbouring large phenotypic diversity.

Chapter 4: Adaptation to environmental stresses can explain diversification of phytopathogen *Ralstonia solanacearum*

Finally, in this third data chapter I concentrate on how abiotic stresses, such as pH and salinity, within the environment can shape the UK *Ralstonia solanacearum* phenotypic diversity using experimental evolution approach. Specifically, this chapter explored the potential role of trait correlations in shaping *R. solanacearum* adaptation by evolving one UK isolate to single stress conditions (acidic, alkaline, and saline conditions) as well as a combination of stresses (acidic with salinity and alkalinity with salinity). I predicted that evolutionary outcomes would depend on how these stress tolerance traits are correlated with one another, with positive correlations resulting in generalist organisms (equally adapted to each stress conditions), while negative trait correlations constraining or driving adaptation into specialist genotypes. After evolving this *R. solanacearum* isolate to different stress conditions these evolved clones were sequenced (whole genome sequencing) to reveal genetic mechanisms underpinning stress tolerance traits.

My results revealed that specialist adaptation to extreme pH occurred, while no adaptation to the combined stress conditions took place, suggesting negative trait correlations between the two stress conditions. Adapting to all stress conditions also resulted in a trade-off in metabolic capacity, suggesting that trait-correlations across metabolism and environmental stresses exists within *R. solanacearum*. Genome sequencing revealed little core genome variation (SNPs and small INDELs) between evolved clones, the only promising causative one being a

few SNPs within a type IV pilus gene, agreeing with the lack of genetic variation seen within the UK population in chapter 3. However, insertion sequence (IS) movement was abundant, suggesting that genome rearrangement of these insertion sequences plays a potentially important role in the initial *R. solanacearum* adaptation. Furthermore, when phenotypic diversity of clones evolved to different stress conditions were compared this revealed a larger diversification of clones adapted to stress conditions compared to the stress-free adapted clones, suggesting that environmental stresses can drive increased phenotypic diversity. Overall, this chapter highlights that trait correlations can drive diversification of *R. solanacearum* and IS movement plays a large role in RSSC adaptation to environmental stresses.

Chapter 5: General discussion

A general discussion considering the results and context of all chapters together and an overview of my three research questions. Potential future directions to further this area of research are also discussed here.

The methods used for each chapter are outlined in their respective chapters. References are provided at the end of the thesis. A list of all the bacterial isolates within each collection, along with supplemental protocols, are found within appendix A at the end of the thesis. Supplementary information for each chapter is also shown at the end of the thesis in separate chapter appendices (B-D).

2 Chapter 2: Two species within the *Ralstonia solanacearum* species complex (RSSC) have diverse but similar ecologies

2.1 Abstract

The *Ralstonia solanacearum* species complex (RSSC) is a diverse group of plant pathogenic bacteria, causing bacterial wilt disease among a wide range of plants of agricultural, horticultural, and environmental importance all over the world. While previous research has been conducted to determine the genetic diversity and host range of RSSC, the ecological relevance of its wider phenotypic diversity remains unclear. This research aimed to compare a broad range of phenotypic traits within two widely distributed RSSC species, *R. pseudosolanacearum* and *R. solanacearum*, and to investigate their potential ecological similarities and differences. To achieve this, high-throughput phenotyping was conducted, on a worldwide collection of 194 strains, with traits involved in virulence, stress tolerance, and metabolism. Only a small amount of trait variation was explained by species identity, with *R. solanacearum* being more oligotrophic, and *R. pseudosolanacearum* more diverse and metabolically efficient. However, strains could be divided into five distinct 'ecotypes' based on variation among 46 ecologically relevant phenotypic traits. All five ecotypes comprised strains from both species, in line with both species belonging to the same species complex and suggesting that ecotype separation may drive the observed diversity among RSSC strains. Linking this trait variation with genomic information revealed that all ecotypes were present across the RSSC phylogeny with no clear clustering. However, differences in accessory genome variation were present, potentially explaining and driving these ecotype differences. A GWAS also highlighted a type II secretion system associated with cold tolerance and three novel genome regions associated with rifampicin resistance. Overall, this research suggests that the environment is selecting for similar ecotype differences within both RSSC species, aligning with the fact that both species still share similar life cycles.

2.2 Introduction

The *Ralstonia solanacearum* species complex (RSSC) is a diverse group of plant pathogenic bacteria and the causative agent of bacterial wilt disease in a wide range of plants across the world (Hayward 1991, Safni et al. 2018, Bragard et al. 2019, EPPO 2022). A species complex has been defined as a group of closely related species that are highly similar, so much so that the boundaries between them can be unclear (Fegan and Prior 2005). While DNA-DNA homology has revealed that the relatedness between RSSC members is often less than 70%, commonly considered as the threshold between one microbial species and another, considerable similarities in phenotypes and disease symptoms have also been observed (Fegan and Prior 2005, Lowe-Power et al. 2018). Gillings and Fahy first used to term "species complex" to reflect the phenotypic and genetic variation observed within the bacterial wilt bacteria (Fegan and Prior 2005). A lack of consistency in pathogen host range and behaviour between these closely-related species in the RSSC, in addition to phenotypic and behavioural similarities between genetically distinct members, further justified the need to define this pathogen as a species complex (Sharma et al. 2022). Based on significant variations in the whole genome, the RSSC currently comprises three separate species, *R. pseudosolanacearum*, *R. solanacearum* and *R. syzygii* (Safni et al. 2014, 2018, Prior et al. 2016, Kumar et al. 2018, Stevens et al. 2018) which aligns with the previous phylotype classification system (Fegan and Prior 2005). *Ralstonia solanacearum* was previously classified as RSSC phylotype II and predominantly originated from the Americas. *Ralstonia pseudosolanacearum* comprises both phylotype I and III classified strains and predominantly originated from Asia and Africa (for phylotype I and III respectively), while *R. syzygii,* comprises previously classified phylotype IV strains, originating from Indonesia (Safni et al. 2014, 2018).

R. syzygii strains are mainly located within Indonesia (Safni et al. 2018), with recent emergence in other countries (EPPO 2022). However, the other two species within the RSSC, *R. solanacearum* and *R. pseudosolanacearum*, have global distributions and can be found within the same country in some regions of the world (EPPO 2022), suggesting that these two species have overlapping niches. The separation of these two widely distributed species is thought to date back to the

geological separation of the continents (Hayward 1991), with the two species reflecting each lineage that arose when ancestors became geographically separated (Genin and Denny 2012). There are also some apparent differences in host preference between the two species. *Ralstonia pseudosolanacearum* has been recovered from a wide range of hosts, the most common belonging to the *Solanum* genus, including potato (*Solanum tuberosum*) and tomato (*Solanum Lycopersicon*) (Bragard et al. 2019). Uniquely *R. pseudosolanacearum* also comprises of strains that can infect ginger (*Zingiber officinale*) and mulberry (*Morus* spp.) (EPPO 2018). *Ralstonia solanacearum* also has a wide host range, with the primary host plants also belonging to the *Solanum* genus (Bragard et al. 2019). The *R. solanacearum* species includes the Moko disease-causing strains infecting banana (*Musa* spp.) and *Heliconia* species (EPPO 2018). Furthermore, *R. solanacearum* also includes the cold-tolerant potato (*Solanum tuberosum*) brown rot causing strain, previously known as race 3 biovar 2 or sequevar 1 of phylotype IIB (Fegan and Prior 2005, Safni et al. 2014). The phylotype IIB sequevar 1 strain is adapted to lower temperatures compared to other more tropical *R. solanacearum* strains, and therefore also poses a serious threat to temperate agriculture (Williamson et al. 2002, EPPO 2018).

Despite these differences, *R. pseudosolanacearum* and *R. solanacearum* species share very similar life cycles. Both can infect a broad range of host plants, producing various virulence factors and evading a wide range of plant immune responses. Additionally, they must also persist within the environment, such as in the soil and river water, during transmission from host to host (Genin 2010). As a result, both RSSC species have evolved to be highly diverse both genetically and phenotypically with multiple complex regulatory networks that respond to internal and environmental cues allowing for high phenotypic plasticity (Schell 2000, Cellier and Prior 2010, Genin and Denny 2012, Perrier et al. 2019, Chen et al. 2022, Yan et al. 2022). This means that they can undergo phenotypic conversions depending on environmental cues and therefore, the overall diversity of this species cannot be determined by its genome alone (Drenkard and Ausubel 2002). Furthermore, RSSC have large accessory genome variation, accredited to horizontal gene transfer within the bacterial community (Ailloud et al. 2011, Geng et al. 2022). Despite current research investigating genetic differences among RSSC members, it is still

unclear how this genetic variation is associated with species-level phenotypic differences and similarities.

This research aimed to explore diversity within the two widely distributed RSSC species, *R. pseudosolanacearum* and *R. solanacearum,* using high-throughput phenotyping to compare a collection of virulence traits, as well as other ecologically important traits involved in stress tolerance and metabolism. As a plant pathogenic bacterium, virulence traits, such as protein production, tolerance to reactive oxygen species, biofilm formation and siderophore production, are important phenotypes driving RSSC adaptation. Furthermore, resource utilisation is important for invading plant rhizosphere microbiomes and establishing infection of hosts, as well as for survival within environmental reservoirs (Genin and Boucher 2002, Mansfield et al. 2012). Selection within environmental reservoirs can also drive RSSC adaptation to abiotic stresses, resulting in changes in tolerance to pH, salinity, temperature, and water potential (Elphinstone et al. 1997, Sgrò and Hoffmann 2004, Parkinson et al. 2013, Stevens et al. 2018). RSSC is also constantly surrounded by a community of other microorganisms (Fukui 2003), some of which, such as *Bacillus* and *Pseudomonas* bacterial species, are known to produce a variety of antibiotics that can kill RSSC strains to reduce competition within the environment (Allen et al. 2010, Yuliar et al. 2015, Wei et al. 2019). Therefore, tolerance of biotic stresses, including antibiotics, could be selected for as they allow RSSC members to persist within these communities.

Therefore, trait diversity was explored across 194 RSSC isolates from both species (*R. pseudosolanacearum* and *R. solanacearum*) (see appendix table A.1 for full list of strains). This was conducted by collecting 46 ecologically relevant traits, regarding three different adaptive strategies: 1) resource utilisation, 2) environmental stress tolerance, both abiotic and biotic, and 3) virulence (see appendix table B.1 for the full list of phenotypic traits collected). This study aimed to improve the understanding of phenotypic diversity among RSSC strains, that are genetically and phylogenetically well characterized (Cellier and Prior 2010), and help distinguish what is driving the observed phenotypic diversity among this complex of plant pathogens. Drivers of trait diversity could be either; 1) genetic differences between the two species, 2) life-history differences between strains, such as host or

location sampled from, or 3) differences between ecological niches to which each strain is exposed. This study also aimed to link this observed phenotypic variation to genetic data, identifying possible underlying genetic causes of phenotypic variation among the RSSC. To do this, a genome wide association study (GWAS) was conducted. This is a unbiased method of exploring associations between genetic information and phenotypes and identifying causative gene or genes for certain phenotypes (Read and Massey 2014, Power et al. 2016). By identifying potential genetic mechanisms involved in trait diversity among the RSSC, novel genes important for plant pathogen survival could then be explored.

Overall, the two RSSC species investigated within this study were found to be phenotypically diverse with large overlaps with their trait diversity. Some differences were observed, with *R. solanacearum* being more oligotrophic, while *R. pseudosolanacearum* was more diverse and metabolically efficient. The continent and host in which isolates were sampled only explained a small amount of trait diversity among RSSC isolates. However, five phenotypically distinct 'ecotypes' were discovered within the RSSC collection (n=194) each differing in their niche preferences and all containing isolates from both species. Linking genotypes to phenotypic variation also revealed that all ecotypes were present across the RSSC phylogeny with no ecotype clustering. However, large differences in each ecotypes' accessory genome was observed, suggesting that horizontal gene transfer within the environment is potentially driving these ecotype differences. Together, these findings supports that environmental conditions could select for similar ecotype differences within both RSSC species. A GWAS also identified novel genetic mechanisms, a type II secretion system associated with cold tolerance and gene regions associated with rifampicin resistance, which requires further investigation. Overall, comparative analysis among a plant pathogenic group has revealed insights into bacterial ecological diversity, providing important insights to RSSC adaptation and identifying potential genetic mechanisms of certain traits.

2.3 Materials and Methods

2.3.1 *Ralstonia solanacearum* species complex (RSSC) isolates and culture conditions

The 194 *Ralstonia solanacearum* species complex (RSSC) isolates used in this study (appendix table A.1) were collected and stored by FERA Science ltd, Sand Hutton, York, UK. These strains were curated from a variety of collections, such as the National Collection of Plant Pathogenic Bacteria (NCPPB) and Protect collection maintained by FERA Science ltd, and spanned a variety of hosts, countries, and years, from 1945 to 2016 (figure 2.1). All RSSC isolates were verified as belonging to the RSSC using 16S rRNA real-time PCR (RT-PCR) with (Weller et al. 2000) primers and the standard protocol advised by the European and Mediterranean Plant Protection Organization (EPPO) (EPPO 2018). All isolates were grown up on SP agar plates (20g/L sucrose, 5g/L peptone, 0.5g/L potassium hydrogen phosphate, 0.25g/L magnesium sulphate heptahydrate, 12 g/L agar) (Mehan and McDonald 1995), before being normalised to 0.1 OD_{600} (optical density at 600nm) using sterile deionised water. They were then suspended in 25% glycerol, in random positions across different 96-well microplates, and cryopreserved at -80°C, ready for high-throughput phenotyping. Among these inoculation microplates were also two well characterised RSSC reference strains, the *R. pseudosolanacearum* type strain GMI1000, and *R. solanacearum* type strain K60, from French Guyana and USA respectively. These strains were repeated three/four times per plate and were placed strategically across the microplates to detect and account for both batch and plate position effects. A collection of 46 different traits were investigated for this RSSC bacterial collection (see appendix table B.1 for full list of traits), categorised into four different groups: resource utilisation, abiotic and biotic stress tolerance, and virulence. A linear model using batch, plate position and technical replicate as random effects was conducted on the two type strains to determine that only small effects were due to these variables (7%, 6% and 0% for batch, plate position and technical replicate respectively), and therefore corrections for batch effect were not conducted.

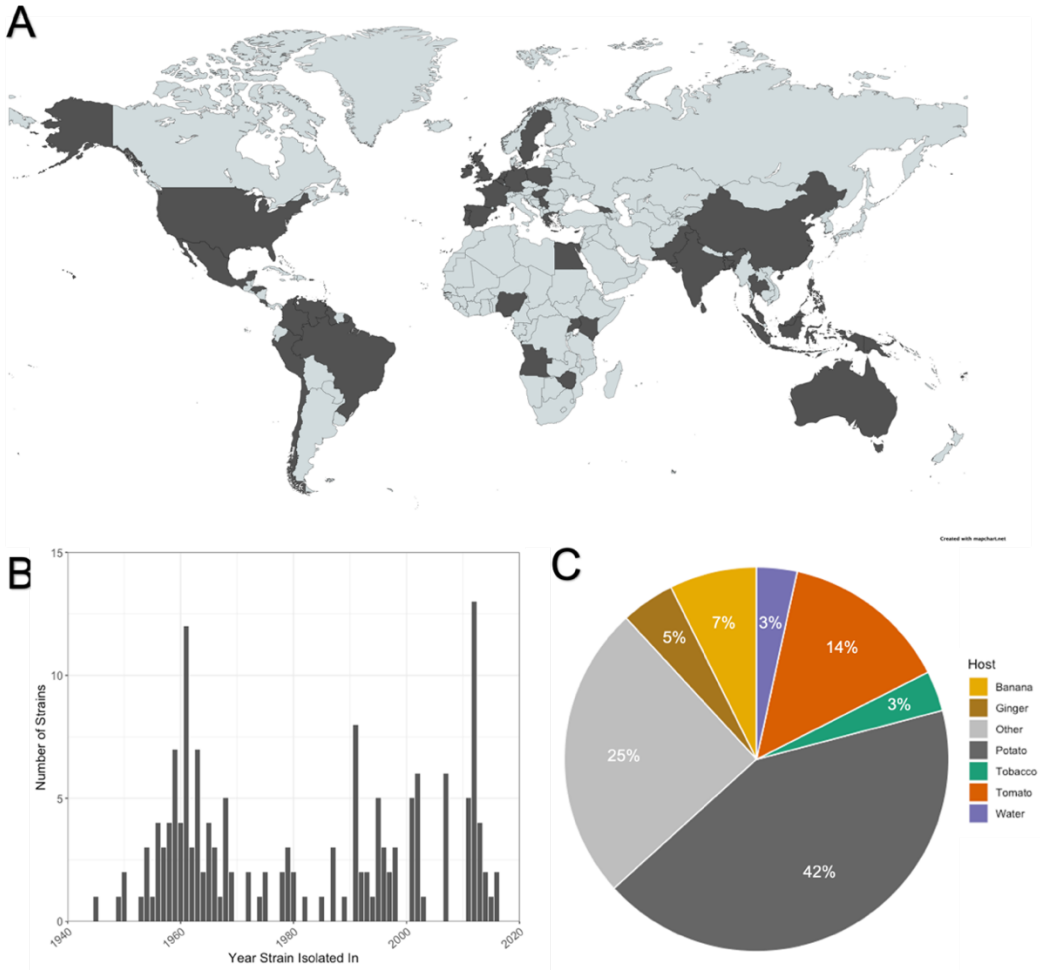


Figure 2.1: *Ralstonia solanacearum* species complex (RSSC) isolate collection and sampling distribution. The RSSC collection used within this study comprised of 194 isolates, however missing metadata for certain isolates was observed. (A) World map showing the geographical distribution of this RSSC collection. Shaded countries indicate that at least one isolate within the RSSC collection is from this region (n=185 in total; map created using mapchart.net). (B) Temporal distribution of this RSSC collection, showing the number of isolates collected per year (n=159 in total). (C) Pie chart showing the host distribution of the RSSC collection, with each section showing the percentage of isolates from this RSSC collection isolated from different hosts. Most hosts have fewer than 3 isolates and are therefore within the `other` category for easier interpretation (n=177 in total). For full list of isolates and metadata see appendix table A.1.

2.3.2 Quantifying phenotypic traits linked with metabolic capacity

The investigation of metabolic traits included growth within four rich standard media commonly used for RSSC culture in the laboratory, as well as 16 single carbon resources (asparagine, glutamine, histidine, proline, serine, glycine, glucose, arabinose, xylose, sucrose, maltose, sorbitol, nicotinamide, citric acid, malic acid, and succinic acid; all at 10mM; appendix table B.1). These carbon resources were chosen because they are commonly present in the rhizosphere and

previous research suggests that they may impact RSSC growth and virulence (Yao and Allen 2006, Stevens and Van Elsas 2010, Jacobs et al. 2012, Zuluaga et al. 2013, Zhang et al. 2014, Wu et al. 2015, 2017, Perrier et al. 2016, Lowe-Power et al. 2018).

2.3.2.1 Complex media

For growth within complex media, casamino acid, peptone and glucose media (CPG: 1g/L casamino acid, 10g/L peptone, and 5g/L glucose) (Kelman 1954), casamino acid sucrose media (CS: 7.5g/L casamino acid, 2g/L sucrose, 0.25 g/L magnesium sulphate heptahydrate, 0.5g/L potassium hydrogen phosphate, and 0.25g/L of ammonium iron(III) citrate) (Remenant et al. 2011), nutrient broth (NB) media (made as per Thermo Fisher scientific's instructions), and sucrose peptone media (SP: 20g/L sucrose, 5g/L peptone, 0.5g/L potassium hydrogen phosphate, and 0.25g/L magnesium sulphate heptahydrate) (Mehan and McDonald 1995). Each media was made and adjusted to pH 7 before being autoclaved at 120°C for 20 minutes. The media (190μl) was then placed into each well of a 96-microplate and 10μl of diluted inoculant bacteria (diluted 100-fold to 0.001OD_{600}) using sterile distilled water, was added to create a final volume of 200μl.

2.3.2.2 Single carbon resources

Growth in the 16 different carbon resources was quantified using a defined minimal media containing salts essential for bacterial growth (OS media). This minimal media was prepared, as in appendix table A.3. Briefly, 5ml of the carbon resource of choice, at concentration 100mM, was added to 45ml of minimal media to make a 10mM solution of each carbon source. For the negative control, or 'no carbon' condition, 5ml of sterile distilled water was added. The pH of the media was then adjusted to pH7 and filter sterilised using a 0.2μm filter. 190μl of this 10mM solution was then added to each well of a 96-well microplate and 10μl of each isolate in the inoculation plates (diluted 100-fold to 0.001 OD_{600}) was then added to these wells, creating an overall volume of 200μl.

Three repeats of each resource, complex and single carbon media, was used. Initial OD_{600} measurements were taken, and microplates were placed in an

incubator at 28°C with humidity to avoid evaporation. OD_{600} reads were then taken every 24 hours for 5 days to measure bacterial growth. Area under the curve (AUC) was calculated as a proxy for growth. Mean average AUC across all replicates (n=3) was taken as the trait value for each isolate in complex media. While relative growth per isolate was calculated, by dividing their AUC in each carbon resource by their AUC in the no carbon control, for single carbon resource traits. Mean average relative growth across all replicates (n=3) was then taken as the trait value for that isolate.

2.3.3 Phenotyping abiotic stress tolerance

Growth within different abiotic stress conditions, including saline, extreme pH, high water potential and nutrient limited conditions, was conducted at 28°C with humidity. Relative growth within a variety of temperatures, 10°C and 35°C, with humidity were also quantified as an abiotic stress tolerance trait (appendix table B.1). SP media (Mehan and McDonald 1995) and 96-well microplates were used for all trait conditions using the same methods as described earlier.

2.3.3.1 Growth in nutrient limited media

For growth within resource limited conditions, 1% and 10% solutions of SP media were made by diluting 100% SP media 10-fold (for 10% SP) and 100-fold (for 1% SP) using sterile deionised water. 190µl of each media, 1% SP, 10% SP, and 100% SP, were added to each well of a 96-well microplate and 10µl of bacteria inoculant (0.001 OD_{600}) was then added.

2.3.3.2 Growth in saline media

For saline stress conditions growth in 3 different concentrations of salt (NaCl); 0.5%, 1%, and 2%, along with a control of 0% NaCl was measured. The isolates (10µl of 0.001 OD_{600}) were grown in 190µl of SP media with different amounts of sodium chloride (NaCl) added to make up the desired salt concentration.

2.3.3.3 Growth in extreme pH

Growth within 4 different pH levels, ranging from acidic to alkali (4.5, 7, 9, 10), was also measured as follows. Drops of sodium hydroxide (NaOH) or hydrochloric acid (HCl) was added to SP media to make the specific pH required before media was autoclaved (120°C for 20 minutes). 10μl of inoculant (0.001 OD_{600}) was then added to 190μl of this media.

2.3.3.4 Growth in water potential stress conditions

To characterize how well *R. solanacearum* grows in water potential stress we took relative growth measurements in 15% polyethylene glycol (PEG)-4000. 15g of polyethylene glycol was added to 100ml of SP media before being filter sterilized with a 0.2μm filter. 10μl of inoculant bacteria (0.001 OD_{600}) was then added to 190μl of this media.

2.3.3.5 Heat and cold tolerance traits

Heat tolerance was quantified as growth in SP media at 35°C, where 10μl of inoculant (0.001 OD_{600}) was added to 190μl of SP media and incubated at 35°C with humidity.

Cold tolerance was quantified as growth in SP media at 10°C, where 10μl of inoculant (0.001 OD_{600}) was added to 190μl of SP media and incubated at 10°C with humidity.

For this trait and all abiotic stress traits above, optical density reads at 600nm (OD_{600}) were taken every 24 hours for 5 days. Area under the curve (AUC) was then calculated as a proxy for growth. Relative growth per isolate was then taken by dividing their AUC within the stress condition by their AUC in SP media without any stresses (100% pH7 SP media, with 0% NaCl added, and incubated at 28°C). Mean average relative growth across all replicates per isolate (n=3) was then taken as the trait value for that isolate.

Another cold tolerance trait was survival at 4°C. 10μl of inoculant (0.001OD) was added to 190μl of SP media (pH adjusted and autoclaved), with three repeats per isolate, and incubated at 4°C for five days. Afterwards, 5μl was spotted onto a

CGP agar plate (1g/L casamino acid, 10g/L peptone, 5g/L glucose, and 17g/L of agar) (Kelman 1954). Survival was measured as a binary growth trait (growth as 1 and no growth as 0) after 48 hours in a 28°C incubator and therefore was only used within the GWAS.

2.3.4 Phenotyping biotic stress tolerance

Antibiotic resistance was calculated as relative growth within four different antibiotics, each with a different mechanism of action, at 'high' and 'low' concentrations per antibiotic. The antibiotics chosen were: ciprofloxacin, a fluoroquinolone which works by targeting the alpha subunits of DNA gyrase preventing it from supercoiling the bacterial DNA which prevents DNA replication (Silver 2016). Gentamycin, an aminoglycoside that works through inhibition of bacterial protein synthesis by binding to 30S ribosomal subunit (Silver 2016). Rifampicin, a rifamycin which acts via inhibiting DNA-dependent RNA polymerase (Silver 2016). Tetracycline, a member of the tetracyclines which works by binding specifically to the 30S ribosome of the bacteria preventing attachment of the aminoacyl tRNA (Silver 2016) (appendix table B.1).

Stocks (0.01g/ml) of each antibiotic were made, and filter sterilised with a 0.2μm filter. Antibiotic stocks were then added to sterile SP media (Mehan and McDonald 1995) in amounts to make up 'high' and 'low' concentrations of each antibiotic, as determined by preliminary experiments on a subsample (10 isolates) of RSSC isolates (appendix figure B.1). Concentrations were as follows; 3 and 5μg/ml for ciprofloxacin, 0.5 and 1 μg/ml for gentamycin, 0.5 and 4μg/ml for rifampicin, and 1 and 5μg/ml for tetracycline. 10μl of inoculant (0.001 OD_{600}) was then added to 190μl of each antibiotic concentration containing media, along with the no antibiotic control (SP media alone). Bacteria were grown in a 28°C incubator, with humidity, for 48 hours and then OD_{600} measurements were taken. Growth (OD_{600}) of each isolate within each antibiotic condition was then divided by their growth (OD_{600}) within the SP control condition to get the relative growth. Mean average relative growth across all replicates per isolate (n=3) was then taken as the trait value for that isolate.

2.3.5 Phenotyping virulence traits

Phenotypes classed as virulence traits included biofilm production, oxidative stress tolerance, siderophore production and protein production (appendix table B.1).

2.3.5.1 Biofilm production

RSSC biofilms have been suggested to filter out nutrients from the flow of xylem fluid, as well as protect bacteria from host plant immune defences, making them an important trait for RSSC host infection (Álvarez et al. 2010, Genin and Denny 2012, Meng 2013). Biofilm assays were conducted, using a crystal violet staining method, where isolates were grown on microplates for 7 days (to make sure carrying capacity had been reached) in the four different complex media, CS (Remenant et al. 2011), CPG (Kelman 1954), SP (Mehan and McDonald 1995) and nutrient broth (NB) (see complex media methods section). The biofilm assay was completed as dictated by (O'Toole and Kolter 1998) with slight modifications, such as those from (Burton et al. 2007). Bacteria were removed from the microplate and microplates were washed three times using sterile deionised water. They were then allowed to dry for 15 minutes before adding 200μl of 0.1% crystal violet. After 15 minutes of incubation at room temperature, the crystal violet was washed out three times using distilled water. 200μl of 70% ethanol was then added to solubilize the crystal violet and OD_{600} was taken after 15 minutes. OD measurements were divided by the OD_{600} of a blank control plate, conducted per batch, and the mean average across the replicates (n=3) was taken as the value of this trait per isolate.

2.3.5.2 Oxidative stress tolerance

Reactive oxygen species (ROS) are produced by plants after infection as a defence system (Flores-Cruz and Allen 2009), and therefore tolerance to ROS is an important virulence trait for RSSC isolates. Growth of all isolates were taken in SP media (Mehan and McDonald 1995) with 100mM of H_2O_2, the sublethal amount of ROS for RSSC (Tondo et al., 2020), within a 28°C incubator with humidity. All media

was freshly prepared by adding the required amount of filter sterilized H_2O_2 (0.2μm filter) to SP media and was kept in the dark using foil as dictated by (Rodríguez-Rojas et al., 2020). Inoculant bacteria (10μl at 0.001 OD_{600}) was added to 190μl media with 100mM of H_2O_2, or no H_2O_2 for the control. OD_{600} was taken every 24 hours for five days and AUC was then calculated as a proxy for growth. Relative growth per isolate was calculated by dividing AUC within H_2O_2 by their AUC in the SP control. Mean average relative growth across all replicates per isolate (n=3) was then taken as the trait value for that RSSC isolate.

2.3.5.3 Siderophore production

Iron chelating ability of bacterial supernatant was measured as a proxy for siderophore production as this has been indicated as an important trait for rhizosphere community invasion and RSSC virulence (Gu et al. 2020). The (Schwyn and Neilands 1987) universal siderophore assay was used with chrome azurol S (CAS) and hexadecyltrimethylammonium bromide (HDTMA) as indicators. This assay works as CAS/ HDTMA complexes tightly bind with ferric iron to produce a blue colour, when supernatant with a strong iron chelator, such as a siderophore, is added this then removes the iron from the dye complex, changing the colour from blue to orange/pink (Arora and Verma 2017).

All bottles were washed with 10mmol/L HCl (6ml of 1M HCl in 600ml of Milli-Q ultrapure water) and milli-Q ultrapure water before stock solutions were made to remove any iron residues. $FeCl_3$ stock solution was made by dissolving 0.135g of $FeCl_3$ $6H_20$ in 500ml of 10mmol/L HCl, the CAS stock solution made by dissolving 0.12g of CAS in 100ml of milli-Q ultrapure water. The HTDMA solution was made by dissolving 0.088g of HTDMA in 200ml of milli-Q ultrapure water and the piperazine buffer by dissolving 17.23g of piperazine in 120ml of milli-Q ultrapure water and adjusting the pH to 5.6 using 37% HCl. All stock solutions were stored at 4°C. On the day the siderophore assay was conducted, the chrome azurol S (CAS) solution was freshly made by mixing 0.75ml of $FeCl_3$ stock solution, 3.75 ml of CAS stock solution, 25 ml of HTDMA solution, 15 ml of the piperazine buffer and 5.5ml of milli-Q ultrapure water.

Isolates (0.001 OD_{600} inoculant) were grown up in 190µl of SP media (Mehan and McDonald 1995) for three days at 28°C with humidity (n=6 per isolate). OD_{600} measurements were taken, and plates centrifuged at 2000 rpm for 15 mins. From these centrifuged plates 100µl of the supernatant per replicate (n=6) was then removed and placed into two 96-well filter plates (with a total volume of 300µl in each plate, totaling to 600µl overall) which were then centrifuged at 2000 rpm for 15 mins. After centrifuging these plates 100µl of filtered supernatant was then placed into a new 96-well microplate with 100ul of CAS solution and mixed carefully, to ensure no bubbles were created, using a pipette. Four technical replicates were conducted per isolate, along with the control (SP media with no inoculant). Plates were left to incubate for 2 hours, and OD reads at 630nm (OD_{630}) were taken. Siderophore production (psu) was then calculated using the below formula:

$$\frac{(A_r - A_s) * 100}{A_r}$$

Where A_r is the absorbance of the reference (CAS solution and uninoculated broth), and A_s the absorbance of the sample (CAS solution and cell-free supernatant of sample) (Arora and Verma 2017). Mean average across the four technical replicates were taken and then divided by average initial OD_{600} to get Siderophore production (psu) per cell values (normalised by growth).

2.3.5.4 Protein production

RSSC strains secrete a lot of virulence proteins while invading plant hosts, to evade plant immune systems, degrade cell walls etc. (Genin and Denny 2012). Therefore, Bradford assays of RSSC isolate's supernatant was conducted to determine the amount of protein produced. RSSC isolates (0.001 OD_{600}) were grown up in 190µl of sterile SP media (Mehan and McDonald 1995) for three days in a 28°C incubator with humidity (n=6 per isolate). OD_{600} measurements were taken, before plates were centrifuged at 2000 rpm for 15 minutes. 100µl of supernatant per replicate was then taken and placed into two 96-well filter plates (each with a total volume of 300µl) which were then centrifuged at 2000 rpm for 15 minutes. The Pierce detergent compatible Bradford assay kit was conducted using the

manufactures 2-25ug/ml detection limit microplate protocol with a few modifications from (Baroukh et al. 2021). 100µl of Bradford dye reagent was added to 96-well EIA/RIA assay microplates, then 100µl of supernatant was added and mixed with a pipette carefully to ensure no bubbles arose. Plates were left to incubate for 30 minutes before OD was taken at 595nm (OD_{595}). Three technical replicates were conducted per isolate. Bovine serum albumin (BSA) was used as the standards with a variety of concentrations, from 0 to 50µg/ml, diluted using sterile SP media.

The amount of protein within the supernatant was calculated by subtracting the blank standard (0µg/ml standard) OD_{595} from all samples and standards OD_{595}. A standard curve was then produced by plotting the average blank-corrected measurements for each BSA standard vs. its concentration in µg/ml. Using this standard curve, protein concentration estimates were calculated for each sample. This protein concentration was then divided by the average growth of that isolate, OD_{600} before supernatant was extracted, to get protein produced per cell. The mean average across all technical repeats per isolate (n=3) was then taken as this trait value.

2.3.6 Whole genome sequencing and genetic variant curation

Whole genome sequencing was conducted on 182 isolates. DNA was first extracted following the Qiagen DNeasy Blood and Tissue Kit, with the optional RNase A step included and slight modifications to the manufacturers protocol (appendix table A.4). The elution buffer used was trisaminomethane hydrochloride (Tris HCl, 10mM, pH 8) as recommended by Earlham Institute for their library preparation. Quality of the extracted DNA was determined using 3 methods. 1) Nanodrop, to detect any impurities present. 2) Quant-iT™ double stranded (ds) DNA broad-range assay, performed as per the manufacturer's microplate assay instructions, to discern double stranded DNA concentration using a fluorescent DNA binding stain. 3) electrophoresis gels using 0.8% agarose to ensure minimal shearing of the DNA occurred. DNA was diluted to 15ng/µl and transported to Earlham for Illumina MiSeq 30x sequencing and raw untrimmed paired FASTQ files were

received. Files can be found publicly available at SRA under project number PRJNA823737.

2.3.7 Genetic analysis: linking phenotypes with genetic properties

A genome wide association study (GWAS) was used to link phenotypic variation to changes within RSSC genomes. Genomic variation across RSSC isolates was determined by differences in the presence and absence in cluster of orthologous genes (COGs), mutations, such as single nucleotide polymorphisms (SNPs) and small insertion and deletions (indels), and different DNA segments, known as unitigs. To create the input COG files, an .Rtab file of gene presence/absence was created using panaroo (Tonkin-Hill et al. 2020). For investigation in unitigs, unitig-counter (v1.1.0) (Jaillard et al. 2018) was used. For SNP/small indel GWAS an input vcf file was curated using freebayes (Garrison and Marth 2012). Bacterial GWAS was conducted using Pyseer v1.3.10 (Lees et al. 2018) which uses linear models with fixed or mixed effects to associate genetic variation with a phenotypic variable of interest, while accounting for confounding population structure in the bacterial population. To account for population structure, all analyses were supplemented with phylogenetic distances from a tree constructed using IQ-TREE (Nguyen et al. 2015). Furthermore, a minimum minor allele frequency (maf) cut-off of 0.05 (and maximum maf of 0.95) was used for the unitig and SNP GWAS analysis. All Pyseer analyses were run using the linear mixed model (LMM).

Five reference strains representing the two RSSC species, *R. pseudosolanacearum* and *R. solanacearum,* were used for unitig gene annotation. These were *R. pseudosolanacearum* strain GMI1000 (NCBI accession number: 000009125), *R. solanacearum* strain UW551 (NCBI accession number: 002251655), *R. solanacearum* strain UY031 (NCBI accession number: 001299555), *R. solanacearum* strain K60 (NCBI accession number: 002251695), *R. pseudosolanacearum* strain CMR15 (NCBI accession number: 000427195) and *R. solanacearum* strain YO199 (this study). Phandango software (Hadfield et al. 2018) was used to visualise Manhattan plots with GMI1000 (accession number: 000009125) as the reference strain.

For phenotype GWAS, trait values (as measured as above for each trait) were taken and mean centered. All 46 traits were inputted as a continuous phenotype, however binary input phenotypes were also calculated for all antibiotic resistant traits, cold tolerance, and reactive oxygen species tolerance, by taking growth (1) as OD_{600} over 0.13 at the last timepoint and no growth (0) as below this value. Survival at 4°C was also measured as a binary phenotype and used as an input phenotype. For the ecotype level GWAS, the input phenotype used was ecotype classification based on k-means clustering and PC1 and 2 values from figure 2.3B.

Accessory genomes of each ecotype were also investigated further as enriched KEGG pathways for each ecotype relative to the pangenome (all genes). Accessory genes were classified as being present in more than one isolate and less than 168 isolates (around 90%). All gene sequences were then functionally annotated with KEGG pathways using eggNOG-mapper (version emapper-2.1.9) (Huerta-Cepas et al. 2019, Cantalapiedra et al. 2021) with sequence searches performed using DIAMOND (Buchfink et al. 2021) with default setting apart from 0.00001 as the minimum e-value threshold and 'betaproteobacteria' as the annotation taxa group. Enrichment analysis was then completed to determine which pathways were enriched within each ecotypes accessory genome compared to the pangenome of the RSSC collection. This was conducted within R version 4.2.1 (R Core Team 2022) using the clusterProfiler package (Wu et al. 2021) with the Benjamini-Hochberg (BH) multiple testing correction method and 0.05 as the p value and q value cut-off.

2.3.8 Statistical analysis and data visualisation

Statistical data analysis and visualisation were conducted within R version 4.2.1 (R Core Team 2022). Figures were produced using the package ggplot2 (Wickham 2016) and heatmaps were curated using the Heatmap function in the ComplexHeatmap package (Gu et al. 2016). Venn diagrams were curated using the ggVennDiagram (Gao 2021) package, while upset plots were curated using the package ComplexUpset (Krassowski et al. 2020) and ggtree (Yu et al. 2017, 2018, Yu 2020, 2022) was used to visualise the RSSC phylogeny.

Area under the growth curve (AUC) was calculated as a proxy for growth for most traits, as AUC captures lag time, growth rate and carrying capacity, using the MESS package in R version 4.2.1 (Ekstrøm 2019, R Core Team 2022). Standardised trait values were z-scores curated using the scale function in the stats R package (R Core Team 2022) and mean standardised trait values across the trait groups (appendix table B.1) was taken as the mean z-score across all traits within that group. Principal component analysis (PCA) was carried out using tidymodels (Kuhn and Wickham 2020) with normalisation of the data. K-means clustering was conducted using the kmeans function in stats R package using Euclidean distances (R Core Team 2022) and optimal number of clusters were determined using the cascadeKM function in the vegan package (Oksanen et al. 2019), using simple structure index (ssi) as the criterium.

Significance of clusters and groups within PCA plots was determined with PERMANOVA analysis using the adonis2 and betadisper functions in the vegan package (Oksanen et al. 2019). Post hoc analysis was conducted using pairwise PERMANOVA using the pairwise.perm.manova function in the RVAideMemoire package (Hervé 2022) with PC1 and PC2 values as the input variables and the fdr p value adjustment method. Significant differences of mean trait values per phenotype group between each species and ecotype was determined using two-way ANOVA with the aov function, and post hoc using the TukeyHSD function, in the stats R package (R Core Team 2022), using the standardised trait value as the dependent variable and species (or ecotype) and phenotype group as the explanatory variables. Mean pairwise distances were calculated by taking the mean Euclidean distance per isolate from all other RSSC isolates, using the dist() function (R Core Team 2022) on a matrix of the 46 phenotypic trait values. The median of the mean Euclidean distance between different species and ecotypes was then compared using Kruskal-Wallis significance test, using the kruskal.test function, and pairwise Wilcoxon test, using the pairwise.wilcoxon.test function, both available within the stats R package (R Core Team 2022). Significance of host, decade and species distribution across the five ecotypes was calculated using chi squared tests with the chisq.test function in the stats R package (R Core Team 2022), and post hoc analysis using the chisq.posthoc.test package (Ebbert 2019). Significance of cold

tolerant gene hit (*hxcR*/*epsE*) distribution across the ecotypes was also calculated using chisq.test function and the chisq.posthoc.test package (Ebbert 2019, R Core Team 2022).

2.4 Results

2.4.1 Two *Ralstonia solanacearum* species complex (RSSC) species show only small phenotypic dissimilarity

To compare the phenotypic diversity of two RSSC species, 46 trait measurements were collected across a collection of 194 global RSSC isolates, spanning eight decades (1945 to 2016). Isolates assigned as *R. solanacearum* (n=119) and *R. pseudosolanacearum* (n=66), through whole genome sequencing analysis, were compared to see if genetics can determine trait diversity patterns and trait specialization within the RSSC. A principal component analysis (PCA) was conducted to visualize the phenotypic differences between the isolates (figure 2.2B). This PCA plot explained 46% of the total variation within the dataset, with the majority (30%) explained by principal component 1 (PC1) and 16% by principal component 2 (PC2). A lower PC1 value corresponds to an increase in growth across most traits, apart from nutrient limited conditions and heat tolerance traits where a positive PC1 value corresponds to higher growth (figure 2.2A). Furthermore, PC2 shows differences in values across the traits explored, with higher values showing higher growth on carbon resources, and an increase in antibiotic resistance and biofilm production traits, while negative values indicate higher stress tolerance within extreme pH and salinity conditions (figure 2.2A). Both species revealed large phenotypic diversity and they differed significantly from one another especially along PC2. However, this only accounted for a small amount of the trait variation (8%) observed between the RSSC isolates (figure 2.2B, PERMANOVA: $F_{1,183}$= 15, R^2= 0.08, p=0.001). A large overlap along PC1 and PC2 in trait diversity was observed across the two species and no significant differences in dispersion was observed (figure 2.2B: ANOVA: : $F_{1,183}$= 2.5, p=0.11). However, exploring each isolate's mean pairwise distance from all other isolates, calculated using Euclidean distances across all 46 traits, also revealed that *R. pseudosolanacearum* isolates were more diverse,

and therefore had a significantly higher mean pairwise distance compared to *R. solanacearum* isolates (figure 2.2C, Kruskal-Wallis: $X^2_1 = 13.7$, p=0.0002).

Next, to explore how these two species differ in specific trait groups, their standardised trait values for each of the 13 phenotypic groups (appendix table B.1) were compared (figure 2.2D). This revealed that the two species only differed significantly in their ability to grow on complex media and nutrient limited conditions, with *R. pseudosolanacearum* having significantly higher growth within complex media compared to *R. solanacearum* (figure 2.2D, ANOVA: $F_{1,1937}$= 0.30, p<0.0001, TUKEY: p=0.0004) and *R. solanacearum* isolates having significantly higher growth in nutrient limited conditions compared to *R. pseudosolanacearum* isolates (figure 2.2D, ANOVA: $F_{1,1937}$= 0.30, p<0.0001, TUKEY: p=0.018). This suggests that each species specializes in different niches, with *R. solanacearum* focusing on survival within nutrient limited conditions, while *R. pseudosolanacearum* has higher metabolic efficiency, with higher growth within complex media (figure 2.2D). However, for most traits the two species were not significantly different and large phenotypic diversity was observed within both species (figure 2.2B). This suggests that each species could have overlapping niches maintaining the large phenotypic diversity observed among both species or more specific test conditions are required.

The RSSC collection metadata was also used to explain variation in the phenotypic data. First, to explore how geographical separation explains RSSC trait diversity the same PCA plot as figure 2.2B was created but with isolates coloured by the continent in which they were isolated. This shows a significant, although small, difference in trait diversity depending on which continent they were from (appendix figure B.2D, PERMANOVA: $F_{5,179}$= 3.2, R^2= 0.08, p=0.002) and no significant dispersion (appendix figure B.2D, ANOVA: $F_{5,179}$= 0.8, p=0.53). This was mainly explained by isolates from Europe being different compared to other isolates from other continents, specifically Asia, Africa, North America, and South America (pairwise PERMANOVA: p=0.037 for all). Next, the amount of trait diversity explained by the isolation host or source was investigated. Due to the large number of unique hosts in this dataset, metadata were filtered so that only hosts represented by five or more isolates were selected for this analysis, which resulted

in a dataset of 133 isolates from the following hosts: potato (n= 75), tomato (n=25), banana (n=13), ginger (n=8), water (n=6), and tobacco (n=6). The analysis showed that isolation source also explained a small but significant amount of trait variation within the collection (appendix figure B.2E, PERMANOVA: $F_{5,127}$= 2.4, R^2= 0.09, p=0.014) with no significant difference in dispersion around the centroid (appendix figure B.2E, ANOVA: $F_{5,127}$= 1.8, p=0.12). However, pairwise PERMANOVA analysis revealed no significant differences between host groups, most likely due to small sample sizes. In conclusion, phenotypic diversity of RSSC was explained by species, isolation source and geographical location of isolation, however the percentage of variation explained by each factor was relatively small (between 8 and 9%).

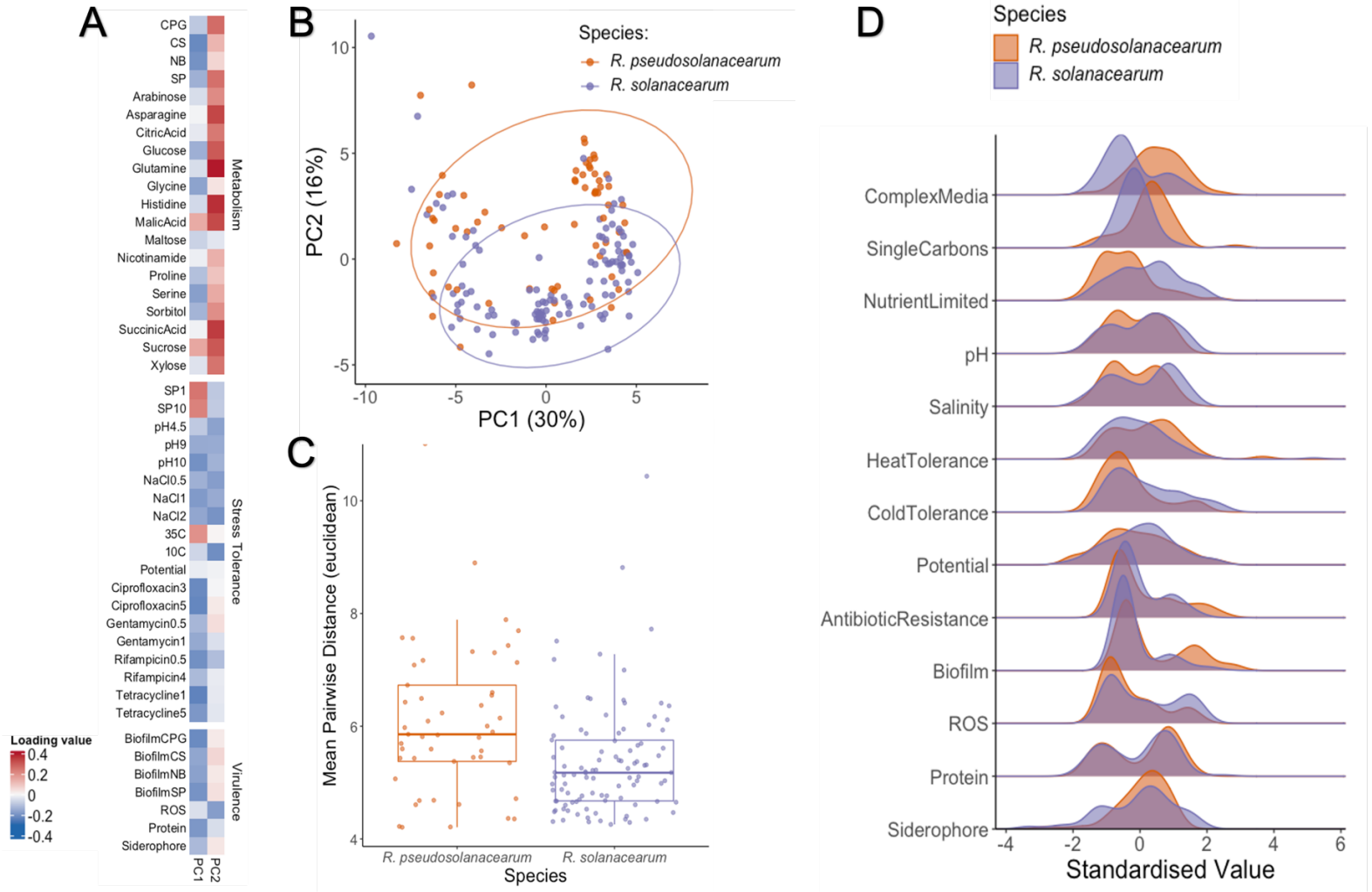

A
CPG
CS
NB
SP
Arabinose
Asparagine
CitricAcid
Glucose
Glutamine
Glycine
Histidine
MalicAcid
Maltose
Nicotinamide
Proline
Serine
Sorbitol
SuccinicAcid
Sucrose
Xylose
Metabolism
SP1
SP10
pH4.5
pH9
pH10
NaCl0.5
NaCl1
NaCl2
35C
10C
Potential
Ciprofloxacin3
Ciprofloxacin5
Gentamycin0.5
Gentamycin1
Rifampicin0.5
Rifampicin4
Tetracycline1
Tetracycline5
Stress Tolerance
BiofilmCPG
BiofilmCS
BiofilmNB
BiofilmSP
ROS
Protein
Siderophore
Virulence
PC1
PC2
Loading value
0.4
0.2
0
-0.2
-0.4
B
Species:
R. pseudosolanacearum
R. solanacearum
PC2 (16%)
PC1 (30%)
C
Mean Pairwise Distance (euclidean)
R. pseudosolanacearum
R. solanacearum
Species
D
Species
R. pseudosolanacearum
R. solanacearum
ComplexMedia
SingleCarbons
NutrientLimited
pH
Salinity
HeatTolerance
ColdTolerance
Potential
AntibioticResistance
Biofilm
ROS
Protein
Siderophore
Standardised Value

Figure 2.2: Two *Ralstonia solanacearum* species complex (RSSC) species share similar ecological diversity, differing in diversity and only some trait groups. (B) Principal component analysis (PCA) plot with each point as one of the isolates within the RSSC collection coloured by species, *R. pseudosolanacearum* or *R. solanacearum*, assigned using whole genome sequencing (n=185). Principal component (PC) 1 explained 30% of the variation within the dataset and PC2 16%. The amount of contribution each trait has on the principal components is shown in a heatmap in panel (A) Orange and purple points show *Ralstonia pseudosolanacearum* and *R. solanacearum* isolates respectively. (C) Mean pairwise difference to all other isolates calculated per isolate and grouped by species. Pairwise distances were measured using Euclidean distances. Orange and purple points show *Ralstonia pseudosolanacearum* and *R. solanacearum* isolates respectively. Boxplots lines show the median per species of the mean pairwise distances with the box indicating the interquartile range and whiskers showing the 95% quantile range. (D) Standardised values (z-score) distribution across each species for traits grouped into 13 ecologically relevant categories are shown in a density plot.

2.4.2 *Ralstonia solanacearum* species complex (RSSC) isolates cluster into five distinct phenotypic 'ecotypes'

As species identity could not explain most of the trait variation seen among RSSC isolates, further analysis exploring if strains could be clustered based on their phenotypic traits was conducted. This revealed that RSSC isolates clustered into five groups based on their phenotypic diversity of 46 traits (figure 2.3A), assigned by k-means clustering and ssi criterion (appendix figure B.3A). These can be seen clearly in the PCA plot (figure 2.3B) where isolates from each group are significantly different from one another phenotypically (figure 2.3B, PERMANOVA: $F_{4,189} = 232$, $R^2 = 0.83$, $p = 0.001$) and ecotype 5 having significant higher dispersion around the centroid compared to ecotypes 2 and 3 (figure 2.3B, ANOVA: $F_{4,189} = 3.8$, $p = 0.005$, TUKEY: $p = 0.02$ for ecotypes 5 and 2 and $p = 0.01$ for ecotypes 5 and 3). As this phenotypic differentiation is likely to affect the fitness and ecology of RSSC isolates within the environment, groups are henceforth called 'ecotypes'. Ecotypes also differ in their trait diversity (appendix figure B.3B, Kruskal-Wallis: $X^2_4 = 126$, $p<0.0001$), as calculated as each isolate's mean pairwise Euclidean distance from all other isolates. Ecotype 5 was the most diverse ($p<0.0001$-0.004 compared to all other ecotypes), followed by ecotype 4 ($p<0.0001$-0.01 for all other ecotypes), then 2 ($p<0.00001$-0.01 for all) then ecotype 1 ($p<0.0001$ for all), and with ecotype 3 being the least diverse of them all ($p<0.0001$-0.01 compared to all other ecotypes).

Isolates from different ecotypes also differed in their trait specialisation (figure 2.3A), with significantly different standardised value distributions (figure 2.3C) across 13 phenotypic groups (for trait groupings see appendix table B.1). RSSC isolates from ecotype 1 (n = 60) could be considered as oligotroph specialists, with higher standardised values within nutrient limited conditions compared to all other RSSC ecotypes (figure 2.3C, ANOVA: $F_{4,2457} = 155$, $p<0.0001$, TUKEY: $p<0.0001$ for all), and lower standardised values across all other traits, especially growth within complex media (figure 2.3C, ANOVA: $F_{4,2457} = 155$, $p<0.0001$, TUKEY: $p<0.0001$ compared to ecotypes 2, 4 and 5). RSSC isolates assigned to ecotype 2 (n=31) could be regarded as the heat tolerant specialist isolates, with significant higher standardised values when grown in high temperatures compared to ecotypes 4 and 5 (figure 2.3C, ANOVA: $F_{4,2457} = 155$, $p<0.0001$, TUKEY: $p<0.0001$ for both ecotypes)

and significant lower trait values when grown in cold temperatures compared to ecotypes 3 and 4 (figure 2.3C, ANOVA: $F_{4,2457}$= 155, p<0.0001, TUKEY: p<0.0001 for ecotype 3 and p=0.003 for ecotype 4). In contrast ecotype 3 isolates (n=49) were cold tolerant with the highest standardised values within this condition compared to all other ecotypes (figure 2.3C, ANOVA: $F_{4,2457}$= 155, p<0.0001, TUKEY: p<0.0001 for all ecotypes). This ecotype also had the highest growth in the presence of reactive oxygen species (ROS) (figure 2.3C, ANOVA: $F_{4,2457}$= 155, p<0.0001, TUKEY: p<0.0001 for all ecotypes) and is one of three ecotypes (along with ecotypes 4 and 5) that had higher standardised values in pH and salinity stress compared to ecotypes 1 and 2 (figure 2.3C, ANOVA: $F_{4,2457}$= 155, p<0.0001, TUKEY: p<0.0001). Ecotypes 4 (n=29) and 5 (n=25) were both antibiotic resistant and biofilm producers, with higher standardised values in these traits compared to the other three ecotypes (figure 2.3C, ANOVA: $F_{4,2457}$= 155, p<0.0001, TUKEY: p<0.0001 for all). These two ecotypes, however, differed in their metabolic traits with ecotype 5 having the highest trait values in complex media, and significantly higher standardized values in single carbon media compared to ecotype 4 (figure 2.3C, ANOVA: $F_{4,2457}$= 155, p<0.0001, TUKEY: p<0.0001). Overall, the two species within the RSSC investigated in this study are phenotypically diverse, clustering into five distinct ecotypes, each differing in their trait specialization.

To investigate ecotype differences in more detail, the significance of metadata across the five ecotypes was compared. This revealed that species distribution among the five ecotypes was significantly different (appendix figure B.4A, X^2_4= 48, p<0.0001), with a higher proportion of isolates from the nutrient limited specialist ecotype (ecotype 1) (r=3.3, p=0.009) and cold tolerant ecotype 3 (r=3.1, p=0.017) belonging to the *R. solanacearum* species, and a higher proportion of heat tolerant ecotype 2 isolates belonging to the *R. pseudosolanacearum* species (r=6.1, p<0.0001). However, representation from both species was seen in all five ecotypes (appendix figure B.4A), supporting the large overlap in phenotypic diversity between the two species observed in figure 2.2. Furthermore, continent distribution was also uneven across the five ecotypes (appendix figure B.4B, X^2_{20}= 37, p= 0.013). Isolates from the heat tolerant ecotype 2 (r=-3.3, p=0.03) were less likely to be from Europe and isolates from the cold tolerant ecotype 3 (r=-3.5

p=0.015) were less likely to be isolated from Asia, with no isolates from either continent in these respective ecotypes. Host distribution among the five ecotypes was also uneven (appendix figure B.4C, X^2_{20}= 44, p= 0.0015). With isolates from the heat tolerant ecotype (ecotype 2) more likely sampled from ginger hosts (r=4.2, p=0.0009) and isolates from the heat tolerant ecotype (ecotype 2) less likely sampled from potato crops (r=-3.1, p=0.05). Some ecotypes were also missing isolates from certain hosts, however due to the low overall sample sizes from certain hosts we cannot conclude if this is a statistically significant trend. Overall, five phenotypically distinct ecotypes within the RSSC were identified, which were evenly distributed within both species.

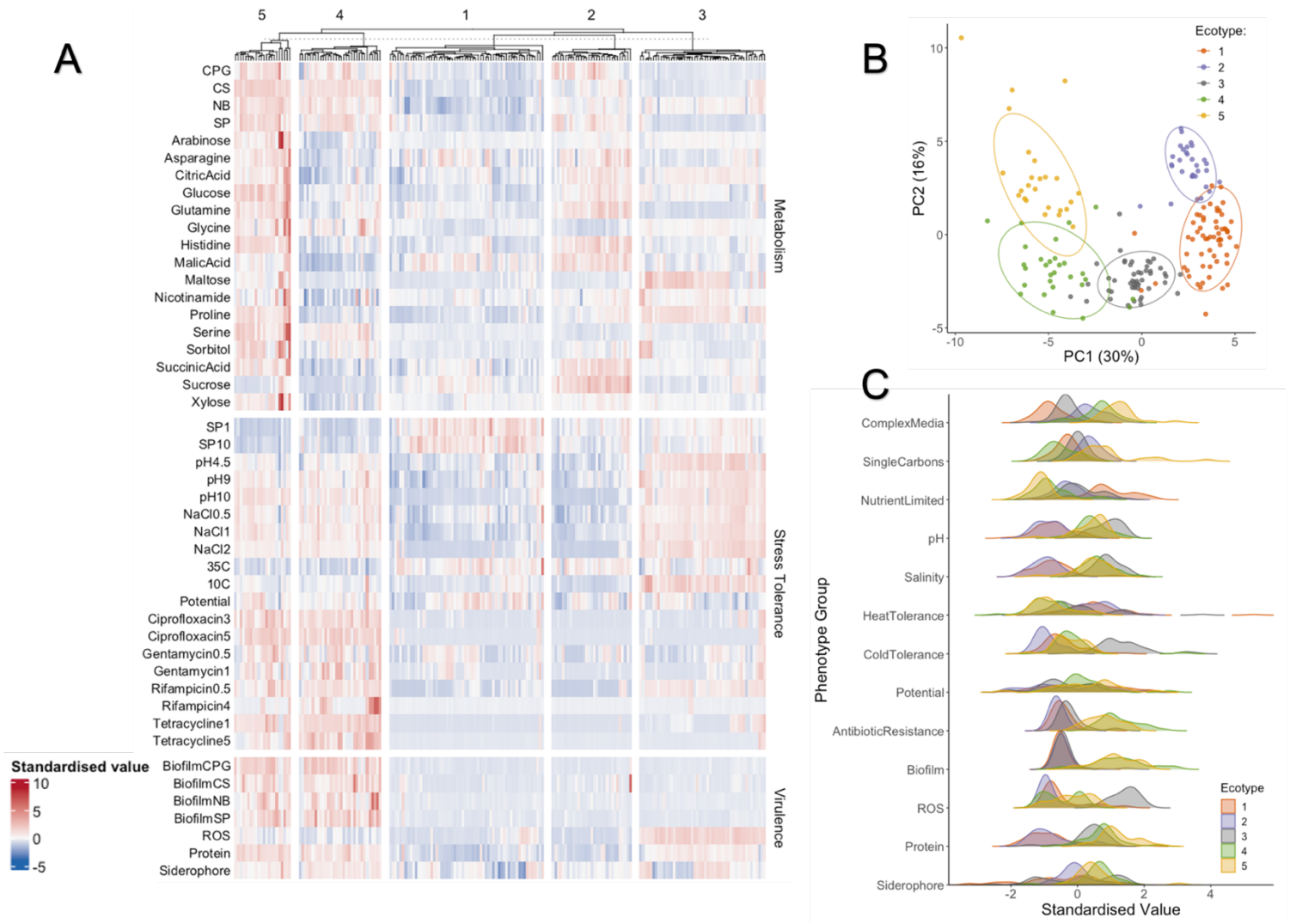

A
5
4
1
2
3
CPG
CS
NB
SP
Arabinose
Asparagine
CitricAcid
Glucose
Glutamine
Glycine
Histidine
MalicAcid
Maltose
Nicotinamide
Proline
Serine
Sorbitol
SuccinicAcid
Sucrose
Xylose
SP1
SP10
pH4.5
pH9
pH10
NaCl0.5
NaCl1
NaCl2
35C
10C
Potential
Ciprofloxacin3
Ciprofloxacin5
Gentamycin0.5
Gentamycin1
Rifampicin0.5
Rifampicin4
Tetracycline1
Tetracycline5
BiofilmCPG
BiofilmCS
BiofilmNB
BiofilmSP
ROS
Protein
Siderophore
Metabolism
Stress Tolerance
Virulence
Standardised value
10
5
0
-5
B
Ecotype:
1
2
3
4
5
PC2 (16%)
PC1 (30%)
C
Phenotype Group
ComplexMedia
SingleCarbons
NutrientLimited
pH
Salinity
HeatTolerance
ColdTolerance
Potential
AntibioticResistance
Biofilm
ROS
Protein
Siderophore
Ecotype
Standardised Value

Figure 2.3: *Ralstonia solanacearum* species complex (RSSC) clusters into five separate phenotypic 'ecotypes'. (A) Heatmap showing all 194 isolates and their z-score standardised values in each of the 46 traits measured. K-means clustering (k=5) was used to split isolated into 5 groups. Traits can be split into three categories, metabolism, stress tolerance and virulence. (B) Principal component analysis (PCA) with each point as one of the 194 isolates within the RSSC collection, principal component (PC) 1 explains 30% of the variation within the dataset and PC2 16%. The amount of contribution each trait has on the principal components is shown in a heatmap in figure 2.2A. Isolates are coloured by ecotype assigned via K-means clustering, five optimal groups were determined using ssi criterion. Ellipses show the 90% confidence intervals around the centroid of each group. (C) Standardised values (z-score) distribution across each ecotype for traits grouped into 13 ecologically relevant categories are shown in a density plot.

2.4.3 Ecotype differences within the RSSC can be explained by accessory genome variation

The ecotype distribution across the RSSC phylogeny, as determined using core genome differences, showed that all ecotypes were present across the phylogeny and across both species (figure 2.4A). Figure 2.4A also highlights that ecotype 1 is more commonly found within *R. solanacearum* (coloured as purple on the phylogeny) and ecotype 2 is more commonly found within the *R. pseudosolanacearum* clade (orange) on the phylogeny (figure 2.4A). However, all ecotypes were present across both clades with no clear ecotype clustering, suggesting core genetic differences may not be contributing much towards ecotype differences.

Next, accessory genome variation between each ecotype was explored (figure 2.4B and C). This revealed that while 44% of genes were shared across all five ecotypes, a vast number of genes were unique within each ecotype, with 11%, 7%, 1%, 6% and 4% of total genes being unique to ecotypes 1, 2, 3, 4 and 5, respectively. As most ecotype specific genes were only present within one isolate, or were annotated as hypothetical proteins, genetic differences between each ecotype were further explored by comparing enriched pathways within the accessory genomes of each ecotype group relative to the whole RSSC pangenome. Accessory genes were determined as those present in more than one isolate and fewer than 90% of isolates (168 isolates). These accessory genes for each ecotype were then compared to the pan genome (all genes across all isolates; Stoycheva et al. 2022: manuscript in progress) to reveal which pathways, using the KEGG database, were enriched within each ecotype's accessory genome (figure 2.4D). Overall, for all ecotypes, the most significantly enriched pathway within their accessory genomes were genes involved in microbial metabolism in diverse environments, suggesting that metabolism differentiation is the main driver of RSSC diversification. Furthermore, genes involved in starch and sucrose metabolism, tyrosine metabolism and xylene degradation were all enriched within the accessory genomes of all five ecotypes. Pathways associated with the degradation of aromatic compounds pathway were enriched within the accessory genomes of all ecotypes, apart from ecotype 5. Benzoate degradation genes were enriched within ecotypes

1, 2 and 4, and aminobenzoate degradation genes were enriched within the accessory genomes of ecotypes 1 and 2. Furthermore, ecotypes 1 and 4 had glycolysis/gluconeogenesis genes enriched within their accessory genomes as well. Unique pathways were also found enriched within ecotypes accessory genomes. Ecotype 1 had the fatty acid metabolism and galactose metabolism pathways enriched within their accessory genome. Ecotype 2 had genes involved in quorum sensing overrepresented within their accessory genome. Ecotype 4 had pyruvate metabolism enriched within their accessory genome and ecotype 5 had genes involved in methane metabolism enriched (figure 2.4D). Ecotype 3, the cold tolerant group, had no unique enriched pathways, however this ecotype also had the smallest number of unique genes (figure 2.4B). This highlights differences between the ecotypes, potentially reflecting their phenotypic differences. This suggests that accessory genome variation could be responsible for driving ecotype differences. Also, a GWAS was conducted, using SNPs/small INDELs, presence and absence of cluster of orthologous genes (COGs) and DNA segments known as unitigs as the input genetic variation. This analysis resulted in no significant hits associated with ecotype. Therefore, despite large trait differences between ecotypes, and high trait similarity between isolates within the same ecotype, our analysis suggests that different genetic mechanisms could be responsible for phenotypic differences between the ecotypes.

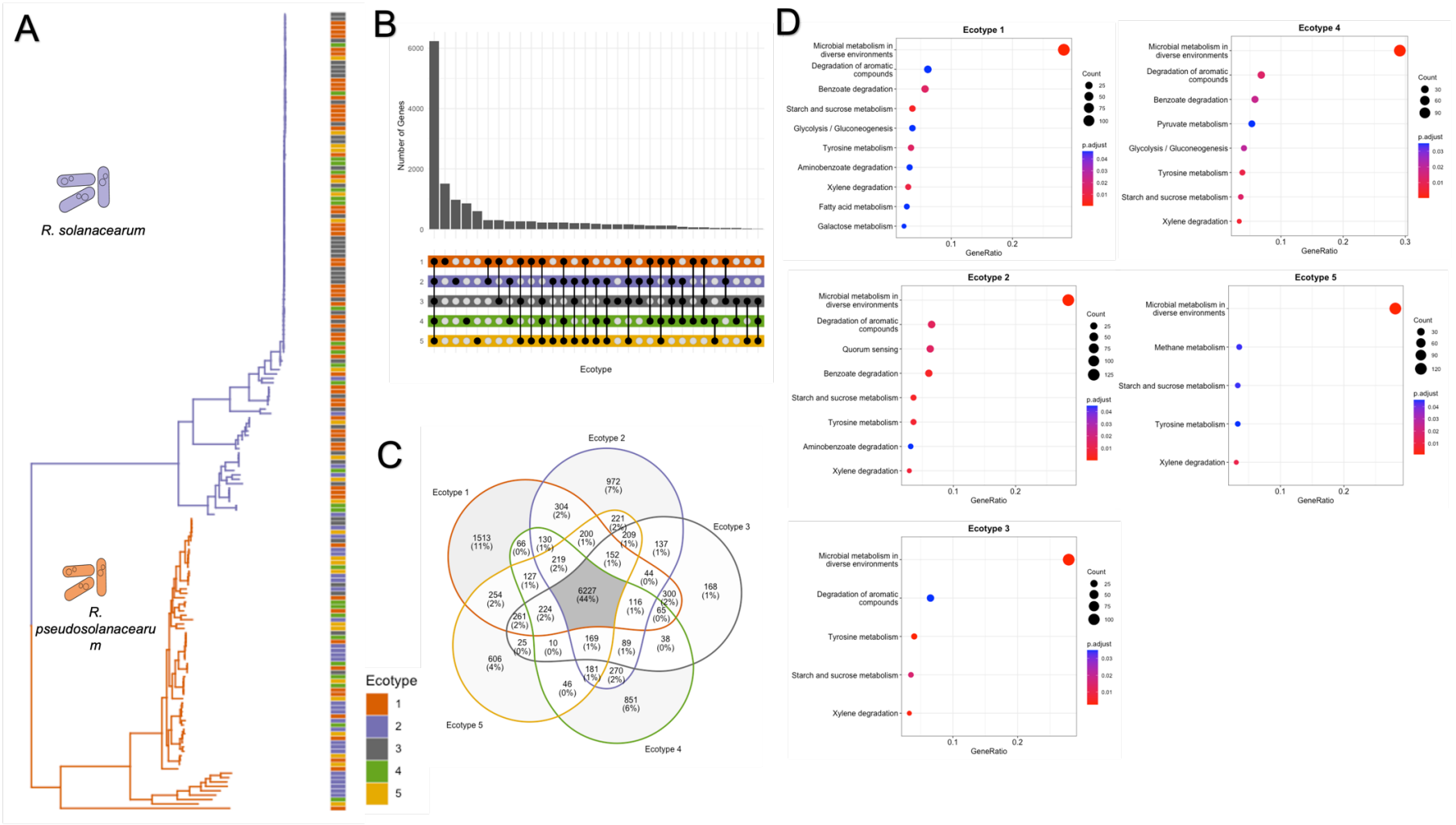

A
R. solanacearum
R. pseudosolanacearum
B
Number of Genes
6000
4000
2000
0
1
2
3
4
5
Ecotype
C
Ecotype 1
Ecotype 2
Ecotype 3
Ecotype 4
Ecotype 5
1513 (11%)
972 (7%)
168 (1%)
606 (4%)
851 (6%)
6227 (44%)
304 (2%)
221 (2%)
200 (1%)
209 (1%)
137 (1%)
66 (0%)
130 (1%)
219 (2%)
152 (1%)
44 (0%)
127 (1%)
254 (2%)
300 (2%)
116 (1%)
65 (0%)
224 (2%)
261 (2%)
25 (0%)
10 (0%)
169 (1%)
89 (1%)
38 (0%)
181 (1%)
270 (2%)
46 (0%)
Ecotype
1
2
3
4
5
D
Ecotype 1
Microbial metabolism in diverse environments
Degradation of aromatic compounds
Benzoate degradation
Starch and sucrose metabolism
Glycolysis / Gluconeogenesis
Tyrosine metabolism
Aminobenzoate degradation
Xylene degradation
Fatty acid metabolism
Galactose metabolism
GeneRatio
Count
p.adjust
Ecotype 4
Microbial metabolism in diverse environments
Degradation of aromatic compounds
Benzoate degradation
Pyruvate metabolism
Glycolysis / Gluconeogenesis
Tyrosine metabolism
Starch and sucrose metabolism
Xylene degradation
GeneRatio
Count
p.adjust
Ecotype 2
Microbial metabolism in diverse environments
Degradation of aromatic compounds
Quorum sensing
Benzoate degradation
Starch and sucrose metabolism
Tyrosine metabolism
Aminobenzoate degradation
Xylene degradation
GeneRatio
Count
p.adjust
Ecotype 5
Microbial metabolism in diverse environments
Methane metabolism
Starch and sucrose metabolism
Tyrosine metabolism
Xylene degradation
GeneRatio
Count
p.adjust
Ecotype 3
Microbial metabolism in diverse environments
Degradation of aromatic compounds
Tyrosine metabolism
Starch and sucrose metabolism
Xylene degradation
GeneRatio
Count
p.adjust

Figure 2.4: Genetic differences between the five identified ecotypes. (A) core genetic phylogeny shows all five ecotypes are spread across the RSSC phylogeny and across the two species, *R. pseudosolanacearum* and *R. solanacearum*. Two species are coloured on the phylogenetic tree, orange being *R. pseudosolanacearum* and purple *R. solanacearum*. (B) Accessory gene differences, gene presence and absences, across the five ecotypes as shown as an upset plot and (C) Venn-diagram. (D) Enrichment analysis on the five ecotypes accessory genomes highlights overrepresented KEGG pathways. Accessory genes were selected as present within more than one isolate and less than 90% of isolates (168 isolates). The genes meeting these requirements within each ecotype was then compared to the pan-genome (all genes) to see which pathways were enriched. Pathways were annotated using the KEGG database. A p-value cut-off of 0.05 was used and multiple testing correction method of Benjamini-Hochberg (BH) was conducted. Gene ratio is calculated as the number of enriched genes belonging to a given gene-set divided by the total number of genes in the gene-set. With separate graphs for each of the five ecotypes.

2.4.4 A GWAS reveals potential causative genes for cold tolerance and rifampicin resistance traits within the RSSC

As the ecotype level GWAS found no significant genetic associations with the five identified ecotypes, another GWAS was conducted for each of the single trait measurements across the whole RSSC collection (182 isolates) including both species, as sample size is critical for having enough power to conduct a bacterial GWAS. This analysis revealed no clear associations for most traits due to strong population structure and linkage disequilibrium present among bacterial isolates. However, some promising genetic associations were discovered. A few genes were found significantly associated with cold tolerance traits, 'cold survival at 4°C' (binary_cold_survival) and 'growth at 10°C' (binary_low_temp), when COGs were used as the input genetic variable (appendix table B.2). A gene annotated as a type II secretion system protein (*hxcR*/*epsE*) was identified as significantly associated with both cold temperature traits, suggesting a higher likelihood of this gene being a true association with cold tolerance. Further exploration of the presence of this gene among the RSSC collection revealed that it was distributed unevenly across the five ecotypes (figure 2.5B, Chi-squared: X^2_4= 13.6, p=0.009), with a significantly lower proportion of isolates within the heat tolerant ecotype 2 having this gene (p=0.005). This ecotype had lowest growth within cold conditions and therefore suggests that this gene could be involved in tolerance to cold temperatures within the RSSC or that the absence of this gene promotes heat tolerance (figure 2.5A and 2.3C).

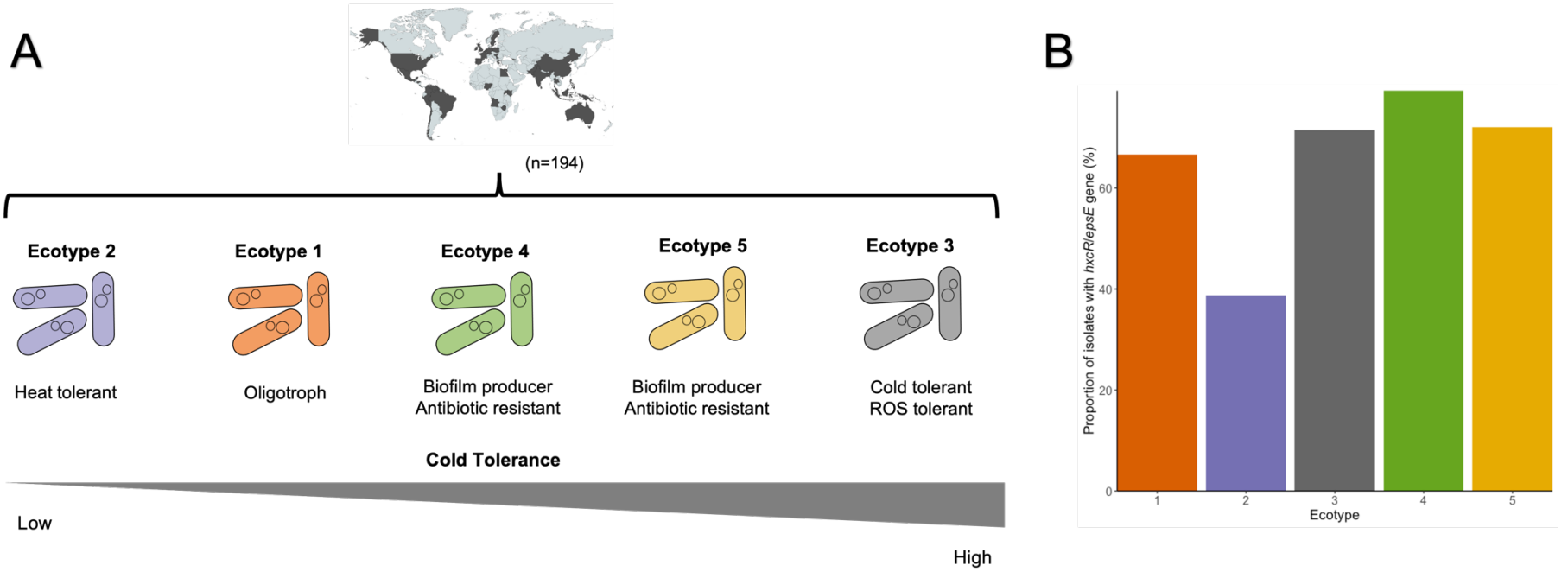


Figure 2.5: A gene associated with cold tolerance is represented less within the heat tolerant ecotype 2. (A) schematic summarising the five different ecotypes identified in the RSSC. Ecotypes are based on phenotypic differences among 46 traits and 194 RSSC isolates. (B) proportion of isolates (%) with the type II secretion system protein (*hxcR*/*epsE*) gene per ecotype. This gene was found significantly associated with cold tolerance using COG GWAS technique (n=182).

Associations between rifampicin resistance at 0.5µg/ml (as a binary trait) was also discovered using both COGs and unitigs as the input genetic variable (appendix table B.3). Most of these hits were annotated as hypothetical genes with some COGs annotated including a PCP degradation transcriptional activation protein (*pcpR*), a toxin (*fitB*) and an HTH-type transcriptional regulator (*gltR*). For the unitig GWAS, qq plots show a normal distribution of p-values suggesting that there was little population structure confounding these results (figure 2.6A). While manhattan plots (figure 2.6B) revealed three regions within the genome, two in the chromosome (between 870-900kb and 3.47-3.5Mb) and one in the megaplasmid (between 700-720kb), that were significantly associated with resistance to rifampicin (0.5 µg/ml). All significant unitigs were however annotated as being hypothetical proteins (figure 2.6C and D). These regions were explored in more detail using the best annotated RSSC isolate, the *R. pseudosolanacearum* GMI1000 strain (appendix figure B.5), as the reference. This revealed that the first region within the chromosome (870-900kb) mainly consisted of hypothetical proteins, with some genes annotated as probable transposase proteins, putative bacteriophage related-proteins and a putative DNA modification methylase protein, suggesting a role in gene regulation and insertion sequence movement for RSSC rifampicin resistance. The second significant chromosome region (3.47-3.5Mb) also mainly consisted of hypothetical proteins, along with a couple mobile elements, probable transmembrane proteins, transposase proteins, bacteriophage related proteins, a type III effector protein (*ripT*) and a putative drug efflux lipoprotein. Finally, the region within the megaplasmid significantly associated with rifampicin resistance consisted of mainly hypothetical proteins, putative transposase proteins (including isrsol6, isrsol10, is1421 insertion sequences), putative transmembrane proteins, a putative type III effector protein, putative bacteriophage proteins, a probable exoglucanase, a hypothetical signal protein and a putative lipoprotein. Together, these analyses suggest that a type II secretion system and insertion sequence movement might explain both cold and rifampicin tolerant RSSC genotypes in this dataset.

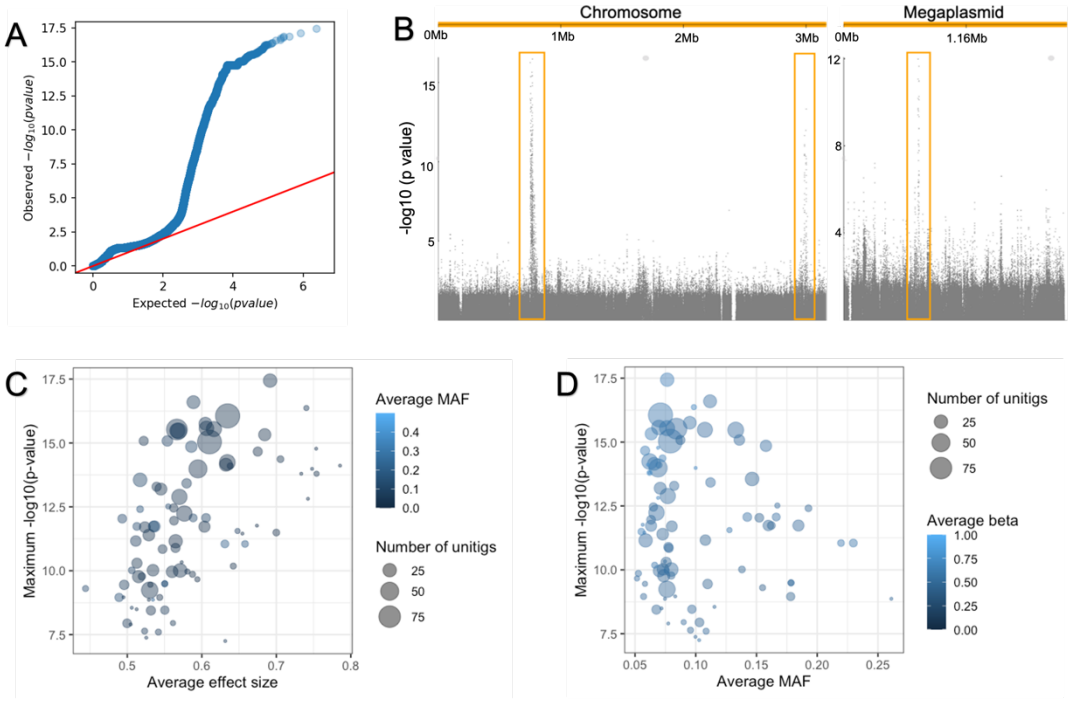


Figure 2.6: Rifampicin resistance at 0.5µg/ml, as a binary trait, is associated with differences within three regions of the genome. Linear mixed model pyseer GWAS technique was used with unitigs as the input genetic variant and rifampicin resistance (0.5µg/ml) as a binary trait as the input phenotype. (A) A qq-plot of unitig p-value distribution. (B) Manhattan plot of annotated unitigs positions along the bipartite genome, both chromosome and megaplasmid are shown. This highlights the positions where unitigs have higher p-value and are therefore associated with this trait. Identifying three regions within the genome that are significantly associated with rifampicin resistance. (C) and (D) show each annotated gene's p-value, average effect size (beta) and average minor allele frequency (MAF) along with the number of significant unitigs that are associated with this gene. All genes are annotated as hypothetical proteins.

2.5 Discussion

The aims of this research were to phenotypically compare two bacterial wilt plant pathogen species in the *Ralstonia solanacearum* species complex (RSSC), based on 46 ecologically relevant traits involved in metabolism, stress tolerance (abiotic and biotic) and virulence. Phenotyping a collection of 194 isolates revealed large similarities between the two species reflecting their similar life cycles. Five phenotypically distinct 'ecotypes', each differing in their trait specialisation (figure 2.7), were discovered across the whole collection, each comprising isolates from both species. Furthermore, genomic differences between these ecotypes could be attributed to accessory genome variation. Together, this suggests that the large diversity among the plant pathogen RSSC could be a product of adaptation to variable environmental niches with horizontal gene transfer driving this adaptation.

First, comparisons in the phenotypic diversity of the two widely distributed RSSC species, *R. pseudosolanacearum* and *R. solanacearum* were conducted to determine how these species differ (figure 2.2). This revealed that *R. pseudosolanacearum* was more diverse than *R. solanacearum,* with a large overlap in trait diversity between the two. This phenotypic diversity could reflect the fact that *R. pseudosolanacearum* is genetically more diverse, comprising of two phylotypes (I and II), compared to *R. solanacearum* which consists of only one phylotype (phylotype II) (Fegan and Prior 2005, Safni et al. 2014). However, we do have an overrepresentation of a single genotype strain, the cold tolerant phylotype IIB sequevar 1, within *R. solanacearum* data which could have also biased this result (Fegan and Prior 2005, Safni et al. 2014). Trait differences were also observed between the two species with *R. solanacearum* having higher trait values in limited nutrient conditions and *R. pseudosolanacearum* having higher growth within complex media. This could reflect life-cycle differences, or niche preferences, between the two species. This finding could also be explained by the smaller host range of phylotype IIB sequevar 1 strains, which survives mainly in vegetatively propagated potato or solanaceous weed species (van der Gaag et al. 2019), being present within the *R. solanacearum* species. Despite these differences there was large overlap in trait diversity between the two species which is surprising considering that it is thought that these species separated long ago, at the time of

continent split (Hayward 1991, Genin and Denny 2012). However, both species occupy similar habitats, with broad worldwide distributions, and cause the same disease symptoms, suggesting both species have extremely similar life cycles, and therefore, might be under similar selective pressures (Hayward 1991, Genin 2010, Bragard et al. 2019, EPPO 2022) Other research has also found that RSSC pathogen specificity was not phylotype specific (Lebeau et al. 2011), supporting that phylogeny or species does not correlate with RSSC isolate phenotypic profile. This implies that the ecological environment can have a large effect on phenotypic characteristics of bacterial species, with genetically different species showing highly similar traits if they occupy and have adapted to similar niches. This is further supported by the fact that the metadata, continent sampled from and isolation source, explained similar amounts of phenotypic variation (8 to 9%) than the two species did. Overall, *R. pseudosolanacearum* was more diverse but there were also large phenotypic and ecological niche overlap between the two species, justifying the species complex definition for this plant pathogen.

Exploration of the trait diversity among all RSSC isolates within this study was also conducted. This revealed a separation into five phenotypically distinct groups, or 'ecotypes', which each differed in their trait specialisations (figure 2.7). Ecotype 1 could be considered an oligotrophic group, with higher growth in low concentrations (1% and 10%) of sucrose peptone media (Mehan and McDonald 1995) compared to the other four ecotypes. This is an important trait to have for RSSC survival as between infections within environmental reservoirs they must cope with the stress of nutrient limited habitats, such is the case within river water (Álvarez et al. 2008). This ecotype also had the lowest growth in rich 'complex' media, suggesting a trade-off in overall growth, and therefore these isolates could be considered oligotrophic specialists. Nutrient availability has been previously linked with competitive ability and virulence of RSSC isolates (Yang et al. 2018, 2019, Li et al. 2021) supporting that growth within oligotrophic habitats could result in trade-offs between other aspects of the RSSC lifecycle.

Two other ecotypes also specialised in growth at either high (35°C) or low (10°C) temperatures. Ecotype 2 could be considered as the heat tolerant ecotype with higher growth within 35°C compared to the other ecotypes, while having

lowest trait values within cold (10°C) temperatures, suggesting a trade-off between the two temperature conditions. As a tropical pathogen, RSSC isolates are well adapted to grow at high temperatures, with optimal growth temperatures of around 35°C (Álvarez et al. 2010), and hence, it is not surprising that this pathogen species can grow well at this temperature. In contrast, ecotype 3 was the cold tolerant ecotype, with higher standardised trait values within cold temperatures (10°C) compared to the other ecotypes. The ability to grow, and therefore infect plants, at cold temperatures (10°C) will be extremely important for RSSC strains residing in temperate regions of the world, such as within Europe (Hayward 1991, Williamson et al. 2002, EPPO 2022). RSSC was first thought to become adapted to cold temperatures within the high altitudes in South America and was then moved to other temperate regions around the world along with the potato trade (Hayward 1991, Wicker et al. 2007, Safni et al. 2014). Therefore, this ecotype could consist of these cold adapted strains (race 3 biovar 2 or phylotype IIB) which has a lower optimal growth temperature at around 27°C (van der Gaag et al. 2019). Reactive oxygen species (ROS) tolerance was also highest within this ecotype which may be an adaptation of RSSC to infect plants within the colder temperate regions of the world, as both growth within colder temperatures and tolerance to ROS, produced by plants as a defence mechanism (Flores-Cruz and Allen 2009), will be required for successful infection.

Ecotype 4 and 5 both had higher standardised values in antibiotic resistance and biofilm production traits and differed in their ability to grow on single carbon media. Antibiotic resistance is an important trait for persisting within microbial communities, where antagonistic bacteria can produce a variety of antibiotics that kill RSSC species to reduce competition (Allen et al. 2010, Yuliar et al. 2015). Biofilm production is an important virulence trait for RSSC as they have been suggested to filter out nutrients from the flow of xylem fluid, as well as protect the pathogen from host plant immune defences (Álvarez et al. 2010, Genin and Denny 2012, Meng 2013). Biofilm production has also been shown to protect bacteria from antibiotics (Høiby et al. 2010), bacteriophage predators (Hosseinidoust et al. 2013) and other stresses within the environment, (de la Fuente-Núñez et al. 2013).

Suggesting that survival of general stresses, such as antibiotics, could be positively linked with biofilm production within the RSSC.

Furthermore, all ecotypes differed in their ability to utilise rich 'complex' media, with ecotype 5 having the highest growth measurements, ecotype 4, 2 and 3 having intermediate values, and ecotype 1 the lowest growth within complex media (figure 2.7). This could be an indication of niche separation between the five ecotypes to avoid competition (Bajic and Sanchez 2020). Other research, such as on the ocean living bacterium *Prochlorococcus*, has found separation into different genetic ecotypes based on environment nutrient availability (Kent et al. 2016), suggesting that nutrient availability is an important trait that divides bacterial species into separate groups. Overall, this presence of five separate ecotypes within the RSSC suggests that the large diversity observed among this plant pathogen (Hayward 1991, Genin 2010) could be due to niche separation or ecological diversification, resulting in separate ecotypes rather than generalist adaptation to a wide range of environments. This ecotype separation resulting in high microbial diversity has also been observed among other bacterial species, such as *Polynucleobacter necessaries* and *Limnohabitans* bacterial species (Jezbera et al. 2011, 2013, Larkin and Martiny 2017). High overall diversity could aid RSSC survival due to ecotype sorting in the face of selection, helping isolates avoid competition with one another and increases survival chances within diverse environments, or could also help survival due to temporal fluctuations within the habitat (Jezbera et al. 2011).

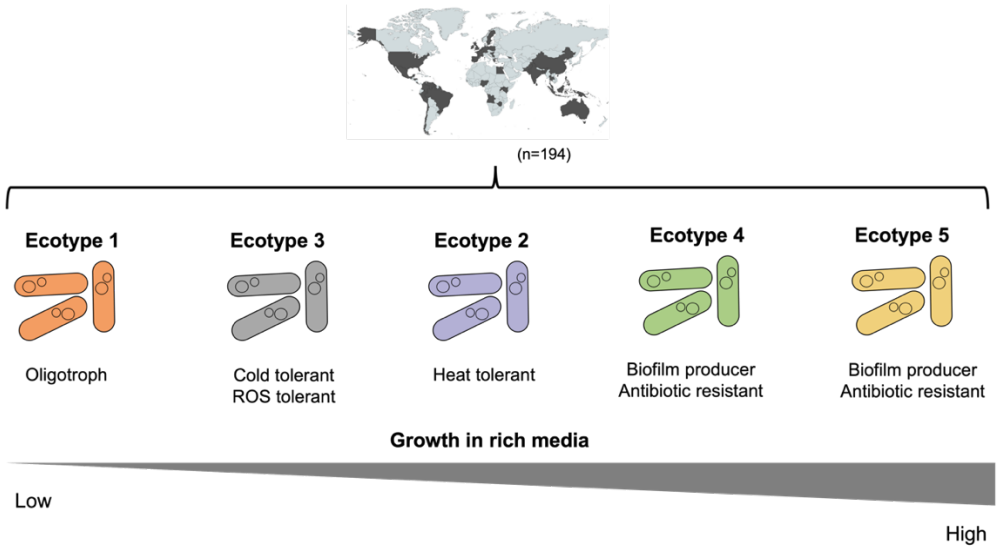


Figure 2.7: Summary schematic of the five assigned ecotypes. Ecotypes were determined by phenotypic variation within 46 separate ecologically relevant traits among a collection of 194 RSSC isolates.

The species distribution, as well as life-history data, was evaluated across the five ecotypes to determine if certain ecotypes can be linked to specific groups of isolates. All ecotypes were present within both species. However, distribution of species per ecotype was not equal with a higher proportion of oligotroph specialist, ecotype 1, and cold tolerant group (ecotype 3), assigned as *R. solanacearum* and a higher proportion of heat tolerant isolates being from the *R. pseudosolanacearum* species. Again, this could be due to the overrepresentation of cold tolerant *R. solanacearum* sequevar 1 phylotype IIB (race 3 biovar 2) strains within the RSSC collection (Fegan and Prior 2005, Safni et al. 2014).

Linking this trait variation with life-history data, the continent and host in which they were isolated from, also revealed that variation among RSSC isolates can be partly explained by these two factors. The continent in which an isolate was sampled from revealed that European isolates were different to other continents and are less likely to be assigned to heat tolerant ecotype 2. European RSSC isolates uniquely consists of the genetically similar cold tolerant *R. solanacearum* isolate (race 3 biovar 2 or phylotype IIB) (Fegan and Prior 2005, Safni et al. 2014), thought to have recently undergone a population expansion across the temperate regions of the world, disseminated through the potato trade (van der Gaag et al. 2019). Therefore, it is unsurprising that a lower proportion of European isolates was found

in the heat tolerant ecotype 2 which lacked adaptation to cool temperatures. Furthermore, a lower proportion of samples from Asia was also found to belong to the cold tolerant ecotype 3, suggesting that cold tolerance isn't an important trait to have within this region or could also be due to sampling bias within this region, such as the type of crops isolates were collected from within Asia.

The host or isolation source also explained small amounts of trait variation among isolates. Isolates sampled from potato crops were less likely assigned to the heat tolerant ecotype (ecotype 2), while isolates from ginger hosts were more likely to belong to the heat tolerant ecotype 2 group. This reflects the climate in which different crops are grown, with potato crops grown within temperate regions of the world, and grown within highlands or as a winter crops in tropical and subtropical regions (Campos and Ortiz 2020). In contrast ginger crops are mainly grown within Asian and South American countries as they have an optimal growth temperature between 27°C and 30°C (Retana-Cordero et al. 2021). It is thought that specific RSSC isolates pathogenic for certain hosts may have evolved only in certain parts of the world and are not found elsewhere (Hayward 1991), however while this research shows that certain ecotypes are more likely to infect certain hosts there are still great variation in host and continent origin across the five ecotypes. Overall, both the continent and host of origin, as well as the species, can partly explain the trait variation observed among the RSSC, suggesting that the life-history of isolates can determine trait diversity among the RSSC. However, the amount of trait variation explained is small suggesting that environmental differences could be the main factor driving phenotypic diversity.

The five identified ecotypes within this RSSC collection were next linked to genetic information to determine genetic mechanisms driving this diversification of isolates. All ecotypes were found present across both clades, representing the two species, in the RSSC phylogeny with no clear ecotype clustering, suggesting core genetic differences may not be contributing much towards ecotype differences. However, a large amount of accessory gene variation between the five ecotypes was observed, supporting theories that horizontal gene transfer plays a large role in the evolution and adaptation of the RSSC pathogen (Geng et al. 2022). KEGG pathways enriched within different ecotypes accessory genomes were explored,

revealing some similarities across all ecotypes and highlighting some interesting uniquely enriched pathways with some ecotype's accessory genomes. Most enriched pathways within each ecotypes accessory genome were involved in metabolism, suggesting that metabolic capacity drives RSSC ecotype diversification. Enrichment of metabolism genes driving ecotype separation within the RSSC is further supported by research on *Listeria,* where genes involved in nutrient transport were found responsible for survival within variable environments (Liao et al. 2021). Furthermore, the large differences in accessory genome variation, along with the absence of any significant genes associated with ecotype in a GWAS, both suggest that multiple different options are available for RSSC to achieve the same phenotype. Also, other genetic mechanisms not studied within this research, such as gene regulation, could also play a role in ecotype trait variation. One potential explanation for this result is that RSSC trait diversity is a product of adaptation towards different environmental niches and is potentially driven by horizontal gene transfer from the local environment.

Finally, by linking genetic data of the 182 RSSC isolates with specific phenotypes, exploration on the genetic causes of certain traits can be conducted. This revealed a novel type II secretion system protein (*hxcR*/*epsE*) associated with cold tolerance and less likely to be found within isolates in the heat tolerant ecotype 2, which had significant lower growth at 10°C compared to other ecotypes. Type II secretion systems (T2S) are a means by which bacteria secrete proteins, such as toxins, degradative enzymes, effectors, and novel proteins (Cianciotto 2005). While T2S are well known in their function during pathogenicity, other studies have also shown that T2S also promote growth of bacteria in environmental niches (Cianciotto 2005, Cianciotto and White 2017) and this study suggests that they could also have a role in cold tolerance within the RSSC. This finding is supported by a study on *Legionella pneumophila* which found that a type II secretion system gene (lsp) was required for growth at low temperatures (12°C to 25°C) (Söderberg et al. 2004). The type II secretion protein identified within this study has a heritability (h^2) of around 0.2 (appendix table B.2) suggesting that it cannot explain all the variation in cold tolerance among RSSC isolates. Therefore, gene regulation could also have a role in cold tolerance among this pathogen species which was found to be the case

in previous studies focusing on cold tolerance genetic mechanisms within the RSSC (Meng et al. 2015, Bocsanczy et al. 2017). Further experiments (i.e., constructing mutant strains) are also needed to confirm if this association between the type II secretion system protein is causative and not an artifact of population structure or chance.

Multiple genes were also found associated with resistance to the antibiotic rifampicin at 0.5µg/ml. These genes were annotated as hypothetical proteins and correspond to three regions of genome, two within the chromosome and one in the megaplasmid. Multiple putative transposase proteins were discovered in each of these three regions. Previous research has also shown that transposase genes are linked to antibiotic resistance genes within RSSC isolates and could be responsible for the gene transfer of these genes between bacterial isolates (Gonçalves et al. 2020) . Furthermore, insertion sequence movement within a *R. solanacearum* strain has been linked to ampicillin tolerance within laboratory (Alderley et al. 2022) further supporting this finding. This provides strong evidence that these are involved in resistance to rifampicin or are associated with a rifampicin resistance gene. However, as these genome regions were large (around 30kb in length), and comprise of mainly hypothetical proteins, further investigation within these regions is needed; for example, engineering of knock-out mutants to find out the causative gene or mutants (SNPs, INDELs, etc) responsible for this phenotype.

Most traits, however, could not be linked to genetic variation among the RSSC due to population structure and linkage disequilibrium present among the bacterial isolates confounding results. Another explanation could also be that genetic variation not captured within this study, such as epigenetics, could also be causing RSSC trait variation. RSSC isolates have multiple complex regulatory networks that respond to internal and environmental cues allowing high phenotypic plasticity (Schell 2000, Cellier and Prior 2010, Genin and Denny 2012, Perrier et al. 2019, Chen et al. 2022, Yan et al. 2022). An example of this are the quorum sensing regulatory systems present within this pathogen, such as the *PhcA* quorum-sensing pathway that controls the production of virulence factors in response to cell density (Genin and Boucher 2002, Genin and Denny 2012). Furthermore, transposable elements, such as insertion sequences (IS), and methylation patterns are highly

diverse among the RSSC (Erill et al. 2017, Gonçalves et al. 2020, Greenrod et al. 2022) and therefore could also cause differences among phenotypes within this bacterial pathogen. Overall, RSSC isolates can undergo phenotypic conversions depending on environmental conditions (Drenkard and Ausubel 2002) and this research highlights that there is still large trait variability among the RSSC with unexplained genetic causes. Future work should therefore focus on exploring these alternative genetic mechanisms causing phenotypic differences among this pathogen species.

In conclusion, the *Ralstonia solanacearum* species complex (RSSC) is a diverse collection of plant pathogenic bacteria, with a large overlap in phenotypic diversity between the two widely distributed bacterial species in this group. Phenotypic differences between the two species revealed *R. solanacearum* as being more oligotrophic, while *R. pseudosolanacearum* was identified as being more metabolically efficient. Furthermore, trait variation among 46 ecologically relevant phenotypes across the whole collection revealed five phenotypically distinct ecotypes, each of which comprised of isolates from both species. Linking this trait variation with genomic information revealed that accessory genome variation potentially drives ecotype differences, while a GWAS identified a type II secretion system associated with cold tolerance and three novel regions of the RSSC genome associated with rifampicin resistance. Overall, comparative analysis among a plant pathogenic group has revealed that phenotypic diversity among RSSC isolates could result from adaptation towards the environment, mediated by horizontal gene transfer from the local environment or other organisms within the environment.

3 Chapter 3: The UK *Ralstonia solanacearum* bacterial plant pathogen population has diversified into three 'ecotypes'

3.1 Abstract

Bacterial plant pathogen, *Ralstonia solanacearum,* a causative agent of bacterial wilt disease and potato brown rot, was first recorded in the UK in 1992. Since then, it has been regularly detected in river water and wild host plant (*Solanum dulcamara*) environmental reservoirs. Here, environmental sampling and high-throughput phenotyping was used to ask how diverse *R. solanacearum* is in the UK. 46 independent traits measured across 182 isolates, that span three decades of sampling, revealed that the UK population forms three distinct 'ecotypes' that differ in their growth in nutrient limited conditions, and overall trait generalism and specialism. Trait correlations also differed between these three ecotypes, implying that niche preferences and their trade-offs can be driving diversification. Linking trait differences with metadata highlighted that neither isolation source nor geographic location can explain this diversity. However, the decade in which isolates were sampled can help explain trait variation, indicating increased diversification over time. Specifically, antibiotic resistance and biofilm production are traits that have increased throughout time, suggesting that these are key phenotypic traits driving diversification within the UK pathogen population despite extremely high genetic similarity between isolates (98% shared genes across all three ecotypes). This study is the first to extensively focus on phenotypic diversity of UK *Ralstonia solanacearum* environmental isolates and therefore improves our understanding of how this bacterial pathogen may adapt in environmental reservoirs over decades.

3.2 Introduction

The *Ralstonia solanacearum* species complex (RSSC) comprises of plant-pathogenic bacteria, the causative agents of bacterial wilt disease, which have a wide host range and global distribution (Hayward 1991, Genin 2010, Safni et al. 2018, Bragard et al. 2019), making them one of the most devastating bacterial crop pests globally (Hayward 1991, Mansfield et al. 2012). The first recorded case of RSSC in the UK was within Oxfordshire in 1992 (Elphinstone and Matthews-Berry 2017) and has since been linked to nearby contaminated river water (Parkinson et al. 2013). UK RSSC belongs to the cold tolerant strain of *Ralstonia solanacearum* (race 3 biovar 2 or phylotype IIB sequevar 1), the causative agent of potato brown rot in Europe (Fegan and Prior 2005, Safni et al. 2014). Since 1992 there have been seven outbreaks of *R. solanacearum* in the UK; one in 1995 (infected potato), 1997 (infected tomato), 1998 (infected tomato), 1999 (infected potato), 2000 (infected potato), 2005 (infected potato) and 2010 (infected potato) (Elphinstone and Matthews-Berry 2017). All except the 2010 outbreak, which can be linked to infected imported seed potatoes, have been associated with contaminated water sources used for irrigation (Elphinstone and Matthews-Berry 2017). The prevalence of *R. solanacearum* in the UK rivers can be linked with an asymptomatic weed host, *Solanum dulcamara*, which resides along the riverbanks (Parkinson et al. 2013). The presence of infected *S. dulcamara,* otherwise known as woody nightshade, along rivers has been constantly correlated with *R. solanacearum* contaminated water sources (Wenneker et al. 1999, Bragard et al. 2019) and is thought to aid survival during the cold winter temperatures by acting as a reservoir for the pathogen (Genin and Boucher 2002, Champoiseau et al. 2009). As *R. solanacearum* is thought to have only recently been introduced to the UK, not long before 1992 (Parkinson et al. 2013, Elphinstone and Matthews-Berry 2017), it is an ideal model organism to investigate how adaptation and diversification of bacterium initially occurs within the environment. Bacterial organisms can diversify rapidly when new environmental niches open, however the underlying causes of diversification among pathogen populations are unclear (Rainey and Travisano 1998). Therefore, using *R. solanacearum* as the organism of interest will also increase understanding on how plant pathogens diversify within the natural environment across time,

which can improve the understanding of pathogen dispersal and spill over from environmental reservoirs as a cause of bacterial wilt disease outbreaks.

There are two key stages in the life cycle of *R. solanacearum*. As a pathogen the first stage is the infection of host plants, including production of virulence factors, evading plant immune responses and general survival within the host. The second stage is survival within external environments between transmissions, such as in soil and water (Genin 2010). In order to cope with this environmental variability, RSSC has evolved to be extremely diverse both genetically and phenotypically (Schell 2000, Cellier and Prior 2010, Genin and Denny 2012). However, only a few studies have been conducted focussing on European *R. solanacearum* diversity (van der Wolf et al. 1997, Timms-Wilson et al. 2001, Stevens and Van Elsas 2010, Cruz et al. 2012, Parkinson et al. 2013, Caruso et al. 2017). These studies mainly concluded that the cold tolerant European strains are genetically very clonal compared to other strains within the RSSC, consistent with a recent invasion of the temperate region disseminated through the international potato trade, and agreeing with the reports of *R. solanacearum* outbreaks in Europe, the first identified one being in Sweden in 1972 (Caruso et al. 2017). They also indicate that European isolates show phenotypic diversity, clustering into groups corresponding to pathogenicity (van der Wolf et al. 1997, Cellier and Prior 2010). However, the phenotypic traits studied so far are growth in specific carbon sources, previously used to differentiate isolates into biovars (Hayward 1991), or specific virulence traits. This means that focus on ecologically relevant traits, including those under high selection for bacterial survival within the external environment between infections, has been limited in these previous studies. This is a clear shortcoming as most European countries have reported *R. solanacearum* isolates surviving in rivers and wild host plants being the main cause of recurring outbreaks of the disease (van der Gaag et al. 2019).

One way to study diversity among this pathogen is to use high-throughput phenotyping (Blumenstein et al. 2015) and link this variation with underlying genetic differences. This research used these techniques to determine how diverse UK *R. solanacearum* population is, using a broad range of ecologically relevant traits driving pathogen adaptation, on a collection of 182 UK isolates. These isolates were

collected from both environmental sources (river water and wild host *Solanum dulcamara)* and crop hosts (potato and tomato), spanning almost 30 years of sampling, from the first recorded outbreak in 1992 to 2019 (see figure 3.1 for geographical location and time/host distribution and appendix table A.2 for the full list of isolates). The general diversity of a species can be constrained by correlations between different traits. Trait correlations can increase the evolvability of an organism, in the case of positive trait correlations, or limit adaptive potential, due to negative trait correlations (also known as trade-offs). Trade-offs can therefore prevent organisms from achieving maximum fitness due to allocation of finite resources (energy, time, molecules etc.) affecting investment in one trait over the other (Saltz et al. 2017) or due to genetic conflicts (multiple loci interacting to produce a single phenotype or a single locus affecting multiple different traits) (Foster et al. 2004). Therefore, phenotypic correlations along with niche preferences could play a role in the diversification of *R. solanacearum* within the environment.

The aims of this research were to first explore the phenotypic diversity of UK *R. solanacearum* (n=182), by collecting and analysing 46 ecologically relevant traits which cover three adaptive strategies*:* 1) metabolic capacity, 2) environmental stress tolerance, both abiotic and biotic, and 3) virulence (see appendix table C.1 for the full list of phenotypic traits collected). Then to determine the potential cause or causes of diversification among this pathogen population by linking trait variation to metadata, such as location, isolation source, and year of isolation. The final aim was to link this trait variation with specific genetic variants, using GWAS techniques (Lees et al. 2018), potentially identifying causative genes of adaptation within this *R. solanacearum* population. This research revealed that within the UK population of *R. solanacearum* three phenotypically distinct clusters, or 'ecotypes', exist differing in their ability to grow in nutrient limited environments, and in their trait specificity. Trait correlations also differed between these three ecotypes, implying that niche preferences and their trade-offs can be responsible for the range of phenotypes available to each ecotype, driving pathogen diversification. Investigating the potential causes of trait differences discovered that neither isolation source nor location can explain this diversity. However, the decade in which an isolate was

sampled in does explain small amounts of variation, indicating increased diversification over time. More specifically, traits involved in antibiotic resistance and biofilm formation was found to be increasing with time, suggesting a role of these phenotypes in UK *R. solanacearum* diversification. Finally, high genetic similarity between isolates resulted in a failure to link this phenotypic differentiation to specific genetic changes indicating that other factors, such as epigenetics, could impact UK *R. solanacearum* evolution. This study is the first to extensively focus on phenotypic diversity of UK *Ralstonia solanacearum* environmental isolates and therefore improves our understanding of how this bacterial pathogen may adapt initially within the environment.

3.3 Materials and Methods

3.3.1 *Ralstonia solanacearum* isolates and culture conditions

182 UK *R.* solanacearum isolates used in this study were collected and stored by FERA Science Ltd, Sand Hutton, York, UK. They were curated from a variety of environment (river water and *Solanum* dulcamara) and crop (potato and tomato) samples between 1992 to 2019 (appendix table A.2 and figure 3.1). All isolates were verified as belonging to the RSSC using the same 16S rRNA real-time PCR (RT-PCR) method as described in chapter 2 (EPPO 2018, Weller et al. 2000). Isolates were grown up on SP agar plates (Mehan and McDonald 1995), normalised to 0.1 OD_{600} (optical density at 600nm) using sterile deionised water and suspended in 25% glycerol, in random positions across different 96-well microplates using the same method as described above in chapter 2 (section 2.3.1). These microplates were cryopreserved at -80°C, ready for high-throughput phenotyping with the two previously described well characterised *R. solanacearum* species complex (RSSC) reference strains (*R. pseudosolanacearum* type strain GMI1000, and *R. solanacearum* type strain K60). A collection of 46 different traits were then collected for this *R. solanacearum* bacterial collection (see appendix table C.1 for full list of traits), categorised into four different groups: metabolic capacity, abiotic and biotic stress tolerance, and virulence. As mentioned in chapter 2, a linear model using batch, plate position and technical replicate as random effects was then

conducted, using the two reference strains, to determine that only small proportion of variation was explained by these variables (7%, 6% and 0% for batch, plate position and technical replicate respectively), and therefore corrections for batch effect were not used.

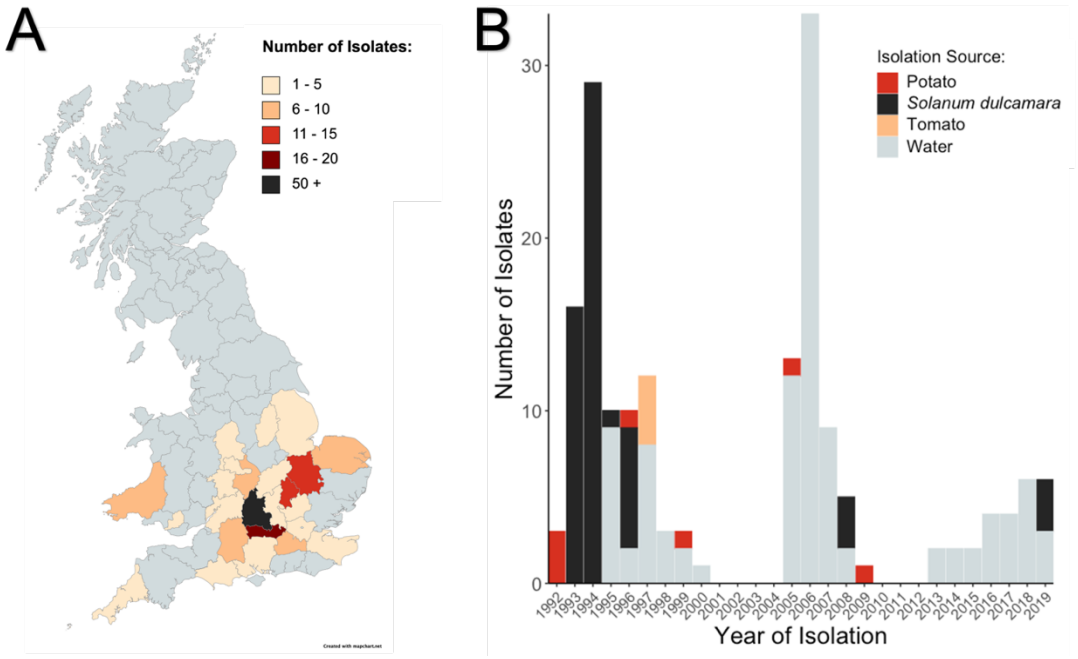


Figure 3.1: UK *Ralstonia solanacearum* isolate collection and sampling distribution. (A) Geographical schematic of the UK *R. solanacearum* collection. Colour indicates that at least one isolate originates from that county, with the shade of colour indicating the number of samples (map created using mapchart.net). (B) Distribution of the year (1992 to 2019) and host (water source, *Solanum dulcamara*, or crop host; potato or tomato) in which the UK collection was isolated from (n=182 in total).

3.3.2 Quantifying phenotypic traits linked with metabolic capacity

Metabolic traits included growth within four rich standard media commonly used for *R. solanacearum* growth in the laboratory (see 'Complex media' section in chapter 2), and 16 single carbon resources: asparagine, glutamine, histidine, proline, serine, glycine, glucose, arabinose, xylose, sucrose, maltose, sorbitol, nicotinamide, citric acid, malic acid, and succinic acid; all at 10mM (appendix table C.1, as described in the 'Single carbon resources' section of chapter 2).

3.3.2.1 Complex media

Traits involving growth in four separate complex media were measured using the same methods as described in section 2.3.2.1 (chapter 2).

3.3.2.2 Single carbon resources

Growth within 16 single carbon resources was also quantified using the same methods as stated in section 2.3.2.2 (chapter 2).

3.3.3 Phenotyping abiotic stress tolerance

Growth within abiotic stress conditions, including saline, extreme pH, high water potential and nutrient limited conditions, was conducted at 28°C within humidity. Relative growth within two different temperatures, 10°C and 35°C, were also quantified as an abiotic stress tolerance trait (appendix table C.1). SP media (Mehan and McDonald 1995) was used for all trait conditions using the same methods as described in chapter 2.

3.3.3.1 Growth in nutrient limited media

Growth within resource limited conditions, 1% and 10% SP media (Mehan and McDonald 1995), were measured as outlined in method section 2.3.3.1 (chapter 2).

3.3.3.2 Growth in saline media

Growth of all UK isolates was also measured in three different concentrations of salt (NaCl); 0.5%, 1%, and 2%, along with a control of 0% NaCl following the protocol described above in chapter 2 (section 2.3.3.2).

3.3.3.3 Growth in extreme pH

Growth within 4 different pH levels, ranging from acidic to alkali (4.5, 7, 9, 10), were also measured as stated in section 2.3.3.3 (chapter 2).

3.3.3.4 Growth in water potential stress conditions

To characterize how well UK *R. solanacearum* grows in water potential stress, relative growth measurements in 15% polyethylene glycol (PEG)-4000 within

SP media (Mehan and McDonald 1995) was characterized as reported above in section 2.3.3.4 (chapter 2).

3.3.3.5 Heat and cold tolerance traits

Growth of UK isolates in SP media (Mehan and McDonald 1995) was also characterised at 35°C and 10°C, following methods outlined in section 2.3.3.5 (chapter 2).

3.3.4 Phenotyping biotic stress tolerance

Antibiotic resistance was calculated as relative growth within four different antibiotics: tetracycline, gentamycin, rifampicin, and ciprofloxacin (appendix table C.1) at two different concentrations per antibiotic. See chapter 2 'Phenotyping biotic stress tolerance' (section 2.3.4) for full methods.

3.3.5 Phenotyping virulence traits

Phenotypes classed as virulence traits included biofilm production, oxidative stress tolerance, siderophore production and protein production (appendix table C.1).

3.3.5.1 Biofilm production

RSSC biofilms filter out nutrients from the flow of xylem fluid, as well as protect bacteria from host plant immune defences (Álvarez et al. 2010, Genin and Denny 2012, Meng 2013). Therefore, biofilm assays were conducted, using a crystal violet staining method, in the four different complex media, CS (Remenant et al. 2011), CPG (Kelman 1954), SP (Mehan and McDonald 1995) and nutrient broth (NB) (see the 'Complex media' methods section in chapter 2 for recipes). The biofilm assay was completed as dictated in section 2.3.5.1 (chapter 2).

3.3.5.2 Oxidative stress tolerance

Reactive oxygen species (ROS) are a defence system used by plants after infection of pathogens (Flores-Cruz and Allen 2009), and therefore tolerance to ROS

was calculated as a virulence trait for UK *R. solanacearum*. ROS tolerance was calculated using the same methods as dictated in chapter 2 (section 2.3.5.2).

3.3.5.3 Siderophore production

Iron chelating ability of bacterial supernatant was measured as a proxy for siderophore production using the (Schwyn and Neilands 1987) universal siderophore assay method (Arora and Verma 2017) as described above in section 2.3.5.3 (chapter 2). Siderophore production has been indicated as an important trait for rhizosphere community invasion and RSSC virulence (Gu et al. 2020). Siderophore production (psu) per cell was also calculated using the same method as stated in chapter 2.

3.3.5.4 Protein production

RSSC strains secrete a lot of virulence proteins while invading plant hosts (Genin and Denny 2012) and Bradford assays of each UK isolates supernatant was conducted to determine the amount of protein produced per isolate. Protein production per cell was measured using methods as stated previously in chapter 2 (section 2.3.5.4).

3.3.6 Whole genome sequencing and genetic variant curation

Whole genome sequencing was conducted on 168 isolates. DNA was first extracted following the Qiagen DNeasy Blood and Tissue Kit (see appendix table A.4 for full protocol). The elution buffer used was trisaminomethane hydrochloride (Tris HCl, 10mM, pH 8) as recommended by Earlham Institute for their library preparation. Quality of the extracted DNA was determined using 3 methods as previously described in chapter 2 (section 2.3.6). DNA was diluted to 15ng/μl and transported to Earlham for Illumina MiSeq 30x sequencing and all raw FASTQ files are publicly available at SRA under project number PRJNA823737.

3.3.7 Genetic analysis: linking phenotypes with genetic properties

Genome wide association studies (GWAS) were used to link phenotypic variation to changes within the *R. solanacearum* genome, as determined by differences in single nucleotide polymorphisms (SNPs) or small insertion and deletions (indels), the presence and absences in cluster of orthologous genes (COGs) and differences in DNA segments, known as unitigs. For identification of SNPs/indels, SNIPPY (Seemann 2015) was used to create alignment files (BAM files), then variant calling was conducted on these files using freebayes (Garrison and Marth 2012). To create the input COG files, an .Rtab file of gene presence/absence was created using panaroo (Tonkin-Hill et al. 2020). For investigation in unitigs, unitig-counter (v1.1.0) (Jaillard et al. 2018) was used. Bacterial GWAS was conducted using Pyseer v1.3.10 (Lees et al. 2018) which uses linear models with fixed or mixed effects to associate genetic variation with a variable of interest, while accounting for confounding population structure in the bacterial population. To account for population structure, all analyses were supplemented with phylogenetic distances from a tree constructed using IQ-TREE (Nguyen et al. 2015). A minimum minor allele frequency (maf) cut-off of 0.05 (and maximum maf of 0.95) was used for the unitig GWAS analysis. 0.10 and 0.90 maf cut-offs were used for the SNP GWAS analysis due to high errors in the SNP calling method due to the high clonality between isolates (Stoycheva et al. 2022: manuscript in progress). All Pyseer analyses were run using the linear mixed model (LMM). For annotations of significant unitigs identified in the GWAS, two reference strains representing the *R. solanacearum* phylotype IIB sequevar 1 strain was used. These were UY031 (NCBI accession number: 001299555) and YO199 (this study).

For phenotype GWAS, trait values (as measured as above for each trait) were taken and mean centred. All 46 traits were inputted as a continuous phenotype, however binary input phenotypes were calculated for all antibiotic resistant traits and reactive oxygen species tolerance, by taking growth (1) as OD_{600} over 0.13 at the last timepoint and no growth (0) as below this value. For ecotype level GWAS, the input phenotype used was ecotype classification based on k-means clustering and PC1 and 2 values from PCA plot in figure 3.2C. A temporal GWAS was

also conducted, where the input phenotype value was time of isolation in years (or decade) as a continuous trait.

The proportion of specific SNPs and COGs present per ecotype was also compared to see if there were overall differences between each phenotypic group. SNPs, from the vcf used for GWAS, were imported into R using vcfR (Knaus and Grünwald 2017) and downstream analysis was conducted on R version 4.2.1 (R Core Team 2022). SNPs that were present in less than 90% of isolates, and within more than one isolate per ecotype, were filtered for better visualisation. The percent of isolates in each ecotype that had a certain SNP was then calculated. For COGs, the same matrix as used in the GWAS analysis was taken and the proportion of genes shared by each ecotype was calculated using R version 4.2.1 (R Core Team 2022).

3.3.8 Statistical analysis and data visualisation

Statistical data analysis and visualisation was conducted within R version 4.2.1 (R Core Team 2022). Figures were produced using the package ggplot2 (Wickham 2016), heatmaps were curated using the Heatmap function in the ComplexHeatmap package (Gu et al. 2016), while visualisation of the correlation matrix was constructed using ggcorr within the GGally suite of packages (Schloerke et al. 2021), and venn diagrams using ggVennDiagram (Gao 2021).

Area under the growth curve (AUC) was calculated as a proxy for growth for most traits, as they capture lag time, growth rate, and carrying capacity, using the MESS package in R version 4.2.1 (Ekstrøm 2019, R Core Team 2022). Standardised trait values were z-scores curated using the scale function in the stats R package (R Core Team 2022) and mean standardised trait values across the ecologically relevant trait groups were taken as the mean z-score across all traits within that group (appendix table C.1). Principle component analysis (PCA) was carried out using tidymodels (Kuhn and Wickham 2020) with normalisation of the data. K-means clustering was conducted using the kmeans function in stats R package using Euclidean distances (R Core Team 2022) and optimal number of clusters were determined using the cascadeKM function in the vegan package (Oksanen et al. 2019), using calinski as the criterium.

Statistical analysis was also conducted in R version 4.2.1 (R Core Team 2022). Significance of clusters and groups within PCA plots was determined with PERMANOVA analysis using the adonis2 and betadisper function in the vegan package (Oksanen et al. 2019). Significant differences of mean trait values per phenotype group between each ecotype was determined using two-way ANOVA with the aov function, and post hoc using the TukeyHSD function, in the stats R package (R Core Team 2022), using the standardised trait value as the dependent variable and ecotype and phenotype group as the explanatory variables. Correlations between phenotype groups per ecotype was calculated using the rrcorr function in the Hmisc package (Harrell Jr 2022), with Pearson as the chosen correlation type. Significance of host and decade distribution across the three ecotypes was calculated using chi squared tests with the chisq.test function in the stats R package (R Core Team 2022), and post hoc analysis using the chisq.posthoc.test package (Ebbert 2019). Mean pairwise distances were calculated by taking the mean Euclidean distance per isolate from all other *R. solanacearum* isolates, using the dist() function (R Core Team 2022) on a matrix of the 46 phenotypic trait values for mean pairwise distance and the longitude and latitude for mean pairwise geographical distance. The median of the mean pairwise distance between different ecotypes and decades were then compared using Kruskal-Wallis significance test, using the kruskal.test function, and pairwise Wilcoxon test, using the pairwise.wilcoxon.test function, both available within the stats R package (R Core Team 2022), with Benjamini-Hochberg (BH) as the multiple test correction method. Statistical differences in antibiotic resistance and biofilm production trait values between each ecotype was determined using one-way ANOVA, with the aov function and post hoc analysis was conducted using the TukeyHSD function both in stats R package (R Core Team 2022), using the trait value as the dependent variable and ecotype as the explanatory variable. To determine correlations between antibiotic resistance and biofilm resistance, along with correlations between these traits and time of isolation, linear regression models were curated using the lm function in the stats R package (R Core Team 2022). Each isolate's trait value (or mean trait value across all isolates from that year) was used as the dependent

variable and time isolated from in years was the independent variable (or the other trait value for correlations between two traits).

3.4 Results

3.4.1 UK *R. solanacearum* population comprises of three distinct 'ecotypes' groups

Multiple independent trait measurements were collected across 182 UK *Ralstonia solanacearum* isolates to explore diversity of this pathogen within the UK. This dataset shows that the UK *R. solanacearum* isolates clusters into three groups, assigned by k-means clustering, based on their phenotypic diversity of 46 traits (figure 3.2A). The optimal number of clusters (k = 3) was determined using calinski criterion (appendix figure C.1B), and the three ecotypes can be seen clearly in the principal component analysis (PCA) (figure 3.2C). Isolates from each ecotype were found to be significantly different from one another phenotypically (PERMANOVA: $F_{2,179} = 528$, $R^2 = 0.85$, $p = 0.001$, pairwise PERMANOVA: p=0.01 for all), with significant lower dispersion within ecotype 2 compared to the other two ecotypes (ANOVA: $F_{2,179} = 11.3$, $p<0.0001$, TUKEY: p = 0.0003 for ecotypes 1 and 2, p = 0.002 for ecotypes 3 and 2, p=0.598 for ecotypes 3 and 1) suggesting lower diversity among isolates within this ecotype (ecotype 2). The PCA plot (figure 3.2C) explained 51% of the total variation within the dataset, with 32% of the variation explained by principal component 1 (PC1) and 19% by principal component 2 (PC2). A higher PC1 value corresponded to an increase in *R. solanacearum* growth across most traits, apart from nutrient limited conditions whereas a negative PC1 value corresponded to high growth in this condition. Furthermore, PC2 showed differences in growth across the 46 traits explored, with higher values showing higher growth on certain carbons and specific abiotic stresses, while negative values indicating higher biofilm production and antibiotic resistance (figure 3.2B). As this trait differentiation is likely to affect the fitness and ecology of *R. solanacearum* isolates, groups are henceforth referred to as 'ecotypes'. Trait diversity of each ecotype group also differed significantly from each other when calculated based on mean pairwise Euclidean distances (figure 3.2C, Kruskal-Wallis: $X^2_2 = 130$, $p<0.0001$), with ecotype 3

being more diverse than the other two ecotypes ($p<0.0001$ for both) and ecotype 2 being the least diverse out of the three ($p<0.0001$ for both ecotypes).

More detailed analysis revealed that isolates from ecotype 1 (n=62) could be considered as heat tolerant oligotrophs with significant higher trait values compared to the other two ecotypes in nutrient limited conditions (figure 3.2D, ANOVA: $F_{2,2327}= 335$, $p<0.0001$, TUKEY: $p<0.0001$ for both) and high temperatures (figure 3.2D, ANOVA: $F_{2,2327}= 335$, $p<0.0001$, TUKEY: $p=0.021$ and $p=0.012$ for ecotype 2 and 3 respectively) compared to the other two ecotypes. Ecotype 1 strains also had significant lower standardised values across most of the other traits, including protein production (figure 3.2D, ANOVA: $F_{2,2327}= 335$, $p<0.0001$, TUKEY: $p<0.0001$ and $p=0.0002$ for 2 and 3 respectively), siderophore production (figure 3.2D, ANOVA: $F_{2,2327}= 335$, $p<0.0001$, TUKEY: $p<0.0001$ for both), growth within complex media (figure 3.2D, ANOVA: $F_{2,2327}= 335$, $p<0.0001$, TUKEY: $p<0.0001$ for both) and single carbon media (figure 3.2D, ANOVA: $F_{2,2327}= 335$, $p<0.0001$, TUKEY: $p=0.0003$ and $p=0.038$ for 2 and 3 respectively), and tolerance to pH (figure 3.2D, ANOVA: $F_{2,2327}= 335$, $p<0.0001$, TUKEY: $p<0.0001$ for both) and salinity (figure 3.2D, ANOVA: $F_{2,2327}= 335$, $p<0.0001$, TUKEY: $p<0.0001$ for both) stress conditions. On the other hand ecotype 3 (n=16) had higher standardised trait values across most traits compared to the other two ecotypes, especially in siderophore production (figure 3.2D, ANOVA: $F_{2,2327}= 335$, $p<0.0001$, TUKEY: $p<0.0001$ for both), biofilm formation (figure 3.2D, ANOVA: $F_{2,2327}= 335$, $p<0.0001$, TUKEY: $p<0.0001$ for both), growth within complex media (figure 3.2D, ANOVA: $F_{2,2327}= 335$, $p<0.0001$, TUKEY: $p<0.0001$ for both) and relative growth across different antibiotics (figure 3.2D, ANOVA: $F_{2,2327}= 335$, $p<0.0001$, TUKEY: $p<0.0001$ for both). The largest cluster, ecotype 2 (n = 104), comprised of isolates with intermediate phenotypes, apart from for tolerance to reactive oxygen species (figure 3.2D, ANOVA: $F_{2,2327}= 335$, $p<0.0001$, TUKEY: $p<0.0001$ for both) and cold tolerance (figure 3.2D, ANOVA: $F_{2,2327}= 335$, $p<0.0001$, TUKEY: $p<0.0001$ for both) where isolates showed higher standardised value compared to the other two groups. Overall, these measurements show that the UK *R. solanacearum* population is phenotypically diverse, clustering into three distinct phenotypic ecotype groups.

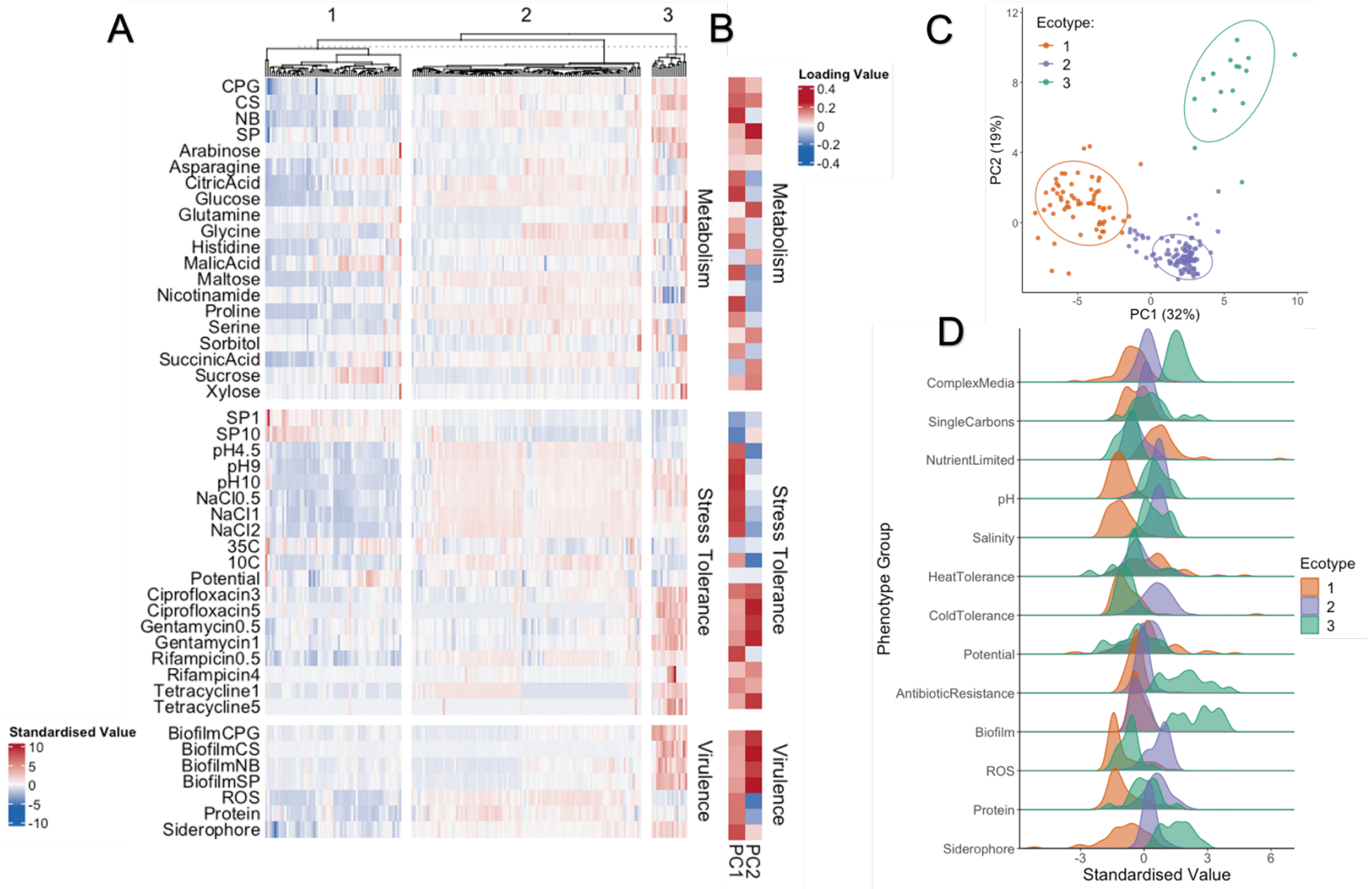

A
1
2
3
CPG
CS
NB
SP
Arabinose
Asparagine
CitricAcid
Glucose
Glutamine
Glycine
Histidine
MalicAcid
Maltose
Nicotinamide
Proline
Serine
Sorbitol
SuccinicAcid
Sucrose
Xylose
SP1
SP10
pH4.5
pH9
pH10
NaCl0.5
NaCl1
NaCl2
35C
10C
Potential
Ciprofloxacin3
Ciprofloxacin5
Gentamycin0.5
Gentamycin1
Rifampicin0.5
Rifampicin4
Tetracycline1
Tetracycline5
BiofilmCPG
BiofilmCS
BiofilmNB
BiofilmSP
ROS
Protein
Siderophore
Standardised Value
10
5
0
-5
-10
B
Loading Value
0.4
0.2
0
-0.2
-0.4
Metabolism
Stress Tolerance
Virulence
PC2
PC1
C
Ecotype:
1
2
3
PC2 (19%)
PC1 (32%)
D
Phenotype Group
ComplexMedia
SingleCarbons
NutrientLimited
pH
Salinity
HeatTolerance
ColdTolerance
Potential
AntibioticResistance
Biofilm
ROS
Protein
Siderophore
Standardised Value
Ecotype
1
2
3

Figure 3.2: UK *Ralstonia solanacearum* clusters into three groups based on phenotypic diversity. (A) Heatmap of z-score standardised trait values across the 47 traits shows trait differences across the 182 isolates. K-means clustering (k = 3) was used to split isolates into 3 groups and was confirmed using calinski criterion (appendix figure C.1). Traits can be split into three categories, metabolism, stress tolerance and virulence. (C) Principal component analysis (PCA) of the 182 isolates was used to visualize the phenotypic diversity of UK R. solanacearum by reducing dimensionality of the 46 traits into a 2D plot. The amount of contribution each trait has on the principal components is shown in a heatmap (B). Each point represents one UK isolate, coloured by ecotype assigned by k-means clustering (k=3). Ellipses show the 90% confidence interval around the centroid for each cluster. Principal component (PC) 1 explains 32% of the variation within the dataset and PC2 19%. PC3 graphs can be seen in appendix figure C.1C-E. (D) Standardised value (z-score standardised) distribution across each ecotype for traits grouped into 13 ecologically relevant categories. Colours indicate the assigned ecotype.

3.4.2 Trait correlations differ between ecotypes

Differences in trait correlations for each ecotype were explored to determine if they not only differed in their phenotypic profile but also regarding trait correlations. To achieve this, the measured 46 traits were grouped into 13 ecologically relevant phenotypic units (appendix table C.1) for easier interpretation using standardised values (same phenotype groups as figure 3.2D) and Pearson's correlation were conducted between all trait combinations (figure 3.3, for trait correlations across all 46 individual traits see appendix figure C.2). These results indicate trait correlations vary between the three ecotypes, with fewer strong significant trait correlations (above 0.4 or below -0.4) observed with ecotype 2 (9) compared to the other two ecotypes (14 and 10 for ecotype 1 and 3, respectively) (appendix table C.2). Furthermore, most correlations were positive, with ecotype 1 having three strong negative correlations and ecotypes 2 and 3 having one and two, respectively (appendix table C.2).

The three ecotypes showed strong trait correlations which were unique to each group. First, ecotype 1 had five unique strong trait correlations, three of which were positive and the other two negative. One of the positive correlations discovered within this group were between the two metabolic traits: growth in complex media and single carbon resources ($r=0.43$, $p=0.0004$). Other positive correlations found were between tolerance towards ROS and pH ($r=0.43$, $p=0.0005$), and high and low temperatures ($r=0.41$, $p=0.0011$). The two strong negative correlations found unique to ecotype 1 were also linked to metabolic traits, with growth within single carbon resources being negatively correlated with growth within nutrient limited conditions ($r=-0.42$, $p=0.0006$) and growth within complex media negatively correlated with tolerance to cold temperatures ($r=-0.50$, $p<0.0001$). This suggests that trait correlations between metabolic traits could be driving ecotype 1 phenotypic diversity.

Secondly, ecotype 2 had four unique positive correlations. Two of which were between virulence traits, with biofilm production being positively associated with growth within single carbon resources ($r=0.59$, $p<0.0001$) and tolerance to reactive oxygen species (ROS) positively associated with high-water potential tolerance ($r=0.44$, $p<0.0001$). Stress tolerance traits were also positively associated

with one another within ecotype 2, with tolerance to pH and cold temperatures being positively correlated ($r=0.45$, $p<0.0001$), in addition to salinity tolerance and antibiotic resistance ($r=0.46$, $p<0.0001$).

Finally, ecotype three showed four unique strong trait correlations, two of which were positive and the other two negative. Salinity tolerance was found positively associated with growth within nutrient limited conditions (0.71, $p=0.0023$), as well as protein production being positively correlated with growth within complex media ($r=0.56$, $p=0.0246$). Negative trait correlations within this ecotype were also discovered between cold tolerance and two other traits, resistance towards antibiotics ($r=-0.52$, $p=0.0384$), and growth in nutrient limited conditions ($r=-0.55$, $p=0.0285$).

Some positive correlations were shared by all three different ecotypes, including correlation between pH and salinity tolerance traits ($r=0.44$, $p=0.0003$ for ecotype 1, $r=0.74$, $p<0.0001$ for ecotype 2 and $r=0.80$, $p=0.0002$ for ecotype 3), as well as between tolerating cold temperatures and reactive oxygen species (ROS) ($r=0.67$, $p<0.0001$ for ecotype 1, $r=0.72$, $p<0.0001$ for ecotype 2 and $r=0.62$, $p=0.0111$ for ecotype 3). Furthermore, multiple strong correlations were shared between two of the three ecotypes. Ecotypes 1 and 2 shared three trait correlations, with strong positive correlations discovered between salinity tolerance and protein production ($r=0.51$, $p<0.0001$ and $r=0.46$, $p<0.0001$ for ecotypes 1 and 2 respectively), and between antibiotic resistance and growth within complex media ($r=0.44$, $p=0.0003$ and $r=0.45$, $p<0.0001$ for ecotypes 1 and 2 respectively). They also both shared a strong negative correlation between growth in complex media and nutrient limited conditions ($r=-0.68$, $p<0.0001$ and $r=-0.69$, $p<0.0001$ for ecotype 1 and 2 respectively). Furthermore, ecotype 1 shared four positive trait correlations with ecotype 3. Three of which were between heat tolerance and other stress tolerance traits, including growth within nutrient limited conditions ($r=0.55$, $p<0.001$ and $r=0.55$, $p=0.027$ for ecotype 1 and 3 respectively), tolerance to salinity stress ($r=0.54$, $p<0.0001$ and $r=0.56$, $p=0.0223$ for ecotypes 1 and 3 respectively) as well as extreme pH stress ($r=0.43$, $p=0.0007$ and 0.68, $p=0.0040$), suggesting that tolerating high temperatures and other stress conditions are closely linked within these two ecotypes. Another shared positive trait correlation was found between

pH tolerance and growth within nutrient limited conditions (r=0.48, p=0.0001 and r=0.57, p=0.0215 for ecotypes 1 and 3 respectively). Interestingly, ecotypes 2 and 3 did not have any unique strong trait correlations shared between them, suggesting that these two ecotypes have very different trait correlation patterns.

Together, these results show that we find both negative and positive strong trait correlations, some of which are shared between the three ecotypes and the majority of which are unique to each ecotype group. Trait correlations could therefore be driving phenotypic differences between UK *Ralstonia solanacearum* isolates and play a potential role in maintaining these ecotypes within the UK-wide population.

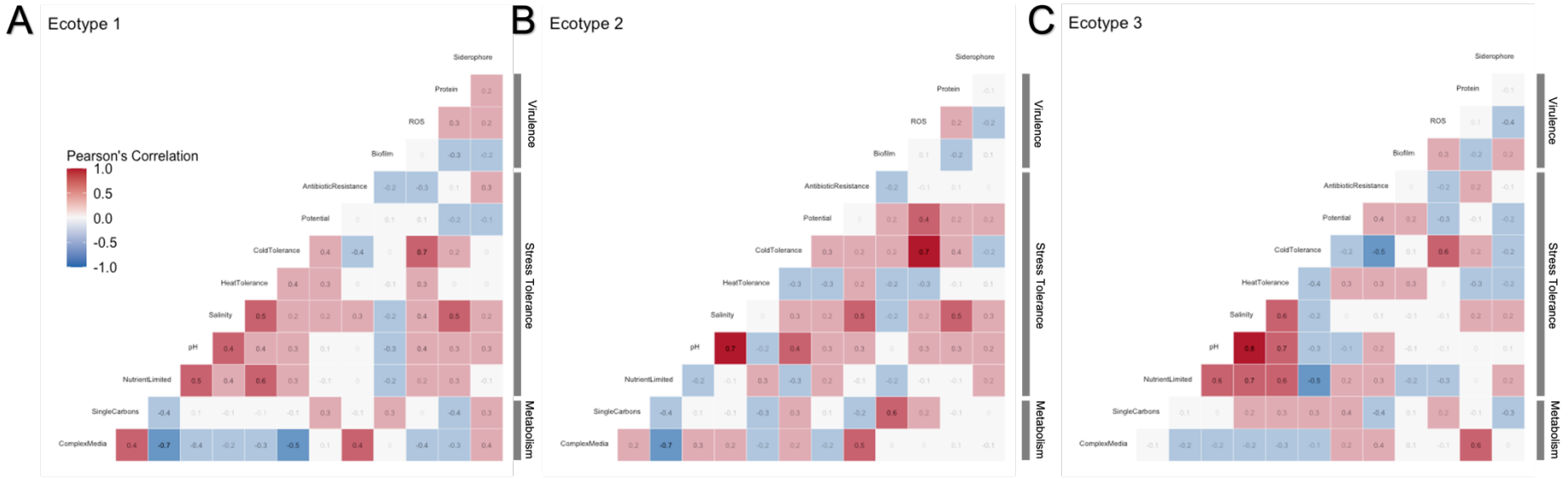


Figure 3.3: Trait correlations differ among the three ecotypes. Pearson's paired correlations of the 46 traits classified into 13 phenotypic groups for easier interpretation. See appendix figure C.2 for correlations among all traits. This analysis was done for each ecotype, assigned by k-means clustering, with 62 isolates in ecotype 1 (A), 104 isolates in 2 (B) and 16 in ecotype 3 (C). Red indicates a positive trait correlation, blue a negative trait correlation and white no correlation. Traits are also ordered by phenotype classification, metabolism, stress tolerance or virulence.

3.4.3 The decade of isolation explains phenotypic diversity among UK isolates, while the isolation source or geographical location do not

Next, exploration on isolate metadata (date, isolation source or location) was conducted to attempt to explain the phenotypic separation observed among UK isolates. To investigate this, the same PCA plot as in figure 3.2C was constructed but with isolates grouped by the source they were isolated from, river water, weed host *Solanum dulcamara* or crops (potato or tomato). This showed that no significant differences in trait diversity was driven by isolation source (figure 3.4A, PERMANOVA: $F_{2,179} = 1.9$, $R^2 = 0.02$, $p = 0.11$). Furthermore, host distribution was found to be non-significantly associated with certain ecotypes (appendix figure C.3A, $X^2_4 = 8.4$, $p = 0.08$), suggesting that host samples were not significantly biased across the three ecotypes.

Isolates were also grouped by the decade in which they were isolated from. Decade was used for easier interpretation and due to the absence of isolates sampled in certain years (figure 3.1B). This showed that significant, although small, differences in phenotypic diversity can be explained by sampling date (figure 3.4B, PERMANOVA: $F_{2,177} = 8.8$, $R^2 = 0.09$, $p = 0.001$). Pairwise PERMANOVA revealed that this was driven by differences between isolates from the 1990s and 2010s ($p = 0.015$) and 2000s and 2010s ($p = 0.015$), suggesting that isolates from the most recent decade (2010s) are more diverse compared to the other two decades (figure 3.4B). This was confirmed by dispersion around the centroid being significantly higher in the 2010s compared to the other two decades (figure 3.4B, ANOVA: $F_{2,177} = 11.8$, $p < 0.0001$, TUKEY: $p < 0.0001$ for both 2000s and 1990s compared to 2010s). Furthermore, mean pairwise distances (compared to all other isolates) agreed with these results, as isolates originating from the 2010s had significantly higher mean pairwise differences compared to the other two decades (appendix figure C.3C, Kruskal Wallis: $X^2_2 = 19.0$, $p<0.0001$, Post Hoc: $p=0.0002$ between 2010s and 2000s, $p=0.0002$ for 2010s and 1990s, $p=0.17$ for 1990s and 2000s). Additionally, isolates from each decade were not evenly distributed across the three ecotypes (appendix figure C.3B, $X^2_4 = 45.0$, $p<0.0001$), with isolates from the 2010s being significantly less likely assigned to ecotype 2 ($p=0.0013$), and significantly more likely assigned to ecotype 3 ($p<0.0001$). Isolates from the 1990s

were also significantly less likely to be associated with ecotype 3 (p=0.013), and no significant differences were found between ecotypes for isolates from the 2000s (p=1.0 for all). This suggest that the decade in which the pathogen was isolated in can partly explain the trait differences between isolates. However, this difference was small and can mainly be explained by isolates from 2010 onwards being more diverse compared to the isolates sampled from the two other decades.

Finally, isolates were plotted by their coordinates of isolation and coloured by the assigned ecotype, but no significant pattern between ecotype and location was found (figure 3.4C, PERMANOVA: $F_{2,171} = 2.4$, $R^2 = 0.03$, $p = 0.07$). Each isolate's pairwise geographical distance from all other isolates (Euclidean distance from the coordinates they were isolated from) was also calculated and compared. This revealed no significant difference in geographical distance from the other isolates among the three ecotypes (appendix figure C.3D, Kruskal Wallis: $X^2_2 = 4.1$, p=0.13). This suggests that neither location nor isolation source can explain the phenotypic separation of UK *R. solanacearum* into the three ecotype groups. However, the time of isolation can partly explain some of these phenotypic differences, suggesting increased diversification over the most recent decade.

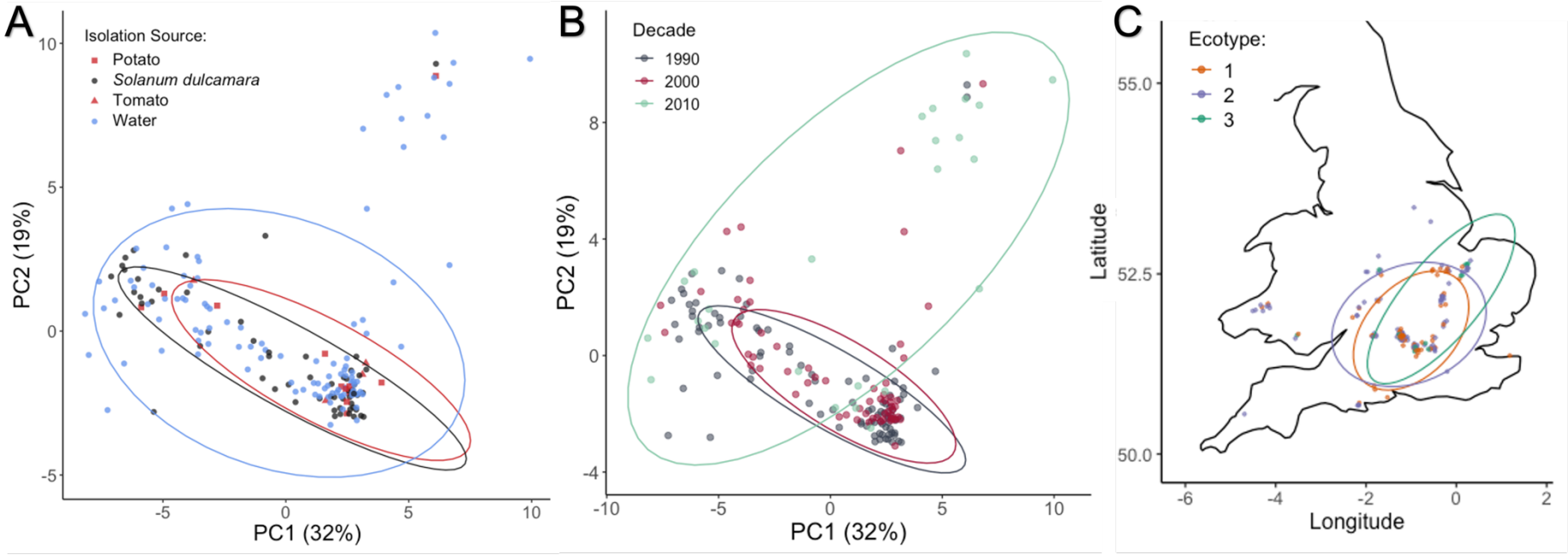


Figure 3.4: Metadata cannot explain clustering of UK *R. solanacearum.* (A) Principal component analysis (PCA) with each point representing an isolate and coloured by the host they were isolated from. This includes isolation from water sources (blue), weed host *Solanum dulcamara* (grey), and crop hosts potato and tomato (red) (differences between potato or tomato are shown by shape). The eclipses highlight the 90% confidence interval around each host group; river water, weed host *Solanum dulcamara* and crops (n=182). (B) PCA with each point representing a UK *R. solanacearum* isolate and coloured by the decade in which they were isolated from. Eclipses shows 90% confidence interval around the centroid of each decade (1990s, 2000s and 2010s) (n=180). (C) Map of location each isolate was collected from across the UK. Each points represents one isolate, plotted by the coordinates of their place of isolation, with jitter to make the points more visible. Isolates are coloured by ecotype assigned by k-means clustering (k = 3). Ellipse shows the 90% confidence interval around the centroid of each ecotype (n=174).

3.4.4 Linking genetic variation with trait differences highlights high genetic similarity within the UK *R. solanacearum* population

To link trait and genetic variation, a genome wide association study (GWAS) was conducted on 168 of the UK *R. solanacearum* isolates using single nucleotide polymorphisms (SNPs), cluster of orthologous genes presence and absences (COGs) and DNA segments known as unitigs as input genetic variables. However, the strong effect of population structure, combined with the little genetic variation discovered among the UK *R. solanacearum* population, confounded results and few promising associations were discovered. After filtering for minor allele frequency (maf), only 734 variable unitigs, 7 SNPs, and 48 variable COGs were found in the UK population. Within the unitig GWAS, only 20 traits had a significant gene associated with them, most of which were antibiotic resistance traits and biofilm production traits within different media (see appendix table C.3). Furthermore, most gene hits within the unitig GWAS were single unitigs, suggesting a high likelihood of them being false positives. However, a tyrocidine synthase gene (*tycC*) was associated with many different traits (including resistance to rifampicin, tetracycline, and gentamycin antibiotics, biofilm and siderophore production, and growth within CS and nicotinamide media), with up to 95 significant unitigs assigned to it for a single trait (appendix table C.3). Further analysis revealed that multiple different annotations were assigned to this gene, suggesting a high likelihood of this gene being a hypothetical protein. Overall, the genetic variation we captured within this study cannot explain most trait differences, supporting genomic analysis, which indicates that the UK *R. solanacearum* population is clonal (Stoycheva et al. 2022: manuscript in progress).

3.4.5 Antibiotic resistance and biofilm production traits are driving UK *R. solanacearum* diversification in time

The few promising associations between a hypothetical protein and antibiotic resistance and biofilm formation were explored in more detail. By using the mean standardised trait value over all antibiotic resistance traits (n=8) and biofilm production (n=4) traits, it was found that both trait values differ significantly between ecotypes (appendix figure C.4A, ANOVA: $F_{2,179}$= 207, p<0.0001 and

appendix figure C.4B, ANOVA: $F_{2,179}$= 371, p<0.0001 for antibiotic resistance, and biofilm production respectively). Specifically, ecotype 3 isolates showed higher trait values compared to ecotype 1 (TUKEY: p<0.0001 for both traits) and 2 (TUKEY: p<0.0001 for both traits) regarding both trait mean averages. In contrast, ecotypes 1 and 2 did not significantly differ in their ability to produce biofilm (TUKEY: p=0.72), but differed in their ability to resist antibiotics, with ecotype 2 being more resistant than ecotype 1 on average (TUKEY: p<0.0001).

It was also found that both antibiotic resistance and biofilm production are increasing with time. With isolates collected more recently having higher trait values compared to isolates collected at the time of the first reported case of *R. solanacearum* in the UK (appendix figure C.4C, $F_{1,178}$= 7.8, adjusted R^2= 0.04, p=0.006 and appendix figure C.4D, $F_{1,178}$= 30, adjusted R^2= 0.14, p<0.0001 for antibiotic resistance and biofilm production respectively). Fitting a linear model over the average trait value per year revealed a significant increase in biofilm production in all four media in time (figure 3.5A: $F_{1,19}$= 15, adjusted R^2= 0.40, p=0.001 and figure 3.5A: $F_{1,19}$= 13, adjusted R^2= 0.38, p=0.001 for CPG and CS respectively and figure 3.5B: $F_{1,19}$= 10, adjusted R^2= 0.32, p=0.004 and figure 3.5B: $F_{1,19}$= 10, adjusted R^2= 0.32, p=0.005 NB and SP media respectively). Similar increase across time was also observed with ciprofloxacin resistance (figure 3.5C: $F_{1,19}$= 4.9, adjusted R^2= 0.16, p=0.039 and figure 3.5C: $F_{1,19}$= 6.1, adjusted R^2= 0.20, p=0.023 for 3μg/ml and 5μg/ml respectively), gentamycin resistance (figure 3.5D: $F_{1,19}$= 9.5, adjusted R^2= 0.3, p=0.006 and figure 3.5D: $F_{1,19}$= 7.3, adjusted R^2= 0.24, p=0.01 for 0.5μg/ml and 1μg/ml respectively), and tetracycline resistance at 5μg/ml (figure 3.5E: $F_{1,19}$= 8.9, adjusted R^2= 0.28, p=0.008). Together, this suggests that ecotype 3 isolates are associated with a higher antibiotic resistance and biofilm production, traits that are overrepresented by isolates collected from the 2010s (appendix figure C.3B). Furthermore, while trait correlations per ecotype (figure 3.3) did not show a positive correlation between overall biofilm production and antibiotic resistance to all measured antibiotics, trait correlations across all three ecotypes were positively correlated with each another (appendix figure C.4E: $F_{1,180}$= 175, adjusted R^2= 0.49, p<0.0001). Selection for antibiotic resistance and increased

production of biofilm could be hence partly driving the UK *R. solanacearum* population diversification into sperate ecotypes.

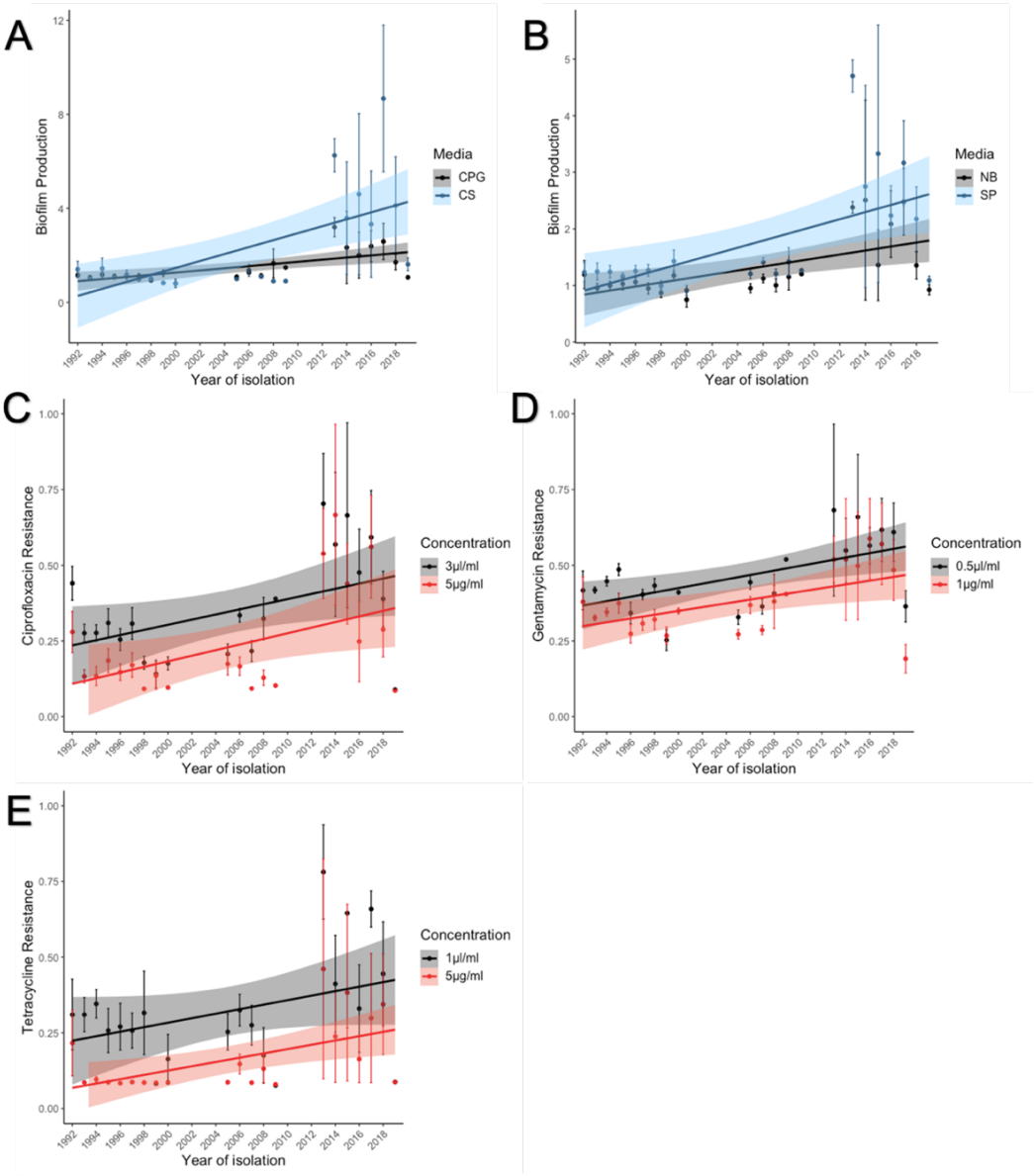


Figure 3.5: Biofilm production and ciprofloxacin, gentamycin, and tetracycline resistance increases across time within the UK *R. solanacearum* population. Mean average biofilm production in CPG and CS media (A) and NB and SP media (B) across all UK isolates each year. Antibiotic resistance was measured as relative growth within the presence of the antibiotic (described in chapter 2.3.4). Mean average ciprofloxacin resistance, at 3µg/ml and 5µg/ml, (C) gentamycin resistance, at 0.5µg/ml and 1µg/ml, (D) and tetracycline resistance, at 1µg/ml and 5µg/ml, (E) across all UK isolates within that year. Error bars show 1 standard error around the mean and coloured line shows the linear model with shaded regions showing 95% confidence intervals around the model for each concentration/media. All linear models but tetracycline resistance at 1µg/ml (black in panel E: $F_{1,19}$=3.0, adjusted R^2= 0.09, p=0.1) show significant positive relationships between resistance/biofilm production and time.

3.4.6 Temporal GWAS (tGWAS) identifies gene associated with time

As the time of isolation appeared to be related to phenotypic diversity within UK *R. solanacearum* (figure 3.4B), a temporal GWAS was also conducted using time in years, as well as the decade in which they were isolated from, as the phenotypic variable. This revealed one SNP, four COGs and multiple unitigs associated with time (appendix table C.5). The SNP hit was identified within a putative deoxyribonuclease RhsC (RSUY_02640) gene. Furthermore, four genes (group_2086, group_1877, group_1933 and group_1905) were also found to be associated with time of isolation. Unfortunately, they are all hypothetical proteins and therefore their functions are unknown. As well as these genes, multiple unitigs (172 for year, and 155 for decade, as the continuous trait) were also found to be associated with time (appendix figure C.6). The most promising unitig hit found associated with time, is the same tyrocidine synthase gene (*tycC*) found associated with multiple traits within the phenotype GWAS. These traits included a few antibiotic resistant and biofilm production traits which were found to be increasing with time (figure 3.5). This tyrocidine synthase gene had 87 or 77 (year and decade as phenotype respectively) unitigs significantly associated with time for homolog 3 (tycC_3) and 66 or 61 (year and decade as phenotype respectively) unitigs associated with homolog 2 (tycC_2), making it highly unlikely that this is a false positive association. However, like previously stated multiple different annotations were found to be assigned to this gene suggesting a high likelihood of it being a hypothetical protein. Overall, temporal GWAS techniques has identified a hypothetical gene linked with time that could be a novel contributor to *R. solanacearum* pathogen survival within the UK.

3.4.7 Genetic variation explains only small phenotypic ecotype differences

Genetic variation (SNPs and COGs) among the UK *R. solanacearum* population was also investigated and compared between the three different ecotypes. The proportion of each single nucleotide polymorphisms (SNPs) identified within each ecotype is summarized in figure 3.6A and B. This highlights that there are fewer SNPs identified among ecotype 3, which could be due to sampling size (n=16 compared to n=62 and n=104 in ecotype 1 and 2 respectively) rather than

actual variation between the three ecotypes. SNPs are spread evenly across the chromosome and megaplasmid (figure 3.6B) and most SNPs are at a low frequency among each ecotype (figure 3.6A and B), suggesting neutral adaptation. The presence and absence of clusters of orthologous genes (COGs) can capture genes inherited through horizontal gene transfer, as well as vertical transmission, which may play an important role in bacterial adaptation to a new niche. However, gene presence and absence information among UK *R. solanacearum* shows a similar pattern than the SNP genetic variation, with little variation seen among these pathogen isolates. Overall, 98% of genes are shared between all three ecotypes, suggesting high genetic similarity between the three groups (figure 3.6C). While very few genes were found to be unique to specific ecotype groups, with ecotype 1 having the most unique genes (40), ecotype 2 having only 8 unique genes and ecotype 3 having no unique genes (figure 3.6C). Overall, this suggests that the genomic variation we have captured cannot explain the large trait differences between the three UK *R. solanacearum* ecotypes identified.

To further study if overall genetic variation between the UK *R. solanacearum* isolates could explain the phenotypic ecotype differences seen, another GWAS was conducted using ecotype, along with principal component (PC) 1 and PC2 values (from figure 3.2C), as the phenotypic variables. This analysis revealed one significant SNP and a couple of unitigs associated with ecotype and PC2 (appendix figure C.5 and appendix table C.4). These unitigs were all unique with only one hit suggesting a high possibility of false positives. The significant SNP identified as being associated with ecotype, was identified using both ecotype and PC2 as phenotypic variables and was annotated as being within a putative deoxyribonuclease *RhsC* gene (RSUY_02640). This SNP can be found at a higher proportion within isolates assigned to ecotype 3 compared to the other two ecotypes (fourth gene from the bottom in figure 3.6A) and may be associated with phenotypic differences between ecotype 3 with the other two ecotypes. Furthermore, this is the same SNP previously found associated with time and suggests that the year of isolation could be the driving force of diversification of UK *R. solanacearum* into three ecotypes. Together, these results highlights a lack of genetic variation observed within the UK *R. solanacearum* population.

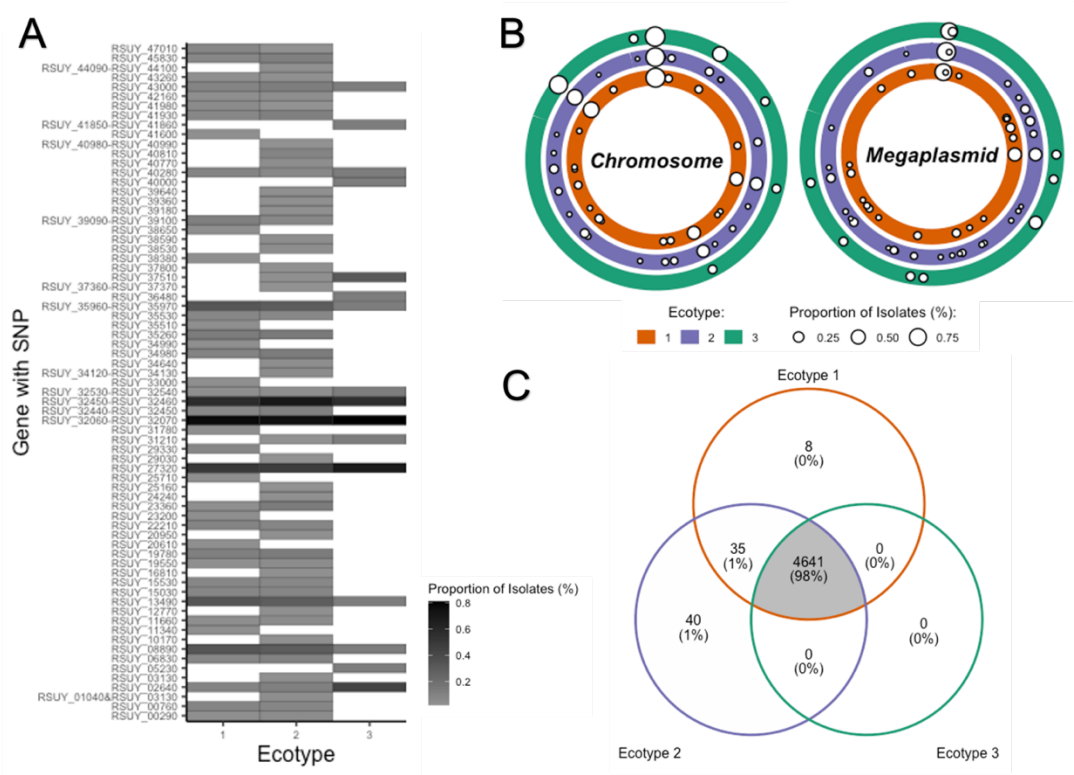


Figure 3.6: Genetic variation captured (SNPs and gene presence/absence) cannot explain ecotype differences. (A) Proportion of isolates (%) in each ecotype that have a SNP within each gene. (B) SNPs found in the chromosome (left, ~3.5Mbp) and megaplasmid (right, ~2Mbp) per ecotype (coloured). Each circle represents a gene in which a SNP has been called in, the position of the circle shows the that genes position within the *R. solanacearum* genome, size of circles indicates the proportion (%) of isolates per ecotype that have that variant present within its genome. Circle highlights the different ecotype group, inner circle being ecotype 1, then ecotype 2 and outer circle showing ecotype 3. (C) Venn diagram showing the number (and percentage) of accessory genes shared across the three ecotypes or are unique to each ecotype.

3.5 Discussion

The aims of this research were to explore the phenotypic and genetic variation of *R. solanacearum* in space and time within the UK. First, by exploring the phenotypic diversity of 46 ecologically meaningful traits (involved in metabolism, stress tolerance and virulence) across 182 UK isolates significant trait variation was observed despite the previously reported clonal population structure and relatively recent introduction to the UK (Elphinstone and Matthews-Berry 2017). From this data, three phenotypically distinct ecotypes were identified, each with specific characteristics (figure 3.7).

Ecotype 1 (n = 62) could be considered a heat-tolerant oligotrophic specialist with lower standardised values compared to the other two ecotypes in most traits apart from low concentrations (1% and 10%) of sucrose peptone media, where it had the highest trait values. Within environmental reservoirs, such as rivers, *R. solanacearum* must cope with the stress of nutrient limited habitats, making this a potentially important trait for surviving and persisting in the environment (Álvarez et al. 2008). Ecotype 2 could be considered as a generalist ecotype (figure 3.7) with intermediate standardised values for all traits, apart from high trait values in tolerance to reactive oxygen species (ROS) and cold temperatures (10°C) (figure 3.2D). The ability to grow, and therefore infect plants, at cold temperatures (10°C) will be extremely important for *R. solanacearum* residing in temperate regions of the world, such as the UK. Average annual temperatures in Oxford, the region where the first outbreak of UK *R. solanacearum* was reported, were recorded to be 15°C from 1991 to 2020 (Met Office 2022), suggesting that in order to survive and infect plants within this environment, tolerating cold temperatures would be beneficial. Therefore, it is unsurprising that this cold tolerant ecotype 2 was the largest ecotype identified in the UK, consisting of 104 isolates. ROS tolerance was also highest within this ecotype. ROS is produced by plants as a defence response to pathogen infection (Flores-Cruz and Allen 2009) and therefore is an important virulence trait for *R. solanacearum*. Ecotype 3 was a specialist pathogen group, with higher standardised values in traits such as growth within complex media, antibiotic resistance, biofilm production and siderophore production (figure 3.7). This ecotype was also the most diverse group out of the three but also the smallest comprising of

only 16 isolates (figure 3.2). Siderophore and biofilm production are both important virulence traits for *R. solanacearum*. Biofilms have been suggested to filter out nutrients from the flow of xylem fluid, as well as protect *R. solanacearum* from host plant immune defences (Álvarez et al. 2010, Genin and Denny 2012, Meng 2013), while siderophores can be used to increase resource utilisation and have been linked to pathogen virulence (Bhatt and Denny 2004, Gu et al. 2020). Antibiotic resistance is an important trait for persisting within microbial communities, where antagonistic bacteria can produce a variety of antibiotics that kill RSSC species to reduce competition (Allen et al. 2010, Yuliar et al. 2015). Also, all three ecotypes differed in their ability to utilise rich 'complex' media, with ecotype 3 having highest growth measurements, ecotype 2 having intermediate values, and ecotype 1 the lowest trait values (figure 3.7). This could be an indication of niche separation between the three ecotypes to avoid competition (Bajic and Sanchez 2020). Overall, this separation of UK *R. solanacearum* into three separate ecotypes, each differing in certain traits (figure 3.7), could help pathogen survival within the environment by each group adopting different adaptive strategies.

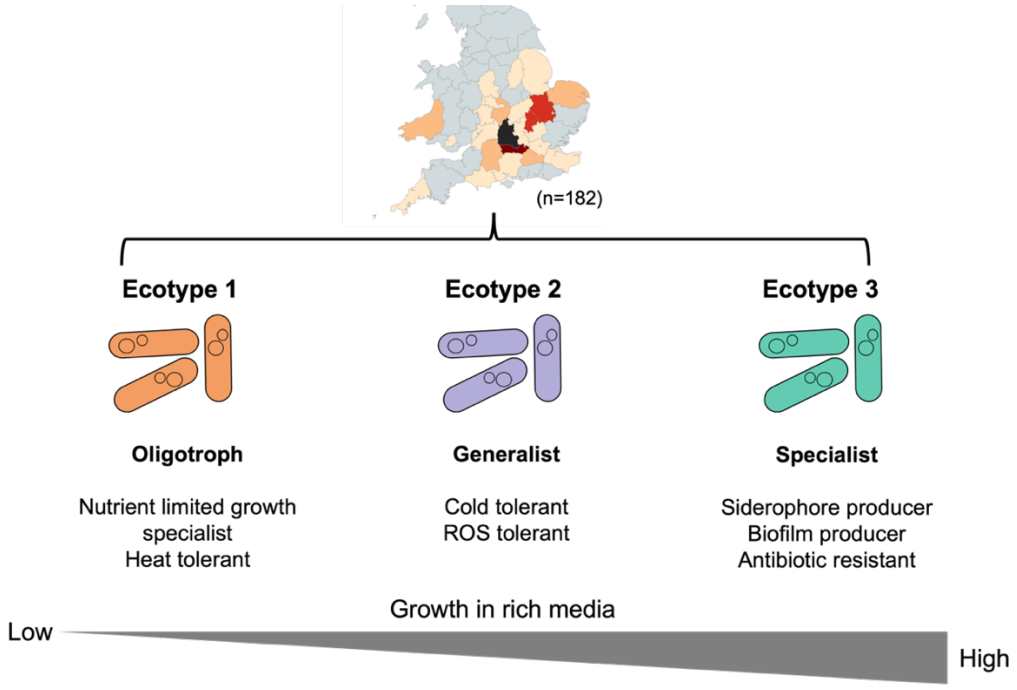


Figure 3.7: Summary schematic of three phenotypically distinct 'ecotypes' identified in this study. From a collection of 182 UK *R. solanacearum* isolates three phenotypically distinct ecotypes have been identified. Ecotype 1, the oligotroph group, has higher trait values in nutrient limited conditions and higher temperatures. Ecotype 2, the generalist group, has higher trait values in cold temperatures and when in the presence of reactive oxygen species (ROS). Ecotype 3, the specialist

group, has higher trait values in producing siderophores and biofilm, as well as being the most antibiotic resistant strains. Ecotype 3 also has highest growth in rich 'complex' media, with ecotype 1 having the lowest growth and ecotype 2 having intermediate values.

Not only do trait values differ between ecotypes, but so do their trait correlations. Exploring trait correlations across the three identified ecotypes revealed that some trait correlations were unique to each ecotype, suggesting that they could help maintain ecotype separation. Negative trait correlations can limit the range of phenotypes open to organisms (Ferenci 2016), while positive traits correlations can increase the evolvability of an organism (Saltz et al. 2017). There are numerous known negative trait correlations (trade-offs) among *Ralstonia solanacearum* species complex (RSSC) strains (Peyraud et al. 2016, Wang et al. 2017, 2019, Khokhani et al. 2017), however a more comprehensive understanding on trait correlations, spanning a wider range of traits, provides a deeper understanding on their role in the phenotypic diversification within this pathogen species. This research illustrates that trait correlations can differ among different groups (ecotypes) within the same pathogen population, highlighting that trait correlations are perhaps more dynamic than first thought within *R. solanacearum.* These differences within trait correlations between the three ecotypes can therefore alter the range of phenotypes open to these isolates, potentially driving them into different adaptive fitness landscapes.

However, other trait correlations were found to be shared by all three ecotypes. Many of the trait correlations observed were also positive, highlighting *R. solanacearum's* generalist nature (Hayward 1991, Genin and Boucher 2002). Two trait correlations were found across all three ecotypes, these were positive correlations between pH and salinity tolerance and tolerance to ROS and cold temperatures. This could indicate that the two traits are genetically linked or environmentally linked, where these two traits commonly occur together within the environment. Salinity has been seen to coincide with high pH stress within the soil and rivers (Sardinha et al. 2003, Jiang et al. 2022) with sewage and agricultural run-off causing increased pH and saline levels within these environments (Rowbury 1997, Shrivastava and Kumar 2015, Zhang et al. 2021). Alkaline pH has also been seen to be increasing within UK rivers through time, from 2000 to 2020 (Jiang et al.

2022). Therefore, tolerating both extreme pH and saline conditions could be linked to environmental survival in the UK environment. ROS and cold tolerance were also found positively associated among all three ecotypes. This may be an adaptation of *R. solanacearum* to infect plants within the colder UK temperatures, as both growth within colder temperatures and tolerance to ROS, produced by plants as a defence mechanism (Flores-Cruz and Allen 2009), will be required for successful infection.

To determine the cause of separation of this UK population into three separate ecotypes, phenotypic variation was correlated with metadata, including the isolation source, year, and location in which isolates were sampled from. The source from which each isolate was taken from, crop host (potato or tomato), weed host (*Solanum dulcamara*), or river water, did not explain any significant phenotypic variation among the UK isolates (figure 3.4A). This research found that UK isolates did not differ significantly based on their isolation source, contradicting another study conducted in the Netherlands which found that river isolates lost their ability to utilise seven resources compared to strains from plant hosts (Stevens and Van Elsas 2010). In the Netherlands, Sweden and UK waterways, *Solanum dulcamara* acts as a reservoir for *R. solanacearum* helping them to survive the cold winter temperatures before they return to the river (Elphinstone et al. 1997, Parkinson et al. 2013, van der Gaag et al. 2019). Therefore, it could be expected that isolates from river water and *S. dulcamara* nearby are from the same population and therefore have very similar phenotypic diversity. Environmental (river water and weed hosts) and crop host environments are also closely linked as river water is used as an irrigation source for crops, causing outbreaks of *R. solanacearum* in the UK (Parkinson et al. 2013). Furthermore, there have also only been eight reported outbreaks in crop plants to date in the UK (Elphinstone and Matthews-Berry 2017), and therefore there could be not enough samples to detect phenotypic differences between them and environmental samples.

Location also did not explain UK ecotype variation, with the three distinct ecotypes identified in this study spread evenly across the UK (figure 3.4C). The UK population of *R. solanacearum* is thought to have been introduced relatively recently, with the first reported outbreak on potato being in 1992 (Parkinson et al. 2013, Elphinstone and Matthews-Berry 2017). Therefore, if diversification is

occurring, we could expect to see some first signs of local adaptation, with isolates closer together being more phenotypically similar. However, this was not the case. This can be explained by constant mixing of *R. solanacearum* populations within UK rivers resulting in no geographical barriers and therefore no localised adaptation (Rainey and Travisano 1998). Another explanation can be due to fluctuations within the environment increasing the benefits of having multiple ecotypes present in the same location.

In contrast to location, the year of isolation could partly explain the phenotypic variation of UK *R. solanacearum* and suggests that the population has become a lot more phenotypically diverse within the past decade (figure 3.4B and appendix figure C.3C). Isolates from the 2010s have a significant higher chance of belonging to the specialist ecotype 3 and a significant lower chance of being in the generalist ecotype (ecotype 2). However, all ecotypes were observed in all three decades, suggesting that the overall phenotypic diversity has always been present and is maintained in the UK population. An increase in antibiotic resistance (gentamycin, tetracycline, and ciprofloxacin) in time was also observed, with lower resistance occurring in the 1990s, when the first recorded outbreak occurred, and higher resistance developing in the more recent years (figure 3.5). Antibiotics are naturally produced by competitor microorganisms within the environment (Allen et al. 2010) and resistance genes are associated with mobile genetic elements, which can be transferred between distantly related bacteria (Wellington et al. 2013). The increase in antibiotic resistance seen among the UK *R. solanacearum* population across time could be a response to tolerate specific antibiotics produced by competitors within the environment or could have been gradually inherited through horizontal gene transfer from nearby microorganisms throughout time. A variety of antibiotics can also be introduced to the environment through medical and agricultural sources. Antibiotics in humans are emitted to the sewage systems where they may be degraded or released into rivers (Wellington et al. 2013). Research has found that the concentration and frequency of detecting antibiotics within the river Thames in the UK increased during the influenza pandemic in November 2009 (Singer et al. 2014). This suggests that there is a transfer of antibiotics consumed by humans to the surrounding environmental rivers, such as

the river Thames where *R. solanacearum* has been consistently detected since the first outbreak (Parkinson et al. 2013, Elphinstone and Matthews-Berry 2017). Alternatively, antibiotics consumed by animals (livestock) are transmitted to agricultural fields as sludge to fertilise crops, and from here they can run off into nearby river systems introducing antibiotics to surrounding waterways (Blackwell et al. 2007, Topp et al. 2008, Wellington et al. 2013). Antibiotics are widely administered for animal health and are poorly absorbed in the gut resulting in a large proportion being excreted and therefore introduced into the environment through fertilisers (Sarmah et al. 2006). The reported concentrations of antibiotics detected in the soil and rivers are generally low, however some antibiotics from both urban and agricultural sources have been seen to persist in environments for longer periods of time (Allen et al. 2010, Wellington et al. 2013) and research has shown that even low environmental concentrations can select for antibiotic resistance (Lundström et al. 2016, Kraupner et al. 2018, Stanton et al. 2020). Furthermore, antibiotic use worldwide has increased both in agriculture and clinical use (Laxminarayan et al. 2020), suggesting a link with antibiotic presence and time. *Ralstonia solanacearum* has thought to have been introduced relatively recently to the UK, and the presence of antibiotics within this new environment could potentially cause a selective pressure for an increase in resistance over time within the population. This increase in antibiotic resistance in UK *R. solanacearum* can have implications in the control of this pathogen, such as reducing the efficacy of antibiotic producing biocontrol agents (Zhou et al. 2012, Singh and Kumar Yadav 2016), and suggests that environmental bacteria can act as reservoirs for antimicrobial resistance genes also having implications for antibiotic efficacy in clinical settings.

Temporal analysis also discovered that biofilm formation increased over time in the UK *R. solanacearum* population. Biofilm production can be a plant pathogen virulence trait (Álvarez et al. 2010, Genin and Denny 2012, Meng 2013), however it can also be used to protect bacteria from general stresses within the environment (de la Fuente-Núñez et al. 2013), such as bacteriophage predators (Hosseinidoust et al. 2013) as well as antibiotics (Høiby et al. 2010). As biofilm production is highly correlated with antibiotic resistance among the UK pathogen

population (appendix figure C.4E), this suggests that *R. solanacearum* could have adapted to have higher biofilm production in response to antibiotic presence in the UK rivers. Another explanation is that biofilm formation could have increased overtime within the UK *R. solanacearum* population for an unrelated reason, for example an environmental stress response such as salinity, bacteriophage or extreme pH within rivers (de la Fuente-Núñez et al. 2013), and an increase in antibiotic resistance is a by-product of this due to trait correlations.

Linking genetic data to trait variation among UK isolates using a genome wide association (GWAS) technique revealed very little genetic variation between the isolates, despite 30 years of sampling, supporting research suggesting that this is a clonal population of *R. solanacearum* (Stoycheva *et al.*, 2022: manuscript in progress). Despite this, a hypothetical protein was found associated with multiple traits, the majority of which are antibiotic resistance and biofilm production traits found to be increasing with time. A temporal GWAS also identified 153/138 sequence unitigs spanning across this same gene, suggesting that it is either linked with time or these specific traits. However, more information is needed to determine the cause of this association. Furthermore, a SNP within a putative deoxyribonuclease *RhsC* (RSUY_02640) gene, was also discovered to be associated with time, as well as the ecotype 3 (figure 3.6A), indicating that this variant may also be involved in adaptation to the UK environment. These identified hits can provide information on how *R. solanacearum* is changing over time within the UK, which could be due to selective pressures within the environment or due to stochastic mechanisms, such as bottlenecking of the population over winter, causing certain genes/SNPs to rise to high frequencies in the UK *R. solanacearum* population by chance. Exploring the genetic variation between the three identified ecotypes also highlighted very few differences between them, with 98% of genes being shared across all three ecotypes (figure 3.6C), despite clear significant phenotypic differences observed between isolates. Most traits are normally controlled by multiple genes and can be environmentally dependent, suggesting that there may be more information not captured within this study, for example epigenetics and phenotypic plasticity which are well-known *R. solanacearum*

adaptive strategies (Genin and Boucher 2002, Genin and Denny 2012), are responsible for the trait variation observed within this research.

In conclusion, UK *R. solanacearum* is phenotypically diverse, comprising of three ecotypes, which differ in their phenotypes and trait correlations. Linking phenotypic variation to isolates metadata revealed that neither location nor isolation source can explain trait variation, however the isolation time was positively associated with increased diversification over time. Antibiotic resistance traits and production of biofilm were also seen to be increasing among this population throughout time, potentially driving diversification of this pathogen. Genetic differences could not be linked with this trait variation, potentially suggesting that phenotypic plasticity or phenotypic switching are causing ecotype variation among this pathogen population. Therefore, future studies should focus on epigenetic variation, such as methylation patterns (Erill et al. 2017) among UK *R. solanacearum*. Overall, this study is the first to extensively focus on phenotypic diversity of *Ralstonia solanacearum* environmental isolates in the context of the whole lifecycle of this plant pathogen, increasing our understanding of the true phenotypic diversity among UK *R. solanacearum*.

4 Chapter 4: Adaptation to environmental stressors can explain diversification of phytopathogen *Ralstonia solanacearum*

4.1 Abstract

Plant pathogen *Ralstonia solanacearum*, the causative agent of bacterial wilt disease, commonly persists within environments, such as water sources, during transmissions from host to host. Within these environments exposure to various abiotic environmental stresses, including salinity and pH, is common. Exploring abiotic stress tolerance across a population of 182 *R. solanacearum* isolates revealed diversification based on acidic and alkaline tolerance, indicative of a trait correlation between the two traits. Trait correlations can determine the adaptation of an organism, with positive correlations resulting in generalists and negative trait correlations resulting in specialist adaptations. To study trait correlations directly, an evolution experiment was conducted where *R. solanacearum* was exposed to acidic, alkaline or salinity stresses alone or in two-stress combinations. It was found that selection resulted in specialist adaptation towards acidic and alkaline stress, whereas no adaptation to salinity was observed. However, the salt stress prevented *R. solanacearum* adaptation to acidic and alkaline conditions in stress combinations likely due to negative trait correlations between different tolerance traits. Furthermore, adapting to all stresses reduced *R. solanacearum* metabolic capacity, indicating a trade-off between stress tolerance and resource utilisation. Genome sequencing revealed little core genome (SNP/small INDEL) variation but stress adaptation was associated with insertion sequence movement in genes encoding virulence and metabolism traits. These results highlight the rapid evolution and diversification of *R. solanacearum* in response to environmental stresses, which could affect pathogen distribution and survival within environmental reservoirs.

4.2 Introduction

Bacteria are commonly exposed to stressful environmental conditions, including pH, temperature, and high osmolarity (Elabed et al. 2019), which can create strong selection for stress adaptations. Theory predicts that selection in stable environments will result in the evolution of specialist phenotypes (Bull 1987, Smits et al. 2006, Rainey et al. 2011). However, environments are rarely constant, but instead fluctuate in time and space (Horner-Devine et al. 2004, Hanson et al. 2012), which could favour the evolution of generalists. As adaptation to one environment can be associated with a fitness loss in another (Elena and Lenski 2003), fluctuating environments could select for generalists that exhibit intermediate fitness across multiple environments (Bull 1987, Smits et al. 2006, Rainey et al. 2011, Ferenci 2016). While the evolution of specialism and generalism have been studied extensively (Bleuven and Landry 2016), we still poorly understand how multiple stressors alone and together drive species adaptation within the environment.

Adaptive traits to multiple stressors within a given environment can be positively or negatively correlated. If positively correlated, adaptation to one stress is expected to lead to an increase in fitness in the other environment, leading to the evolution of a generalist organism that tolerates multiple stresses (figure 4.1). For example, beneficial mutations in *Escherichia coli* selected for in glucose-limited environments, also increased its fitness within other sugars (Ostrowski et al. 2005). However, more often traits are negatively associated with each other, resulting in conflicts due to trade-offs (Ferenci 2016). Trade-offs prevent organisms from achieving maximum fitness due to finite resources, where energy allocated for one trait reduces the investment in the other trait (Saltz et al. 2017). Alternatively, trade-offs can result from genetic conflicts due to epistasis (multiple loci interacting to produce a single phenotype) and pleiotropy (single locus affecting multiple different traits). Examples of such trade-offs include, changing protein function of one trait preferentially, such as reducing porin size to limit antibiotic uptake, which also reduces the metabolism via reduced nutrient uptake (Ferenci 2016). Similarly, evolving cefotaxime antibiotic resistance can result in reduced resistance towards ceftazidime in *Escherichia coli* (Schenk et al. 2015), while sharing molecular

resources, such as transcriptional machinery, which is regulated by the sigma transcription factor in *E. coli* can split resource allocation between stress resistance and metabolism (Ferenci 2016). Trade-offs can therefore constrain adaptation to multiple different stressors causing a reduction in fitness compared to adapting to a single stresses alone (Brennan and Collins 2015), limiting the range of phenotypes 'open' to organisms (Ferenci 2016). While many of the reported cross-tolerance interactions come from antibiotic research, it is less known how bacteria adapt to natural environmental stresses and to what extent selection in environmental reservoirs might affect the evolution of specialism and generalism.

Ralstonia solanacearum, the causative agent of bacterial wilt disease and potato brown rot, is an example of a phytopathogenic bacterium that encounters many different environmental selection pressures (Genin and Denny 2012). It has a large host range and can be found all over the globe, being one of the most important bacterial crop pests in the world (Hayward 1991, Genin 2010, Mansfield et al. 2012, Safni et al. 2014, Bragard et al. 2019). Not only can *R. solanacearum* infect a wide variety of plants, but it also survives across many natural environments outside hosts between transmissions, such as in water, and asymptomatic wild hosts, like *Solanum dulcamara*, where the bacteria can reside and survive for long periods of time (Genin and Denny 2012). *Ralstonia solanacearum* is hence capable of infecting host plants through the production of multiple virulence factors and evading plant immune responses (Genin 2010), as well as surviving external environments and tolerating a variation of abiotic stressors. Within Europe, *R. solanacearum* persists for long periods of time within environmental reservoirs, such as river networks, and several outbreaks of *R. solanacearum* in agricultural fields have been linked to these contaminated water sources (Parkinson et al. 2013, van der Gaag et al. 2019). Therefore, survival within these habitats is a critical part of *R. solanacearum's* life cycle and adaptation to environmental stresses can increase the prevalence of this pathogen thereby increasing risk of disease outbreaks.

One stressor particularly relevant to *R. solanacearum* is the change in pH within the environment, which can occur within the soil, surface water, as well as within the plant xylem (Bahrun et al. 2002, Secchi and Zwieniecki 2016, Li et al.

2017). Optimal pH for *R. solanacearum* is said to be between 5 and 6 (Li et al. 2017) and acidification or alkalization of environments, such as soils or water sources, can occur due to sewage, agricultural waste and chemical run offs (Rowbury 1997). For example, excessive application of nitrogen fertilizer plays a key role in the progress of soil acidification (Li et al. 2017). The spread and prevalence of wilt causing bacteria are closely related to soil pH and other soil properties (Bhanwar 2022), with research suggesting that bacterial wilt disease incidence is a lot higher in acidic soils (pH 4.5 - 5.5). Because of this, calcium amendments, like CaO and CaCO3, have been used to control bacterial wilt disease in China by changing pH and nitrite accumulation in fields (Jiang et al. 2017).

Another stressor that affects *R. solanacearum* survival is salinity, which is considered one of the major stress factors that inhibit microorganism survival (Zhang et al. 2021). *R. solanacearum* has been seen to survive in saline concentrations of up to 2% (Álvarez et al. 2010). Ions released from weathering minerals into the soil and rivers can cause naturally high salinity levels (Parkinson et al. 2013, Shrivastava and Kumar 2015). Salt may also be applied to rhizospheres through irrigation water or by excessive application of chemical fertilizers, especially when precipitation is insufficient (Shrivastava and Kumar 2015, Zhang et al. 2021). Salinized areas are also increasing at a rate of 10% annually for various reasons including, drought, irrigation with saline water, poor cultural practices, and global warming (Shrivastava and Kumar 2015). This indicates that tolerating high levels of salt will be beneficial for *R. solanacearum*, especially in combination with other stresses such as rising temperatures.

The aim of this study was to see if selection to abiotic stresses, pH and salinity, can drive *R. solanacearum* diversification using both comparative analysis of environmental samples and experimental evolution approach. In the UK, *R. solanacearum* has been recorded within environmental reservoirs, such as river networks, since the first reported outbreak in 1992 and since then several outbreaks in agriculture fields have been linked to these contaminated water sources (Parkinson et al. 2013). Therefore, tolerating these abiotic stresses will be critical for pathogen persistence within these environments and for transmission to crop hosts. Exploring the diversity of abiotic stress tolerance traits among a

population of *R. solanacearum* isolates from the UK (n=182) determined that there is large trait variation in tolerance towards acidity, alkalinity, and salinity stresses. This diversification towards abiotic stresses was then replicated within the laboratory by evolving one isolate to a variety of stresses and combined stress conditions. The overall hypothesis was that exposing *R. solanacearum* to different environmental stress conditions (pH and salinity) will result in different evolutionary outcomes depending on how these stress tolerance traits are correlated with one another (figure 4.1). Predictions included that exposure to acidity or alkalinity stress will drive *R. solanacearum* into two specialist niches if traits are negatively correlated, while positive trait correlations will result in generalist adaptations (figure 4.1). Furthermore, exposing *R. solanacearum* simultaneously to multiple stresses could lead to the evolution of generalists that are equally adapted to both stresses if these traits are positively correlated with one another. Alternatively, generalists with intermediate fitness in both stress conditions, or lack of adaptation, would occur if stress tolerances are negatively correlated (figure 4.1).

To test these hypotheses, an evolution experiment was conducted where one UK *R. solanacearum* isolate was exposed to different stress conditions alone (acid, alkaline and saline stress) or in combination (acidity with salinity and alkalinity with salinity) in liquid microcosms containing ten carbon resources for 40 days. Afterwards fitness assays were conducted, measuring the growth of evolved clones within different stress conditions and carbon resources to determine if adaptation had occurred. Furthermore, whole genome sequencing of a subset of evolved clones was conducted to determine genetic mechanisms of stress adaptation. Results showed that *R. solanacearum* adapts to acid and alkaline stress conditions with no correlated increase in fitness in the other stress condition, suggestive of specialist adaptation. Adapting in the presence of salinity prevented *R. solanacearum* evolution to either pH, suggesting that the double stress condition limited adaptation and was highly stressful for the pathogen causing extinctions when under extreme pH and high salinity concentrations (1% NaCl). Additionally, pH adaptation was accompanied with reduced growth on carbon resources, implying a trade-off between stress adaptation and metabolism. At the molecular level, adaptations were associated with movement of insertion sequences, mainly within

the megaplasmid, with a very few mutations (SNPs/small INDELs). Together, these results highlight the potential of different environmental stressors to alter *R. solanacearum* evolution and diversification within ecosystems, potentially affecting pathogen distribution and prevalence within environmental reservoirs.

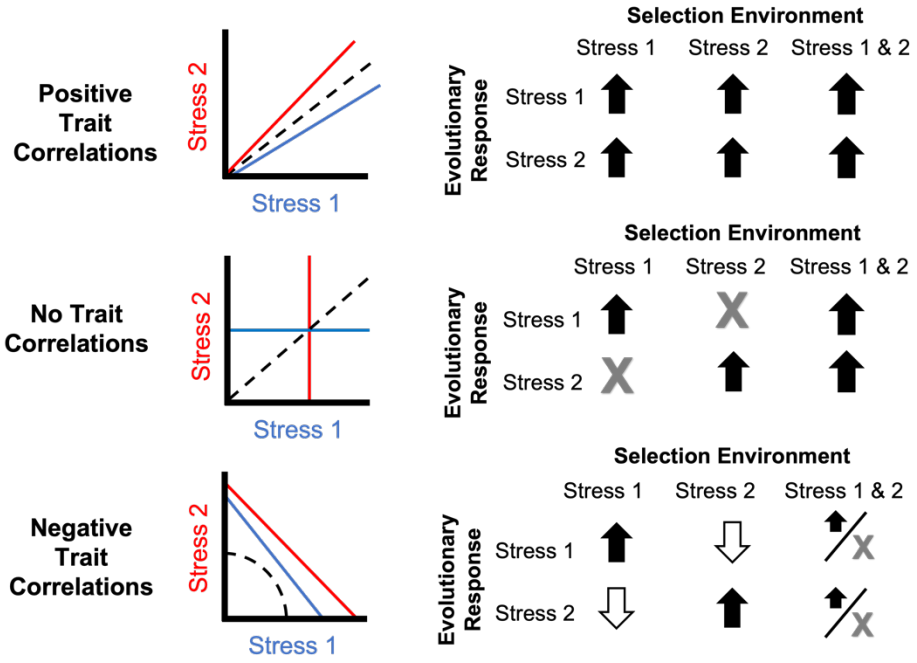


Figure 4.1: Schematic of how different correlations between stress traits determines evolutionary outcomes, leading to a specialist or generalist bacterium. Positive traits correlations will result in generalists for bacteria exposed to single stress (red and blue lines) and combined stress conditions (black dashed line). No trait correlations between the two stress conditions will result in specialists when exposed to single stress conditions (red and blue lines) with no trade-offs, while generalists will arise when exposed to the two stress conditions combined (black dashed line). However, negative correlations between the two traits will result in specialists with trade-offs in the other stress condition (red and blue lines) and intermediate generalists, or no adaptation when exposed to the combined stress condition (black dashed line).

4.3 Methods and Materials

4.3.1 *Ralstonia solanacearum* isolates and culture conditions

The 182 *Ralstonia solanacearum* bacterial isolates, from the United Kingdom (UK), used in this study were collected and stored by FERA Science ltd, Sand Hutton, York, UK. This sample set were curated from a variety of environment (river water and wild host *Solanum* dulcamara) and crop host (potato and tomato) samples collected between 1992 to 2019 (see appendix table A.2 for complete list). All isolates were verified as belonging to the *R. solanacearum* species complex (RSSC)

using 16S rRNA real-time PCR (RT-PCR) with (Weller et al. 2000) primers and the standard protocol advised by the European and Mediterranean Plant Protection Organization (EPPO) (EPPO 2018). One UK *R. solanacearum* isolate (York number: YO336, Protect number: 6941), collected from river water in 2007 from Cambridgeshire, was chosen for further evolutionary experiments as it was positioned in the 'middle' of the trait-space of the whole population (highlighted in figure 4.2B).

Modified OS media was used throughout the evolution experiment, this is a minimal medium containing salt required for bacterial growth (appendix table A.3). Media was prepared from various stock solutions to avoid co-precipitation of the salts with 10mM of carbon source added. SP media (20g/L sucrose, 5g/L peptone, 0.5g/L potassium hydrogen phosphate, 0.25g/L magnesium sulphate heptahydrate) (Mehan and McDonald 1995) was also used for relative growth in stressful conditions to explore diversity of abiotic stress tolerance among all isolates. CPG agar plates (1g/L casamino acid, 10g/L peptone, 5g/L glucose and 17g/L agar) (Kelman 1954) were used for colony-forming unit per ml counts (CFU/ml) with phosphate buffer solution (PBS) (8 g/L sodium chloride, 0.2 g/L potassium chloride, 1.44 g/l disodium phosphate, 0.24 g/l potassium dihydrogen phosphate) used to dilute bacterial populations. SPA plates (sucrose peptone media (Mehan and McDonald 1995) with 12g/L of agar) were also used for growth of single bacteria colonies isolated from evolved populations after the selection experiment for fitness assays and sequencing.

4.3.2 Abiotic stress tolerance diversity among the UK *R. solanacearum* population

Phenotypic traits involving tolerance to abiotic stresses were collected for 182 UK *R. solanacearum* isolates (see appendix table A.2) to investigate variation among the UK population due to abiotic stress tolerance (for geographical distribution of isolates see figure 4.2A). The 182 *R. solanacearum* isolates were standardised to 0.1 OD_{600} (optical density read at 600nm) in 25% glycerol before being cryopreserved at -80°C. On the day of the experiment, these isolates were

thawed and diluted 100-fold using sterile deionized water before 10µl of the inoculant was placed into 190µl of media in a 96-well microplate.

SP media was used, adjusted to pH 7, or 4.5, 9 and 10 for the pH stress conditions, and supplemented with 0.5%, 1% or 2% NaCl for salinity stress treatment before sterilisation (autoclaved at 120°C for 15 minutes). Microplates were placed in a 28°C incubator for 5 days with humidity to avoid evaporation. Change in bacterial growth was measured as optical density at 600nm (OD_{600}) with a spectrophotometer every 24 hours for a total of 5 days. Area under the growth curve (AUC) after five days of growth was used as a proxy of growth. AUC values were divided by isolate growth in the absence of any stress (pH 7 0% NaCl) to get the relative growth for each isolate to control for variation between isolates due to media preference. This data was then converted into a principal component analysis (PCA) plot to visualise the variation among UK *R. solanacearum* isolates due to abiotic stress tolerance (figure 4.2B).

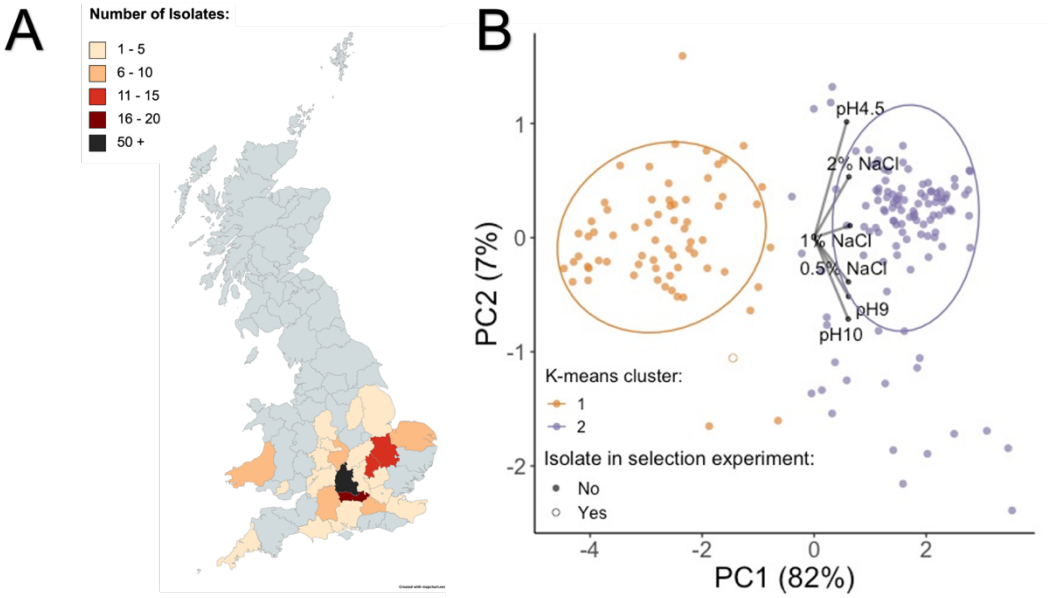


Figure 4.2: Abiotic stress tolerance causes diversification of UK *R. solanacearum* population into two groups. (A) Geographical locations where the 182 UK isolates were sampled, counties are shaded based on the number of samples originating from that region. Map curated using mapchart.net. (B) PCA plot showing UK *R. solanacearum* diversity in its ability to tolerate extreme pH (pH 4.5, 9 and 10) and salinity (0.5%, 1% and 2% NaCl). Each point represents one of the 182 UK isolates. Tolerance is measured as relative growth in that stress (growth in stress is divided by growth without stress). Each point represents a UK isolate (N=182) and the hollow circle indicates the isolate used in the selection experiment shown in Figure 2. PC1 and 2 both explain 91% of the variation in the data together. Loadings of the PCA plot are shown by arrows and in appendix figure D.1A. Isolates are coloured by cluster assigned using k-means clustering. Optimal of 2 clusters were determined using calinski method (appendix figure D.1B). Ellipses shows the 90% confidence interval around the centroid of each cluster.

4.3.3 Selection Experiment

As variation was observed among UK *R. solanacearum* in their abiotic stress tolerance (figure 4.2B), evolution towards different stress conditions was then tested experimentally by conducting a selection experiment. During the selection experiment a single colony of a UK *R. solanacearum* isolate (York number: YO336) was grown in several different stressful conditions including: control (pH 7 + 0% NaCl), salinity stress alone (pH 7 + 0.5% NaCl, pH 7 + 1% NaCl), extreme pH conditions alone (pH 4.5 + 0% NaCl, pH 8.5 + 0% NaCl), and a combination of the two stressors (pH 4.5 + 0.5% NaCl, pH 4.5 + 1% NaCl, pH 8.5 + 0.5% NaCl, pH 8.5 + 1% NaCl) (see figure 4.3A for schematic). Stress values were chosen based on preliminary results (appendix figure D.2). *R. solanacearum* was grown in these conditions at 28° C for 40 days, with 8 replicates per treatment, and 5% of the population was transferred into fresh OS media every 4 days (total of 40 days; 10 transfers; around 100 generations). OS media contained equal amounts of its constituent carbons (arabinose, glucose, maltose, sucrose, glutamine, histidine, serine, citric acid, malic acid, and succinic acid) resulting in at a total carbon concentration of 10mM in all treatments. Media was prepared, pH adjusted and NaCl added (for 0.5% and 1% conditions only) before filter sterilisation, using a 0.2µm filter, and storing the media in the fridge (4°C). Fresh media was prepared every week.

At every transfer 190ul of media was placed into each well of a 96-well microplate and 10µl of *R. solanacearum* was added to each well. Starting *R. solanacearum* densities were $2.2x10^6$ CFU/ml in all treatments. Microplates were kept in a 28°C incubator with humidity to prevent evaporation. Before each serial transfer, optical density (OD) reads were taken at 600nm using a spectrophotometer with shaking, and at every second transfer (8-day intervals) colony-forming units per ml (CFU/ml) were recorded. CFU/ml were calculated by spotting 5µl of bacteria from each of the serial dilutions (10^0-10^{-7}) onto CPG (Kelman 1954) agar plates, for the dilutions phosphate buffer solution (PBS) was used. These plates were then placed in a 28°C incubator for 48 hours before

colonies were counted. Populations were also cryopreserved every second transfer in 25% glycerol at -80°C.

4.3.4 Fitness assays for quantifying evolutionary changes in stress and metabolic adaptations

At the end of the selection experiment, evolved *R. solanacearum* populations were spread onto SPA plates, including the ancestor, and placed in a 28°C incubator for 48 hours. 12 colonies per population (16 for the ancestor) were then randomly picked and cryopreserved in 25% glycerol at -80°C. These clones (n=96 per stress condition) were then compared with the ancestor (n=16) in a variety of conditions to determine if they had adapted to certain stress conditions (pH 4.5, pH 8.5, 0.5% NaCl, pH 4.5 + 0.5% NaCl and pH 8.5 + 0.5% NaCl), and how that impacted *R. solanacearum's* ability to utilise certain carbons resources in the media (10mM of arabinose, glucose, maltose, sucrose, glutamine, histidine, serine, citric acid, malic acid, and succinic acid) as well as all of the ten resources combined at pH 7 (see figure 4.3B for schematic).

Before these fitness assays were conducted, all evolved clones, along with the ancestor isolate, were grown up in a common environment (10 carbon OSG media at pH 7 0% NaCl) to control for any changes due to phenotypic plasticity and to ensure cell densities were high and in exponential phase for the fitness assays. For this 10µl of each cryopreserved clone was grown in 190µl of mixed carbon media for 24 hours at 28°C and humidity to avoid evaporation. After 24 hours, clones were diluted 10-fold using sterile deionized water and 10µl of this inoculum was transferred to 190µl of each assay condition, this included growth in individual stressors (pH 4.5, pH 8.5 and 0.5% NaCl), combined stresses (pH 4.5 + 0.5% NaCl, pH 8.5 + 0.5% NaCl), or the OS ten carbon mixture (pH 7 0% NaCl), as well as in each single constituent carbon media, with carbon resources always at an overall concentration of 10mM. Clones were incubated at 28°C with humidity to avoid evaporation and growth was measured as optical density (OD_{600}) every 24 hours for five days. This data (see appendix figure D.4 and D.5) was then used to calculate area under the curve (AUC), growth rate (r) and carrying capacity (k) as a proxy of bacterial growth.

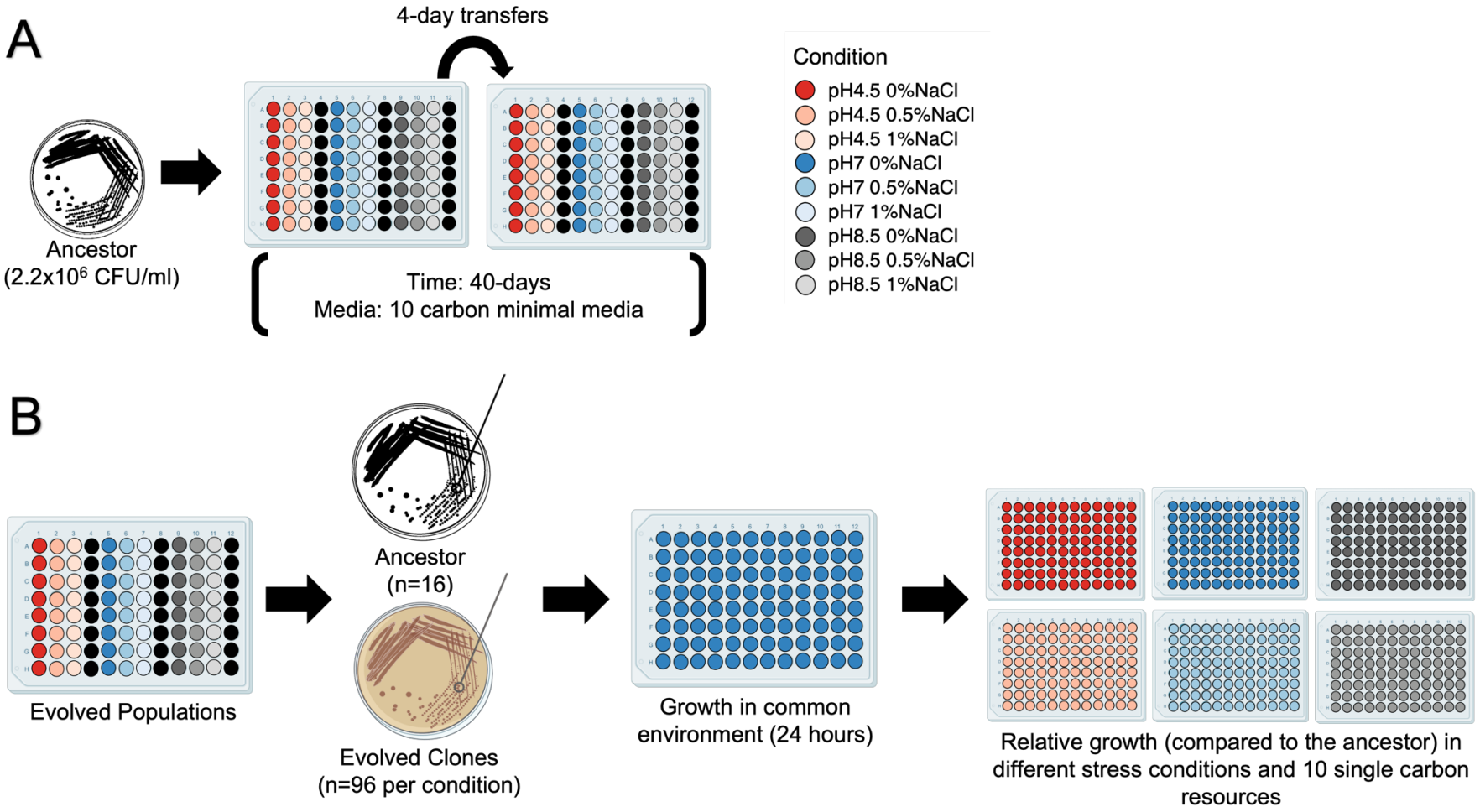


Figure 4.3: Schematic of experimental evolution methods. (A) The selection experiment layout using 96-well microplates. 10µl of the ancestral *Ralstonia solanacearum* isolate (at 2.2x10^6 CFU/ml) was added to 190µl of 10 carbon minimal media at different pH and salinity concentrations. 8 replicate populations per condition were

conducted and populations were transferred into fresh media every 4 days, for a total of 40 days. Optical density reads were taken every 4 days and colony forming units (CFU) per ml were calculated every 8 days. The colour of wells within the 96-well microplates indicate the condition (pH or salinity stress) and back wells indicate empty wells. (B) Fitness assay experimental schematic. After the selection experiment 12 evolved clones per evolved population (n=96 per condition), and 16 ancestral clones, were taken from SPA plates. They were then grown for 24 hours in neutral conditions (10 carbon minimal media at pH7 and 0% NaCl) to control for phenotypic plasticity and increase cell densities for sufficient replication. 10µl of these grown clones were then diluted 10-fold and placed into multiple 96-well plates at different pH, salinity and carbon media to determine relative growth compared to the ancestor in a variety of conditions.

4.3.5 Identification of genetic variants

One clone from each of the 8 replicate populations, evolved in neutral (pH 7 0% NaCl), saline (pH 7 0.5% NaCl), acidic (pH 4.5 0% NaCl) and alkaline (pH 8.5 0% NaCl) conditions was randomly chosen for Illumina sequencing. DNA was extracted using the Qiagen DNeasy blood and tissue kit according to the manufacturer's protocols with a couple of modifications (see appendix table A.4 for protocol). Quality checks of extracted DNA were undertaken using nanodrop and gel electrophoresis imaging of 0.8% agarose gels, and DNA was quantified using the Quant-iT™ dsDNA broad-range assay kit as per the manufacturer's microplate protocol. Samples were then diluted to 15 ng/µl in 10mM tris-HCl (pH 8.4) and 50µl was sent to MicrobesNG for sequencing (30x coverage). MicrobesNG conducted library preparation using Nextera XT Library Prep Kit following the manufacturer's protocol with the following modifications: 2 ng of DNA were used as input, and PCR elongation was increased to 45 seconds. Hamilton Microlab STAR automated liquid handling system was used for DNA quantification and library preparation. Pooled libraries were quantified using the Kapa Biosystems Library Quantification Kit for Illumina on a Roche light cycler 96 qPCR machine. Libraries were sequenced on the Illumina sequencers using a 250 bp paired end protocol. Reads were adapter trimmed by MicrobesNG using Trimmomatic 0.30 (Bolger et al. 2014) with a sliding window quality cut-off of Q15 . De novo assembly was performed using SPAdes version 3.7 (Bankevich et al. 2012) and contigs annotated using Prokka 1.11 (Seemann 2014) (https://microbesng.com/).

For the variant calling, reference genome of a highly related *R. solanacearum* UY031 strain was downloaded from RefSeq (accession number: 001299555.1) as this is currently the best assembled phylotype IIB *R. solanacearum* strain. Variant calling was then conducted using SNIPPY version 4.6.0 (Seemann 2015) or with custom filtering requirements. For personalised filtering, alignment files (bam) were created using bwa version 0.7.17 (Li and Durbin 2009) and used as the input files for Freebayes version 1.3.2 (Garrison and Marth 2012), which was used to identify single nucleotide polymorphisms (SNPS) and insertion/deletion (INDEL) variants compared to the reference strain. These variants were then filtered using vcftools (Danecek et al. 2011) to only include those with a minimum overall

read depth of 5, a minimum number of reads to include that genotype as 5, and a quality score above 30. Finally, SnpEff version 4.3 (Cingolani et al. 2012) was used to annotate and predict the impact of those variants. Annotated variant files from both methods were then imported to R version 4.1.2 (R Core Team 2022) using the vcfR package (Knaus and Grünwald 2017), where all further downstream analyses were performed, including the removal of variants present within the ancestor. Variants that were detected by only one of these filtering methods were confirmed manually using the integrative genomics viewer (IGV) platform (Robinson et al. 2011), along with manually determining that the parallel mutations detected were also not present within the ancestor clone. A full table of each variant can be found in appendix table D.4.

For insertion sequence (IS) identification, IS that were identified within a highly similar UK isolate (York number: YO199), from (Greenrod et al. 2022), were blasted against the ISFinder database (https://isfinder.biotoul.fr/). IS with the highest score, which resulted in 8 different IS elements (IS1021, IS1421, ISAzo23, ISRso10, ISBma3, ISRso20, ISRel26 and ISRso7), were then used as reference IS to search for within our evolved clones. IS within the evolved clones, along with the ancestor, were then identified using short read data with ISMapper version 2.0.2 (Hawkey et al. 2015) using methods as described by (Greenrod et al. 2022). Overall, ISMapper maps short read data to our reference IS sequences, identifying reads that map to and overhang the 3' and 5' IS flanks. Mapped reads were then further aligned to an annotated reference bacterial genome, we used the highly related *R. solanacearum* UY031 strain downloaded from GenBank (NCBI accession number: 001299555.1, gbff format file), and IS positions are determined where 3' and 5' flanking reads both map to similar genomic locations. After running ISMapper, IS hits were filtered to remove potential false positives within R version 4.1.2 (R Core Team 2020). Filtering included removing IS with unknown 5' or 3' coordinates, along with removing IS elements present within all 33 clones. IS hits that overlapped with one another, by plus or minus 100 bases, were reannotated to the same position due to high chances of this being a misalignment. IS elements within transposons were also removed as IS that map inside multi-copy genes will map to all copies

irrespective of the true IS content and therefore may generate spurious hits (Hawkey et al. 2015).

4.3.6 Statistical Analysis

All data and statistical analyses were conducted using R version 4.1.2 (R Core Team 2020). Data manipulation was performed using the tidyverse suite of packages (Wickham et al. 2019) and production of graphs with ggplot2 (Wickham 2016). Principle component analyses were conducted using the prcomp function in the stats package (R Core Team 2020) and PERMANOVAs were conducted using the adonis2 and betadipser function in the vegan package (Oksanen et al. 2019). Growth rate, carrying capacity and area under the curve data was all taken using up to four days' worth of growth data. Growth rate (r), the speed at which the number of organisms in a population increases, was calculated using the below formula:

$$Growth\ Rate\ (r) = ((log_{10}N - \ log_{10}N_0)2.303)/(t - t_0)$$

Where t_0 is the time (in hours) before exponential growth occurred (time at which the lowest OD_{600} was recorded). t is the time at which the highest OD_{600} was recorded (after exponential growth). N_0, the OD_{600} measurement at the time before exponential growth and N the OD_{600} value at the timepoint designated after exponential growth. Carrying capacity (k), the maximum population size that can be sustained by that specific environment, was calculated by taking the maximum optical density (OD_{600}) read of all recorded time points. Area under the growth curve, which captures both growth rate and carrying capacity, was calculated using the auc function in the MESS package (Ekstrøm 2019). Relative growth was then calculated by dividing each clone's growth (AUC, r or k) by the ancestors in the same media condition (appendix figure D.6).

Significance between population densities across the selection experiment was determined via repeated measures ANOVA using the aov function (R Core Team 2020). This model included density (OD_{600} or CFU/ml) as the response variable, and pH and salinity as the predictor factors, with replicate populations set as random effects across time. Post-hoc pairwise comparisons were then computed

using Tukey's significance test through the emmeans package (Lenth 2021). Significance between evolved isolates growth in different stressors or carbons was determined with repeated measures ANOVA using the aov function. This model had growth (AUC, r, or k) as the response variable, and evolutionary history and growth condition (stress or carbon) as the predictor factors. Population was inputted as a random effect and ancestor information was excluded due to lack of populations. Separate one-way ANOVAs were then conducted, including the ancestor growth measurements (n=16), for each condition once significance was observed overall. Mean growth (AUC, r or k) was calculated across all clones (n=12) per population (n=8) to control for pseudo replication. The model consisted of growth (AUC, r or k) as the response variable and evolutionary history as the predictor factor and the aov function (R Core Team 2020) was used. Post hoc tests were also conducted Tukey's significance test through the emmeans package (Lenth 2021). Mean pairwise distances were calculated by taking the mean Euclidean distance per evolved clone from all other clones (along with the ancestor clones), using the dist() function (R Core Team 2020) on a matrix of the relative growth in each stress condition (acid, alkaline and saline stress). The median of the mean pairwise distance between different evolutionary history groups were then compared using Kruskal-Wallis significance test, using the kruskal.test function, and pairwise Wilcoxon test, using the pairwise.wilcoxon.test function, both available within the stats R package (R Core Team 2020), with Benjamini-Hochberg (BH) as the multiple test correction method.

4.4 Results

4.4.1 Abiotic stress tolerance drives diversification of *R. solanacearum* UK population into two distinct groups

To investigate whether exposure to different stress conditions can cause diversification of a *R. solanacearum* population a collection of 182 UK isolates was used (see appendix table A.2) and tolerance to abiotic stresses, such as extreme pH and salinity, were quantified. These isolates were collected from environmental reservoirs, primarily from water sampled from different rivers and the associated

wild host *Solanum dulcamara*, across a timescale of 30 years, beginning in 1992 with the first recorded case to 2019 (for geographical distribution of sampling see figure 4.2A). Relative growth in stress conditions was measured and converted into a principal component analysis (PCA) plot to visualise the variation of abiotic stress tolerance among UK *R. solanacearum* isolates (figure 4.2B). This PCA plot explained 91% of total variation of the dataset, with the majority (82%) explained by principal component 1 (PC1) and 7% by principal component 2 (PC2). A higher PC1 value corresponded to an increase in abiotic stress tolerance across all stresses, suggesting that an increase in growth under one stress was correlated with an increase in growth under another stress. However, PC2 showed differences in abiotic stress tolerance with higher values contributed to higher tolerance to acidic stress (pH 4.5) and high salinity concentrations (2% and 1% NaCl), while negative values indicated higher tolerance to alkaline stress (pH 9 and pH 10) and low salinity concentrations (0.5% NaCl) (figure 4.2B and appendix figure D.1A). Clustering of UK isolates into two groups was observed based on their overall abiotic stress tolerance (figure 4.2B, PERMANOVA: $F_{1,180} = 669$, $R^2 = 0.79$, $p = 0.001$) along PC1, possibly suggesting divergence of UK isolates into two phenotypically distinct groups. To explore this further, metadata (time, location, and source isolated from) were analysed to determine if time, location, or isolation source can explain the observed diversity in abiotic stress tolerance among UK isolates. However, no significant differences was observed between isolates from different hosts (appendix figure D.3A, PERMANOVA: $F_{3,178} = 1.03$, $R^2 = 0.02$, $p = 0.37$), decades (appendix figure D.3B, PERMANOVA: $F_{2,177} = 0.92$, $R^2 = 0.01$, $p = 0.42$) or location (appendix figure D.3C, PERMANOVA: $F_{1,172} = 0.56$, $R^2 = 0.003$, $p = 0.52$). Together, this analysis suggests that UK *R. solanacearum* population varies regarding stress tolerance (figure 4.2B), which cannot be explained by location, host, or time of isolation. How this diversity arose and how different types of stress might drive distinct stress adaptations is still unclear, however the presence of different stresses could drive *R. solanacearum* adaptation into different directions, leading to increased diversity. Therefore, stress adaptation was replicated within laboratory conditions to explore this hypothesis.

4.4.2 Abiotic stresses reduce bacteria densities with extinctions occurring in combined stress conditions

To causally test whether selection by pH and salinity can drive *R. solanacearum* adaptation, and how the presence of multiple stressors affects pathogen survival, one UK *R. solanacearum* isolate (YO336), positioned in the 'middle' of the trait-space of the whole population (highlighted in figure 4.2B), was selected and evolved in a variety of stresses in the lab using experimental evolution (n=8 per stress condition). The effect of stress on the growth of *R. solanacearum* populations throughout the experiment was analysed using optical density (OD), to determine total cell densities, and colony forming unit (CFU) data, to determine viable cell densities. Population densities, as measured by OD, in the no stress control (pH7) were relatively constant at around 0.6 OD, with a slight peak occurring at day 28 and a gradual decrease afterwards (dashed line in figure 4.4A-C). Bacteria densities measured by CFU were more variable with a gradual increase over time, with a drop in population densities at day 24 and a gradual increase afterwards (figure 4.4D). As these fluctuations in OD and CFU occurred in all stress conditions this suggests that uncontrolled changes within the laboratory environment are likely responsible.

Overall, 0.5% salinity significantly reduced total *R. solanacearum* population densities (OD) compared to 0% NaCl (figure 4.4B; Repeated Measures ANOVA: $F_{2,63}$= 2281, p<0.0001; Tukey: p < 0.0001) especially at the start of the selection experiment (figure 4.4B). However, there were no changes in CFU compared to the no stress control with very similar population densities across time (figure 4.4D). A higher concentration of salinity (1% NaCl) reduced population densities even further, with both OD and CFU measurements being extremely low at the start and gradually increasing throughout the selection experiment, but never reaching comparable densities to the no stress control (figure 4.4B and D). Both OD and CFU population density measurements were significantly lower when exposed to 1% NaCl compared to 0% NaCl (figure 4.4B; Repeated Measures ANOVA: $F_{2,63}$= 2281, p<0.0001; Tukey: p < 0.0001 for OD and figure 4.4D; Repeated Measures ANOVA: $F_{2,63}$= 7.9, p= 0.0009; Tukey: p=0.005 for CFU).

Alkaline stress also significantly reduced pathogen densities compared to the neutral control (pH 7), both with OD (figure 4.4B-C; Repeated Measures ANOVA: $F_{2,63}$= 793, p<0.0001; Tukey: p < 0.0001) and CFU measurements (figure 4.4D; Repeated Measures ANOVA: $F_{2,63}$= 6.2, p= 0.003; Tukey: p=0.01). In contrast, acidic stress did not significantly reduce *R. solanacearum* densities relative to control treatment (pH7) in case of both OD (figure 4.4A-B; Repeated Measures ANOVA: $F_{2,63}$= 793, p<0.0001; Tukey: p=0.09) and CFU (figure 4.4D; Repeated Measures ANOVA: $F_{2,63}$= 6.2, p= 0.003; Tukey: p=0.3) data.

Exposing *R. solanacearum* simultaneously to two stresses, magnified the negative effects of each pH stress. For example, all populations went extinct in the presence of both high salt (1% NaCl) and either acid (pH 4.5) or alkaline (pH 8.5) stresses (figure 4.4D). All 8 populations exposed to alkalinity with 1% salinity became extinct after 8 days of exposure, and populations exposed to both acidity and 1% salinity combined went extinct within 24 days. Also, the additional effect of low salt concentrations (0.5% NaCl) alongside acidic stress (pH 4.5) significantly reduced *R. solanacearum* population densities (OD) compared to pH 4.5 treatment alone (figure 4.4A; Repeated Measures ANOVA: $F_{4,63}$= 141, p<0.0001; Tukey: p<0.0001). Similarly, combining salinity with alkaline pH stress significantly reduced *R. solanacearum* population densities (OD) compared to populations growing in the pH 8.5 treatment alone (figure 4.4C; Repeated Measures ANOVA: $F_{4,63}$= 141, p<0.0001; Tukey: p<0.0001), with consistently lower OD throughout the selection experiment. The addition of pH stress, pH 4.5 or pH 8.5, with 0.5% salinity also significantly reduced *R. solanacearum* populations OD compared to populations exposed to salinity stress alone (0.5% NaCl at pH 7) (figure 4.4A-B; Repeated Measures ANOVA: $F_{4,63}$= 141, p<0.0001; Tukey: p<0.0001 for pH 4.5 with salinity and figure 4.4B-C; Repeated Measures ANOVA: $F_{4,63}$= 141, p<0.0001; Tukey: p<0.0001 for alkaline with salinity). However, these stress specific effects on population densities were not observed based on the CFU data. Overall, the presence of stress conditions reduces *R. solanacearum* population densities, apart from acidic stress, while when combined with salinity stress led to even more reduced population densities, causing extinctions when high salinity concentrations (1% NaCl) was combined with extreme pHs.

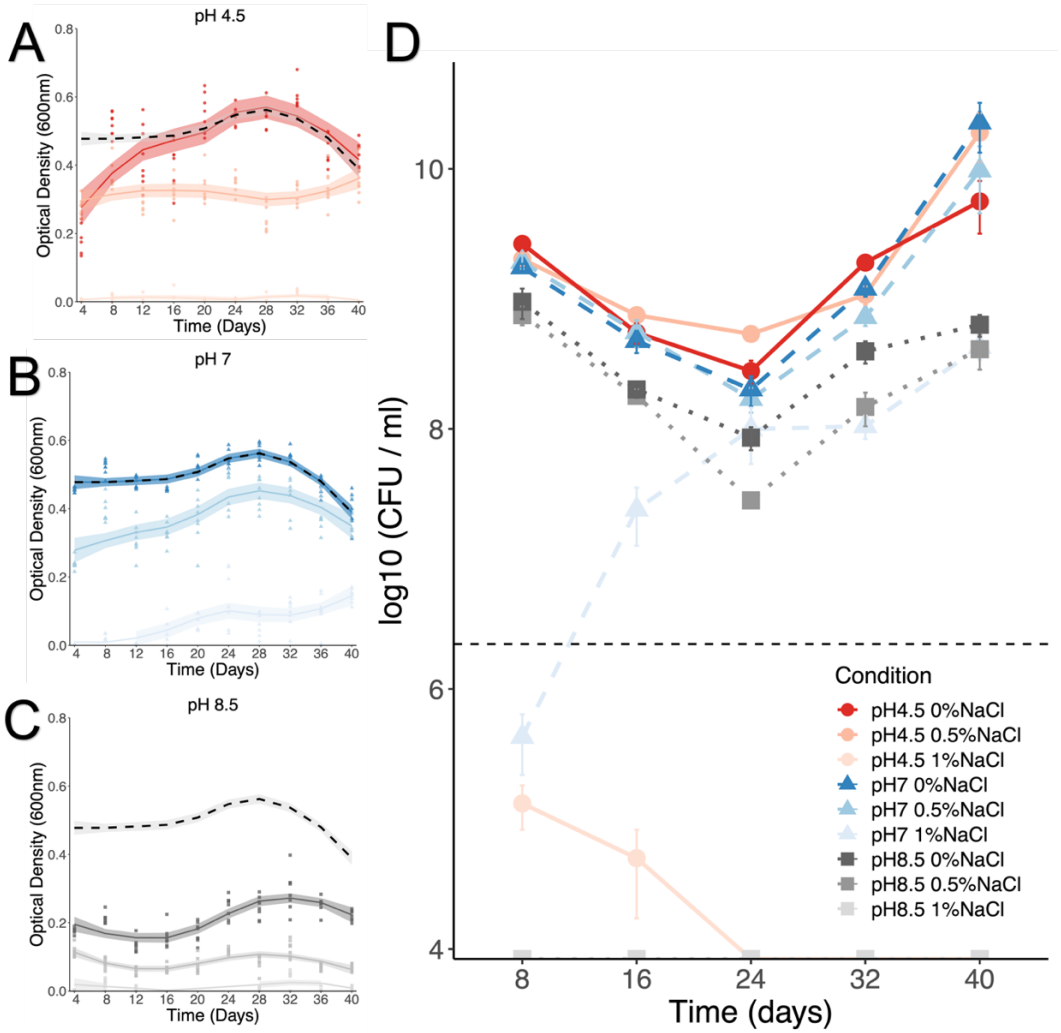


Figure 4.4: *R. solanacearum* population dynamics during the selection experiment. Changes in population densities were measured as optical density (OD) reads at 600nm before each transfer over the course of the evolution selection experiment. Panels show the pH to which the bacteria were exposed to; (A) pH 4.5 (B) pH 7 (C) pH 8.5 and lighter colours indicate increasing salinity concentrations in the media. Points indicate individual replicate populations OD measurement, and the solid line shows the fitted model (loess method) across the 8 populations, with 95% confidence intervals shaded around this line. The dashed line shows the fitted model (loess method) for the no stress control populations (pH 7 with 0% NaCl). (D) colony-forming units per ml (CFU/ml) were calculated every two transfers (8-day interval) throughout the selection experiment. Dashed black line indicates initial starting densities in the media at the start of the 40-day selection experiment. Points shows log base 10 of the mean average CFU/ml across the 8 replicate populations and error bars indicate one standard error from this mean.

4.4.3 *R. solanacearum* adapts to pH abiotic stresses

To test if specialist adaptation to abiotic stresses occurred during the selection experiment 12 clones per replicate population from all treatments were isolated and fitness assays conducted where growth (growth measured as area under the curve (AUC), growth rate (r) and carrying capacity (k)) of ancestral and evolved clones was compared in three different stress conditions: pH 4.5, pH 8.5 and 0.5% salinity. Growth compared to the ancestor was then calculated per clone and the average across all clones per evolutionary history (a total of 96 clones, 12 per treatment replicate) was taken as a measure of relative growth in that condition (figure 4.5). Metabolism utilisation of the ten individual carbon resources within the media for each evolutionary treatment was also measured to understand the effect stress adaptation has on *R. solanacearum* growth. Overall, variation in clones' growth was explained the interaction between evolutionary history and stress condition (figure 4.5A, Repeated Measures ANOVA: $F_{25,175}$= 6.2, p<0.0001, figure 4.5B, Repeated Measures ANOVA: $F_{25,175}$= 7.8, p<0.0001 and figure 4.5C, Repeated Measures ANOVA: $F_{25,175}$= 8.8, p<0.0001 for AUC, r and k respectively) and between evolutionary history and carbon condition (figure 4.5A, Repeated Measures ANOVA: $F_{45,315}$= 5.2, p<0.0001, figure 4.5B, Repeated Measures ANOVA: $F_{45,315}$= 6.9, p<0.0001 and figure 4.5C, Repeated Measures ANOVA: $F_{45,315}$= 7.6, p<0.0001 for AUC, r and k respectively). Therefore, comparisons of stress adaptation within each condition was conducted to understand *R. solanacearum* adaptation under single and two stresses, as well as to compare growth in each carbon recourse per evolutionary history to better understand potential metabolic costs of *R. solanacearum* adaptation.

Clones exposed to no stress (pH 7 with 0% salinity) showed a non-significant increase in growth on pH 7 mixed carbon media compared to the ancestor (figure 4.5, appendix tables D.1-3), indicative of media adaptation. This can be attributed to a significant increase in growth on histidine (k and r) and citric acid (r) (figure 4.5B and C, appendix tables D.2 and D.3). However, a significant reduction in AUC on one of the ten single carbon resources, glucose, was also observed (figure 4.5A, appendix table D.1). The pH7 evolved clones also showed a significant increase in growth (r and k) compared to the ancestor in acidic media (figure 4.5, appendix

tables D.2 and D.3) and an increased carrying capacity (k) compared to the ancestor in alkaline conditions (figure 4.5C, appendix table D.3), suggesting that adaptation to the media also resulted in increased growth in acidic and alkaline stress conditions. However, the increase in growth (AUC, r, and k) compared to the ancestor within salinity stress conditions was non-significant (figure 4.5A, appendix tables D.1-3). Overall, in the absence of stress *R. solanacearum* adapts to the ten mixed carbon media, which is also correlated with increased growth in acidic and alkaline pH.

Ralstonia solanacearum adaptation towards acidic pH (pH 4.5) was observed as evolved clones had significant higher growth (AUC, r and k) in pH 4.5 stress compared to the ancestor (figure 4.5, appendix tables D.1-3). However, adaptation to acidic pH did not lead to changes in clones' growth in the mixed carbon media (figure 4.5, appendix tables D.1-3), suggesting that this adaptation was not due to increased growth on the media. Clones also suffered a reduced growth (AUC) on six of the ten individual carbon resources that were within the media (figure 4.5A, appendix table D.1), suggestive of a trade-off between acid stress adaptation and metabolism. Moreover, adaptation towards acidic pH also did not result in changes in growth within salinity stress conditions compared to the ancestor (figure 4.5A, appendix tables D.1-3). However, it did result in lower growth (AUC and r) in alkaline stress conditions compared to the ancestor (figure 4.5B, appendix tables D.1 and D.2), indicative of a trade-off between the two extreme pH stress conditions. In summary, acid stress adaptation occurred which resulted in reduced metabolic ability and alkaline tolerance suggestive of trade-offs between these traits.

Ralstonia solanacearum clones exposed to extreme alkaline stress (pH 8.5) throughout the evolution experiment evolved clearly higher carrying capacity in the alkaline environment, compared to the ancestor, indicative of adaptation to pH 8.5 (figure 4.5C, appendix table D.3). Alkaline adapted clones had no significant difference in their growth in acidic conditions (figure 4.5A, appendix tables D.1-3), and salinity stress compared to the ancestor (figure 4.5A, appendix tables D.1-3). Furthermore, alkaline evolved strains also did not show any signs of media adaptation, with no significant difference in their AUC, r and k compared to the

ancestor (figure 4.5A, appendix tables D.1-3) within the 10 carbon media. Furthermore, they had a reduced growth (AUC and k) in all ten of the single carbon resources present within the media compared to the ancestor (figure 4.5A, appendix tables D.1 and D.3), suggesting a reduced ability to utilise resources within the media and a trade-off between alkaline adaptation and metabolism.

Adaptation to salinity was not observed even though salinity evolved clones at pH 7 showed a slight increase in their growth (AUC, r and k) in salinity compared to the ancestor, which was however non-significant (figure 4.5A, appendix tables D.1-3). Salinity exposed clones also had no significant difference in growth on the media without stresses (figure 4.5A, appendix tables D.1-3), and a reduction in growth (AUC) in alkaline conditions compared to the ancestor (figure 4.5A, appendix table D.1). Furthermore, salinity exposed clones did have an increase in growth rate (r) and carrying capacity (k) within acidic media compared to the ancestor (figure 4.5C, appendix tables D.2 and D.3). Despite no differences in growth within the mixed ten carbon media, growing in the presence of salinity led to reduced growth (AUC) compared to the ancestor on nine of the ten individual carbon resources (figure 4.5, appendix table D.1), suggesting negative trait correlations between salinity exposure and metabolism. No significant growth differences were observed between the ancestor and 1% NaCl exposed isolates at pH7 (appendix figure D.3 and D.4), possibly due to the low population densities and too few cycles of replication for adaptation to occur. Overall, salinity adaptation was not clearly observed during the selection experiment, even though growing in the presence of salinity increased *R. solanacearum's* carrying capacity in acid stress conditions and reduced their ability to utilise resources within the media.

4.4.4 Presence of multiple stresses hinders *R. solanacearum* adaptation to extreme pH

Next, the aim was to determine if adaptation towards multiple stresses can occur and if this resulted in the evolution of generalists adapted to both stresses. It was found that isolates that were exposed to two stressors during the evolutionary experiment did not adapt to either stress, with pH 4.5 and salinity evolved clones having significant lower AUC in acidic conditions compared to the ancestor (figure

4.5A, appendix table D.1) and salinity conditions compared to the ancestor (figure 4.5A, appendix table D.1). This suggests that not only have these strains failed to adapt to pH and salinity stresses, they also have a reduced ability to grow in these stress conditions compared to the ancestor. Alternatively, population sizes could be smaller within these stress conditions hindering bacteria adaptation. Combined alkaline and salinity stress exposed *R. solanacearum* also showed significant reduction in growth (AUC) in alkaline (figure 4.5A, appendix table D.1) and a non-significant reduction in growth within salinity (figure 4.5, appendix tables D.1-3) stress conditions compared to the ancestor, also indicating no adaptation towards either stress environment. This suggests that simultaneous exposure to multiple stresses limits *R. solanacearum's* adaptation, which is supported by the extinctions that occurred in the combined acid and alkaline stress with 1% salinity during the evolution experiment.

The presence of double stress also constrained growth on the ten carbon mixed media (pH 7 0% NaCl), with clones grown in pH 4.5 with 0.5% salinity having significant reduced growth (AUC and r) compared to the ancestor (figure 4.5A, appendix tables D.1 and D.2), and reduced AUC compared to the ancestor in eight of the ten individual carbons this mixed media is composed of (figure 4.5A, appendix table D.1). Clones exposed to combined alkaline stress with salinity also showed significant reduced growth in AUC on five of the ten single carbon resources compared to the ancestor (figure 4.5A, appendix table D.1). This suggests that exposure to double stress can lead to reduced resource utilisation even in the absence of improved growth in stressful environmental conditions.

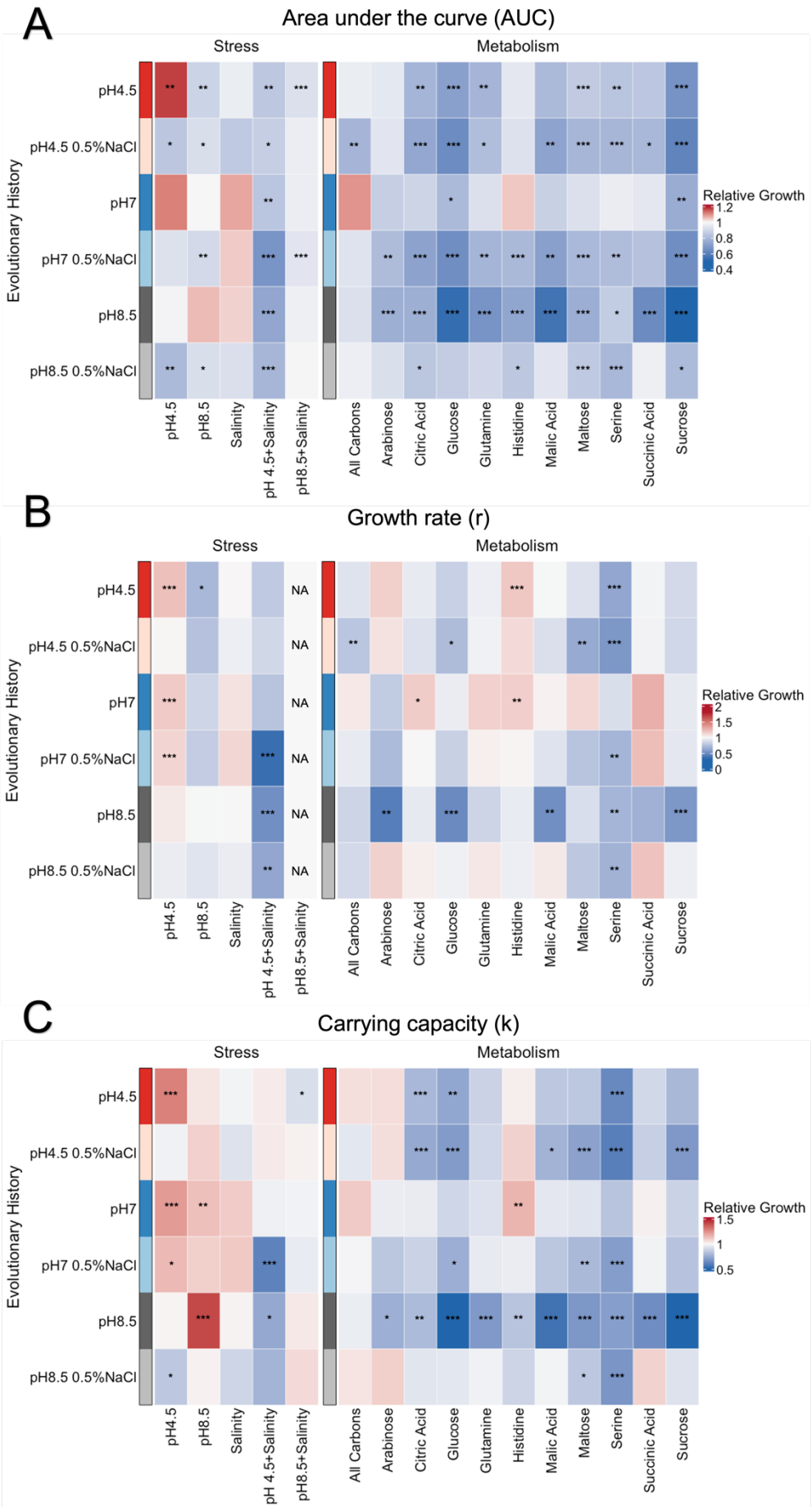

A
Area under the curve (AUC)
Stress
Metabolism
Evolutionary History
pH4.5
pH4.5 0.5%NaCl
pH7
pH7 0.5%NaCl
pH8.5
pH8.5 0.5%NaCl
pH4.5
pH8.5
Salinity
pH 4.5+Salinity
pH8.5+Salinity
All Carbons
Arabinose
Citric Acid
Glucose
Glutamine
Histidine
Malic Acid
Maltose
Serine
Succinic Acid
Sucrose
Relative Growth
1.2
1
0.8
0.6
0.4
B
Growth rate (r)
Stress
Metabolism
Evolutionary History
NA
Relative Growth
2
1.5
1
0.5
0
C
Carrying capacity (k)
Stress
Metabolism
Evolutionary History
Relative Growth
1.5
1
0.5

Figure 4.5: Average relative growth compared to ancestor for each evolved condition in the presence of abiotic stressors and ability to utilise different carbon resources. (A) Area under the curve (AUC), (B) growth rate (r) and (C) carrying capacity (k), was calculated for each evolved clone after 4 days of growth and relative growth was calculated by dividing AUC/r/k in each condition by the ancestor's value in the same condition. Growth was conducted within different stress conditions (pH 4.5, pH 8.5 and salinity), combined stress conditions (pH 4.5 + salinity and pH 8.5 + salinity), as well as in the mixed 10 carbon media at pH 7 (all carbons), and in all 10 individual carbons, with an overall concentration of 10mM. The mean average across all 96 clones per evolved condition was then taken. Each row shows the evolutionary history of these clones, and each column indicates the condition they have been grown in. The colour of the boxes indicates the relative growth compared to the ancestor. White indicates same growth compared to the ancestor, blue a reduced growth and red an increase in growth. Stars indicate significant differences in growth compared to the ancestor (*: 0.01<p<=0.05, **: 0.001<p<=0.01, ***: p<=0.001). NAs indicate that growth rate calculations for these conditions were unreliable due to extremely small growth measurements, not exceeding 0.1 optical density (OD), and therefore have been excluded (see appendix figure D.5 and D.6 for growth curves).

4.4.5 The molecular mechanisms of specialist stress adaptations were not associated with insertion sequence (IS) movements

To understand potential molecular mechanisms of stress adaptations, one clone was selected at random from each replicate population that was exposed to acidity, alkalinity, salinity, and no stress conditions (8 populations per stress condition) as these treatments showed phenotypic adaptation based on the fitness assays. Whole genome sequencing of these clones, along with the ancestor (n=1), was then conducted to observe if there are any unique genetic changes within stress treatments. This was done by comparing mutations (SNPs/small indels) and insertion sequence (IS) movement, in both the chromosome and megaplasmid of the bipartite genome.

Overall, very few (22) genetic mutations (SNPs/small indels) were found among the 32 sequenced *R. solanacearum* clones that were not present within the ancestor (figure 4.6). These variants were spread across the two chromosomes, the 3.6 Mb chromosome and the 2.1 Mb megaplasmid (appendix table D.4 and figure 4.6). A few mutations were present in one clone only, with no suggestion of parallel adaptation within treatment. Some mutations occurred across multiple clones from all four treatments, suggesting either neutral or media/lab adaptations, including mutations in a DUF2778 domain-containing protein (RSUY_RS17095), thought to be involved in the type VI secretion system (Sibinelli-Sousa et al. 2020), and a single SNP present in an Ig-like domain containing protein (RSUY_RS21215). However, one interesting gene to note is the type IV pilus twitching motility protein (RSUY_RS03815) where four separate missense SNPs were present within four clones, three of which originated from alkaline adapted clones in addition to one salinity adapted clone. This could indicate that mutations within this gene provides increased tolerance to stress conditions, especially towards alkaline environments. Overall, few unique genetic mutations were found with strong support as a cause for stress adaptation, indicating few specialist SNP/small INDEL adaptations towards stress conditions in *R. solanacearum.*

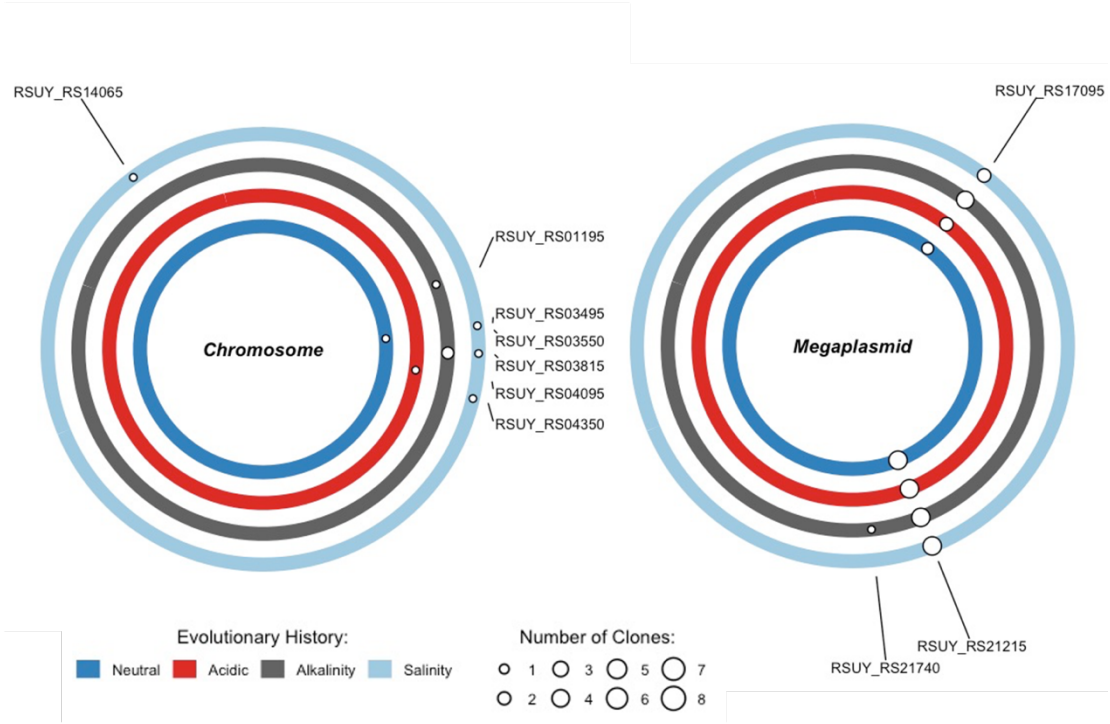


Figure 4.6: Genetic variants within adapted clones compared to the ancestor. Genetic variants found in the chromosome (left, ~3.5Mbp) and megaplasmid (right, ~2Mbp). Colour indicates the evolutionary history of the clones (N=8 per evolved condition), either from neutral (pH 7 with 0% NaCl), acidic (pH 4.5 with 0% NaCl), alkalinity (pH 8.5 with 0% NaCl) and salinity (pH 7 with 0.5% NaCl) stress conditions. Each circle represents a gene in which a genetic variant has been called in, the position of the circle shows the that genes position within the *R. solanacearum* genome, size of circles indicates the number of clones per evolutionary history that has that variant present within its genome.

In contrast, ISs appeared to be highly mobile, especially within the megaplasmid, as evidenced by 46 variable positions in the chromosome and 105 variable positions within the megaplasmid (appendix figure D.7). In most variable positions gain and loss of IS was infrequent, occurring in a few clones per evolved condition, indicative of chance IS movement with non-neutral fitness effects. However, some of the IS movement was treatment specific, indicative of potential adaptation. Especially, the movement of the IS1021 element within the megaplasmid was associated with six genomic areas in the megaplasmid, which differed between clones exposed to different stress conditions (figure 4.7). Regions 1, 2, and 3 all highlight that a few IS elements are commonly missing, compared to the ancestor, within acidic and alkaline adapted clones. Region one highlights four positions where IS1021 was absent in 2 or 3 acidic clones and 1 to 5 alkaline adapted clones but were present within all other clones (8 clones per condition). These positions are surrounded by a PhoPQ-activated pathogenicity-related protein, an endoglucanase precursor, a thioesterase superfamily protein, a TRP repeat-containing protein (*YrrB*), a Leukotokin, and hypothetical proteins, suggesting that these genes may play a role in acidic and alkaline adaptation in *R. solanacearum*. Region 2 highlights a position where IS1021 is absent in 3/8 acidic adapted clones and 4/8 alkaline adapted clones, between a cyclic di-GMP phosphodiesterase Gmr and a H-NS histone family protein. The third region shows that an IS1021 IS element has been removed in 2/8 acidic and alkaline adapted clones between a hypothetical protein and an integrase core domain protein. Four of the eight alkaline adapted clones had an IS1021 element missing between a tRNA-Pro and a filamentous hemagglutinin gene in region four, which indicates that these genes may be involved in specialized alkaline adaptation within *R. solanacearum*. Similarly, regions five had 3 positions where IS1021 or ISRso20 has been lost within neutral (2/8 to 6/8), alkaline (4/8 to 8/8) and salinity (4/8 to 6/8) adapted clones but has been retained within all 8 acidic adapted clones. These positions within the megaplasmid are near genes annotated as aminopeptidase N, type II secretion system protein F, type II secretion system protein G precursor, and hypothetical proteins. The sixth region had two IS positions (747499 and 759889) where the IS1021 element was present within the ancestor clone but missing within

4 of the stress-free clones and 5 to 6 of the alkaline adapted clones, while only missing in 2 of the salinity adapted clones and 1 clone adapted to acidic conditions. These positions have a tRNA3(Ser)-specific nuclease WapA precursor, a minor extracellular protease Epr precursor and two hypothetical proteins nearby. The other two positions in region six show an IS gain in position 928390, within 4/8 stress-free evolved clones, 5/8 acidic adapted clones, 2/8 alkaline and 1/8 salinity adapted clones, suggesting an adaptation towards stress-free and acidic conditions. However, position 931174 shows an IS loss within 1 stress free clone, 3 acidic adapted clones, 2 alkaline clones and 0 salinity adapted clones (out of 8), suggesting a potential region associated with both acidic and alkaline adaptation. Both IS positions are within hypothetical proteins. Overall, this research suggests that there is a large amount of IS movement within *R. solanacearum,* especially within the megaplasmid, with some IS movements being specific to stress treatments and hence potentially adaptive.

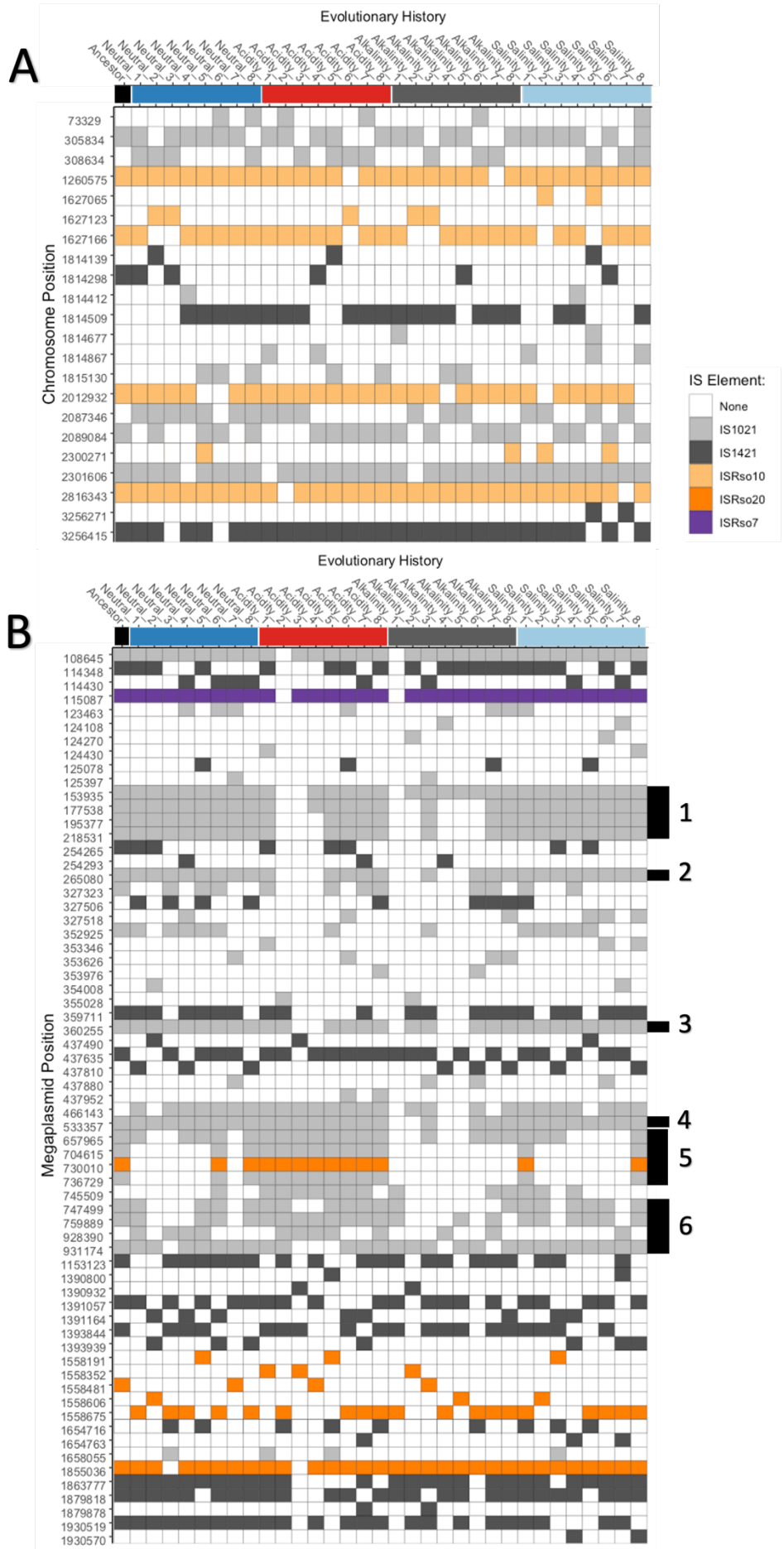

A
Evolutionary History
Ancestor
Neutral_1
Neutral_2
Neutral_3
Neutral_4
Neutral_5
Neutral_6
Neutral_7
Neutral_8
Acidity_1
Acidity_2
Acidity_3
Acidity_4
Acidity_5
Acidity_6
Acidity_7
Acidity_8
Alkalinity_1
Alkalinity_2
Alkalinity_3
Alkalinity_4
Alkalinity_5
Alkalinity_6
Alkalinity_7
Alkalinity_8
Salinity_1
Salinity_2
Salinity_3
Salinity_4
Salinity_5
Salinity_6
Salinity_7
Salinity_8
Chromosome Position
73329
305834
308634
1260575
1627065
1627123
1627166
1814139
1814298
1814412
1814509
1814677
1814867
1815130
2012932
2087346
2089084
2300271
2301606
2816343
3256271
3256415
IS Element:
None
IS1021
IS1421
ISRso10
ISRso20
ISRso7
B
Evolutionary History
Megaplasmid Position
108645
114348
114430
115087
123463
124108
124270
124430
125078
125397
153935
177538
195377
218531
254265
254293
265080
327323
327506
327518
352925
353346
353626
353976
354008
355028
359711
360255
437490
437635
437810
437880
437952
466143
533357
657965
704615
730010
736729
745509
747499
759889
928390
931174
1153123
1390800
1390932
1391057
1391164
1393844
1393939
1558191
1558352
1558481
1558606
1558675
1654716
1654763
1658055
1855036
1863777
1879818
1879878
1930519
1930570
1
2
3
4
5
6

Figure 4.7: Insertion sequence (IS) movement contributes to stress adaptation. Presence (coloured) and absence (white) of IS elements at certain genomic positions across the (A) chromosome and (B) megaplasmid of *Ralstonia solanacearum*. Specific colours indicate the type of IS present and only IS elements whose movement occurred within more than one clone are shown for easier interpretation, for all IS element movement see appendix figure D.7. Evolutionary history indicates the stress conditions in which the clone was exposed to during the selection experiment (n=8 per stress condition) plus the ancestor clone. Black numbered blocks along the right-hand side indicate the regions of interest which differ between different evolutionary history groups. Region 1 includes positions near a PhoPQ-activated pathogenicity-related protein, an endoglucanase precursor, a thioesterase superfamily protein, a TRP repeat-containing protein (YrrB), a leukotokin, and hypothetical proteins. Region 2 involves a cyclic di-GMP phosphodiesterase Gmr and a H-NS histone family protein. Region 3 a hypothetical protein and integrase core domain protein. Region 4 a tRNA-Pro and a filamentous hemagglutinin.Region 5 an aminopeptidase N, a type II secretion system protein F, a type II secretion system protein G precursor and hypothetical proteins. Finally, region 6 includes a putative deoxyribonuclease (RhsB), tRNA3(Ser)-specific nuclease WapA precursor, minor extracellular protease Epr precursor and hypothetical proteins. IS movement positions along the genome can be seen in appendix figure D.8.

.

4.4.6 Abiotic stress tolerance causes diversification of *R. solanacearum* clones

Finally, comparisons on how exposure to these stress conditions affected diversification of *R. solanacearum* was conducted by comparing growth of evolved clones (acid, alkaline, saline and stress-free exposed clones), along with the ancestor, in different stress conditions (pH 4.5, pH 8.5, 0.5% NaCl). Area under the curve (AUC) after 4 days was used as a proxy for growth and was divided by their growth (AUC) on the ten carbon mixed media (pH7) to control for variation amongst the clones due to media preference rather than the presence of the stress. A principal component analysis (PCA) was then conducted, where all clones from the single stress evolutionary history groups along with the ancestor and stress-free (pH7 with 0% NaCl) control (figure 4.8A) were compared. Results showed that principal component (PC) 1 explains most of the variation within the dataset (66%), and PC2 18%. PC1 separated isolates with negative values indicating increased growth on all conditions, while negative values along PC2 showed higher growth on saline and alkaline conditions and higher PC2 values corresponded to high acidic stress tolerance (figure 4.8A and appendix figure D.9). This is similar to the trends observed among the environmental isolates, where higher PC2 values indicated higher tolerance to acidic stress (pH 4.5) and high salinity concentrations (2% and 1% NaCl), while negative values indicate higher tolerance to alkaline stress (pH 9 and pH 10) and low salinity concentrations (0.5% NaCl) (figure 4.2B). Overall, there were significant differences between clones from different evolutionary history (figure 4.8A, PERMANOVA: $F_{4,395}$= 29, R^2= 0.23, p=0.001), post hoc analysis showed significant differences between all groups (figure 4.8A, Pillai: p=0.0011), apart from between the ancestor clones and clones evolved within neutral conditions (pH 7) (figure 4.8A, Pillai: p=0.386), suggesting that adaptation to environmental stresses does drive an increased diversity in abiotic stress tolerance within the *R. solanacearum* population as a whole as clones in different treatments adapt differently. Furthermore, dispersion around the centroid was significantly different between the evolved groups (figure 4.8A, ANOVA: $F_{4,395}$= 11.6, p<0.001), with all evolved clones having significant higher dispersion compared to the ancestor (TUKEY: p < 0.0001 – 0.002) and clones evolved in pH 8.5 having significant higher

dispersion compared to clones evolved in pH 7 with and without salinity (TUKEY: p = 0.002 without salinity and p = 0.0003 with salinity)

To confirm this the mean pairwise distance (as calculated using Euclidean distances) of each clone from all other clones was calculated (figure 4.8B). The results showed that the mean pairwise distance, and therefore diversity, was different between the different evolutionary history groups (figure 4.8B, Kruskal-Wallis: X^2_4= 65, p<0.0001). The least diverse group was the ancestral clones with a significantly lower mean pairwise distance compared to all other groups (p<0.0001 for all). This is expected as these clones do not come from 8 replicate populations unlike the other treatments. Furthermore, evolved clones from no stress control conditions (pH7 0% NaCl) were less diverse compared the clones evolved in stress conditions, with significantly lower mean pairwise distance compared to acidic (p=0.0001), alkaline (p<0.0001), and salinity (p=0.0002) evolved clones. Out of the stress evolved clones, the most diverse groups were those evolved in alkaline conditions followed closely by salinity evolved clones and acidic evolved clones. However, these differences were not significantly different. This suggests that exposure to stress conditions does increase the pathogens diversity, with alkaline stress increasing diversity the most, suggesting that this diversification of UK isolates may be caused by past exposure to different stress conditions.

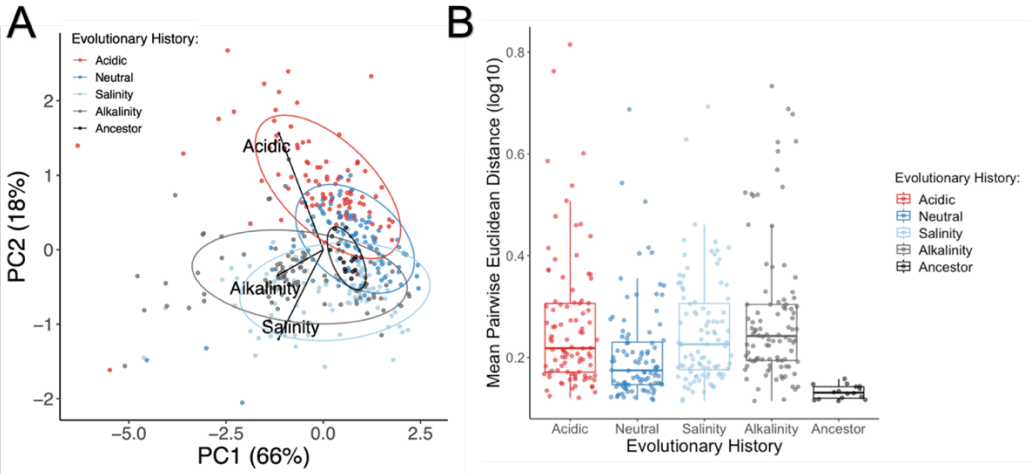


Figure 4.8: Abiotic stress tolerance can cause diversification of *R. solanacearum*. (A) PCA plot showing that evolving clones to environmental stress conditions increases *R. solanacearum* diversity. Each point represents a clone, coloured by their evolutionary history in the selection experiment. Red is acidic (pH 4.5 0% NaCl), blue neutral (pH 7 0% NaCl), light blue salinity (pH 7 0.5% NaCl), grey alkalinity (pH 8.5 0% NaCl), and black the ancestor (YO336). Together the PCA explains 81% of the

total variation within the dataset, with PC1 explaining 66% and PC2 15%. Ellipses show the 90% confidence intervals around the centroid for each evolutionary history. (B) Each clone's mean pairwise distance from all other clones (Euclidean distance) per evolutionary history showing increased diversity in stress adapted clones. Each point shows an individual clone's mean pairwise Euclidean distance (log10). Boxplot lines show the median mean pairwise Euclidean distance (log10) of all clones in each evolutionary history group, with the box indicating the interquartile range and whiskers showing the 95% quantile (n = 96 per evolutionary group apart from ancestor where n = 16).

4.5 Discussion

The aim of this study was to determine how *R. solanacearum* diversifies and evolves when exposed to different stress conditions using both comparative analysis of environmental samples and experimental evolution techniques. Analysis of a population of *R. solanacearum* (n=182), spanning 30 years of sampling (1992-2019) within the UK, revealed great diversity in abiotic stress tolerance (pH 4.5, pH9, pH10, 0.5% NaCl, 1% NaCl and 2% NaCl) (figure 4.2) despite UK *R. solanacearum's* clonal nature (chapter 3). Overall, specialist acidic and alkaline stress adaptation was observed in the selection experiment. However, negative trait correlations existed between these two stresses, resulting in a reduced growth rate in alkaline environments if *R. solanacearum* had been adapted in acidic conditions. The evolutionary experiment also failed to generate stress-tolerant generalists, as exposure to a combination of two stress conditions constrained *R. solanacearum* adaptation, resulting in extinctions when extreme pH was combined with 1% salinity. The lack of stress generalists arising could be due to the high metabolic costs observed linked to all stress conditions or a potential trade-off in *R. solanacearum* ability to tolerate stress and utilise resources simultaneously. Together, these results support the hypothesis that exposure to acidity or alkalinity stress can drive *R. solanacearum* into specialist niches.

Throughout the selection experiment *R. solanacearum* was less affected by acidic and neutral pH compared to alkaline conditions. This is unsurprising as previous research indicates that bacterial wilt disease incidence can be high in lower pH conditions suggesting that the *Ralstonia solanacearum* species complex (RSSC) pathogen is well adapted to these conditions within the environment (Li et al. 2017, Wang et al. 2020). However, it also contradicts other research which states that RSSC is favoured by alkaline conditions (Álvarez et al. 2010, Bhanwar 2022).

The conflicting results within the literature could be due to bacteria changing their pH preference within different environments, for example rich lab media was not used within this experiment unlike in Bhanwar 2022, instead minimal mixed carbon media (10mM) was provided which could be more realistic to the natural resources available to *R. solanacearum* within the environment. Another explanation could be that different *R. solanacearum* strains, or different species within the RSSC, have different pH preferences. However, both Bhanwar 2022 and Li et al. 2017 investigated *R. pseudosolanacearum* phylotype I (race 1 biovar 3) strains and yet found contradicting results, suggesting more evidence for the former argument that nutrient availability is changing pH preference.

The selection experiment also showed that salinity had a larger effect on growth compared to extreme pH conditions, persistently reducing *R. solanacearum* densities regardless of pH condition. Cell densities were a lot lower in 1% salinity, even becoming extinct when combined with extreme pH, while for neutral pH (pH 7) it took until about halfway through of the selection experiment (20 days) for *R. solanacearum* to recover and reach to similar densities as in the other stress conditions. It also took longer to become extinct in pH 4.5 compared to pH 8.5 suggesting, again, that *R. solanacearum* is better adapted to acidic conditions. Cell densities were also lower in the 0.5% NaCl combined stress conditions compared to 0.5% salinity or extreme pH alone, suggesting that adaptation to multiple stresses is more challenging than adaptation to a single stressor. Overall, these results suggest that salinity and alkalinity are extremely stressful for *R. solanacearum,* especially within minimal media, while acidic pH is less stressful in comparison.

Fitness assays revealed evolution of specialists to the two extreme pH stressors (acidic and alkaline). Adaptation to acidic pH did not increase tolerance to alkaline pH and vice versa, suggesting that specialist adaptation occurred with no positive correlations between the two traits. In fact, a reduction in growth within alkaline conditions was observed if the pathogen was previously adapted to acidic conditions, suggestive of a negative trait correlation between the two traits. Research on *Escherichia coli* indicates that exposure to acid pH can in fact induce alkaline sensitivity by creating high internal Na+ (Rowbury 1997), also suggesting that these traits are not positively correlated with one another. Furthermore, an

increased media adaptation within the no stress control clones was also seen to increase their tolerance to acidic conditions, suggesting that a positive interaction between media adaptation and acid stress tolerance exists.

Extreme pH, like many other environmental signals, has strong effects on numerous pathways (Olson 1993). Changes in external pH can bring about alterations in internal cytoplasmic pH, changing the concentration of ions within the membrane, as well as changing membrane potential (Olson 1993, Saito and Kobayashi 2003). An enzyme's activity is also dependent on the binding of protons to its specific sites and consequently, enzymatic activity may decrease with cytoplasmic alkalization or acidification (Saito and Kobayashi 2003). Therefore, in response to changes in pH bacteria have many different adaptive strategies. For example, changing the lipid composition of their cytoplasmic membrane to maintain functionality (Siliakus et al. 2017), such as producing periplasmic proteins (Saito and Kobayashi 2003), or producing acid shock proteins (Olson 1993). Another adaptation towards extreme pH is to prevent the build-up of ions in the cytoplasm via ATP driven exclusion systems, outer membrane porins or sodium/proton antiporters (Saito and Kobayashi 2003). For example, *Streptococcus mutans* relies on its F-ATPase to protect itself from acidification by pumping protons out of the cells (Quivey et al. 2000) and extrusion systems for potassium ions have also been found in *Escherichia coli* and *Enterococci* (Saito and Kobayashi 2003). Research on evolution to long-term acid stress in a closely related bacterial species, *Ralstonia pseudosolanacearum,* has revealed that inactivation of the *PhcA* transcriptional regulator gene resulted in increased growth (Liu et al. 2022). The *PhcA* gene is involved in global regulation of virulence and metabolism in early stages of infection among *R. solanacearum* strains (Genin 2010). Overall, this implies that multiple adaptive strategies to tolerate extreme pH are present for *R. solanacearum,* however this research has shown that acidic or alkalinity adaptation occurs with no cross tolerance towards the other extreme pH condition.

No clear adaptation to salinity was observed in the fitness assays. Salinity is highly stressful to bacteria, with high amounts in the soil creating toxic intermediate products which can weaken bacterial cell metabolism (Zhang et al. 2021). Salt stress can also result in the build-up of cations/anions within the cytoplasm, influencing

the osmotic balance of the cell, and therefore can affect cellular processes such as catalytic activity (Gregory and Boyd 2021, Zhang et al. 2021). Research on multiple bacteria organisms show that adaptation towards high salinity results in differential expression of multiple genes (Chen et al. 2017, Elabed et al. 2019, Zhang et al. 2021), suggesting that adaptation is complex and relies on multiple gene products, pathways, and gene interactions (Elabed et al. 2019, Zhang et al. 2021). Other research has also found that genes involved in cell motility, including flagella genes, was down regulated and genes involved in energy production and conversion were found to be up regulated in high saline environments, suggesting that salinity tolerance is costly in terms of energy and a decreased cell motility allows more energy to be conserved for osmoprotection (Chen et al. 2017). Both reasons, high complexity of salinity adaptation (multi gene regulation) and high energy costs of adaptation, could explain why no clear salinity adaption in *R. solanacearum* was seen in this evolution experiment.

Many of the differentially expressed genes in bacteria within saline environments are involved in proton transmembrane transport and metabolic pathways (Zhang et al. 2021). Sodium efflux and hydrogen or potassium ion uptake across the membrane can help bacteria maintain osmotic balance in high salinity environments. This is accomplished through a number of different transport systems that vary in kinetics, energy coupling and regulation (Chen et al. 2017, Elabed et al. 2019). For example, an increase in expression of H+ transport T3SS ATPase were seen, along with a significant increase in expression of T3SS proteins and 2 cytotoxin secretion factor exoenzymes, in salt adapted *Pseudomonas aeruginosa* (Elabed et al. 2019). Bacteria can also offset osmotic stress by accumulating compatible solutes or osmolytes which maintain cell metabolism (Chen et al. 2017, Zhang et al. 2021). Wang *et al*., found that knocking down the *R. solanacearum* AVT05_RS03545 gene (homologous to the Rsp1238 protein in GMI1000) reduced growth in 1% NaCl but not extreme pH, suggesting a role in salinity tolerance and not pH adaptation. AVTO5_RS03545 is predicted to be a polyisoprenoid-binding protein, YceI, which is shown to be involved in cellular stress response in *Helicobacter pylori* and *Escherichia coli* (Wang et al. 2021). Overall, this suggests that there are a multitude of potential adaptive strategies for *R.*

solanacearum salinity evolution, however no clear adaptation to salinity stress was observed within this study.

This research also revealed that generalist adaptation did not occur when *R. solanacearum* was exposed to double stress conditions as acidic and alkaline adaptation were only observed when *R. solanacearum* was exposed to the pH stress alone. This result indicates that the presence of salinity stress constrained *R. solanacearum* adaptation to both extreme pH conditions, and potentially suggests that there could be negative correlations between the two stresses. Other research, in *Escherichia coli,* has shown that the presence of high NaCl concentration increases bacteria susceptibility to high or low pH, and is thought to be due to the induction of PhoE, an outer membrane proton pore (Rowbury 1997). Trade-offs between two different stresses have been seen in *R. solanacearum* before*,* but with temperature affecting the pathogens' ability to grow in extreme pH conditions (Wang et al. 2020). This indicates that *R. solanacearum* has different mechanisms of stress tolerance for pH and salinity, and that these mechanisms possibly interact negatively with one other. In the environment it is quite likely stressors occur together, so this negative interaction can limit *R. solanacearum*'s ability to specialise and adapt to one stressor. Another explanation could be that the combination of adapting to both stressors results in an even further reduction in their metabolic capacity compared to adaptation to the pH stress alone. Alternatively, the double stress conditions resulted in smaller population sizes throughout the selection experiment (figure 4.4) which could result in not enough bacterial replication for adaptation to occur.

Fitness assays also revealed that exposure to all stresses reduced *R. solanacearum's* metabolic capacity*.* Research has shown that acid or alkaline responses can be affected largely by nutrients and metabolites (Rowbury 1997), suggesting that stress response and metabolism are closely linked. However, I have observed that reduction in metabolism was a lot more extreme in alkaline conditions. This could explain why *R. solanacearum* is naturally better adapted to acidic conditions, with a lower fitness cost being associated with adaptation to this pH. *R. solanacearum* is naturally more likely and persistently exposed to acidic conditions compared to alkaline conditions, such as in the soil where nitrogen

fertilizer increases soil acidification (Li et al. 2017), and in plant xylems which can be slightly acidic, with pH dropping even further when plants are stressed (Secchi and Zwieniecki 2016). Therefore, trade-offs in metabolism when adapting to acidic conditions could have higher consequences for *R. solanacearum's* survival and competitive ability compared to alkaline conditions.

Although no clear adaptation to salinity was seen within this experiment a trade-off between metabolism and exposure to saline conditions was also observed. *Ralstonia solanacearum* clones exposed to 0.5% NaCl conditions over the course of the selection experiment had clear reduction in growth in multiple different carbon resources. This suggests that although adaptation to salinity was not clear, growing in the presence of NaCl negatively impacted *R. solanacearum's* ability to utilise most resources. This could be due to slight adaptation to salinity occurring, as indicated by a non-significant increase in growth in salinity conditions compared to the ancestor, and resulted in a reduction in uptake of nutrients, including NaCl and all other resources, or a conflicting use of resources for metabolism and compatible solutes resulting in lower growth (Gregory and Boyd 2021). Compatible solutes (e.g. glutamate, glutamine, trehalose, proline, glycerol, ectoine and glycine betaine to name a few) are low molecular weight compounds that can be synthesised and/or taken from the environment by bacteria to maintain cellular turgidity and electrolyte concentration to sustain osmotic equilibrium (Chen et al. 2017, Elabed et al. 2019, Gregory and Boyd 2021). These compounds are also in different metabolic pathways in the cell and can be used as carbon sources, therefore the use of these compounds as osmolytes in some species can be costly (Gregory and Boyd 2021). Transporters, such as the *Vibrionaceae* BCCT transporter, can also be involved in both uptake of substrates for metabolism as well as the uptake of compatible solutes (Gregory and Boyd 2021), suggesting a costly trade-off between the two. Elabed *et al.* also found that *Pseudomonas aeruginosa* adaptation to salinity not only depended on the salt concentration but also on the composition of the growth media (Elabed et al. 2019), supporting this theory. As we used minimal growth media (10mM carbon solution) there could be higher costs imposed on growth when adapting to abiotic stress such as salinity compared to rich media. Overall, all stress conditions, acidic, alkalinity, and salinity stress, have

large impacts on *R. solanacearum* metabolic capacity, indicating a trade-off between abiotic stress tolerance and metabolism.

To study the mechanisms of stress adaptation the DNA of one clone from each of the 8 replicate evolved populations for neutral, acidic, alkaline and salinity conditions was sequenced, and genomic variation (SNPs and small INDELs) compared. However, only one candidate gene responsible for stress adaptation was observed. This candidate gene possessed multiple missense SNPs across four different clones exposed to alkalinity (n=3) and salinity (n=1) stress conditions and was annotated as a type IV pilus twitching motility protein. This gene has previously been linked to virulence of this pathogen and therefore could have consequences on this pathogen's ability to infect hosts (Genin and Denny 2012). Therefore, changes within this type IV pilus gene could be to preserve energy, otherwise used for virulence, to increase stress tolerance which is known to be costly in terms of energy requirements (Chen et al. 2017).

Within *R. solanacearum* insertion sequences (IS) have been seen to insert and disrupt genes, such as type III effectors and the global virulence regulator, PhcA, resulting in changes within virulence and phenotypic plasticity (Jeong and Timmis 2000, Gonçalves et al. 2020). Furthermore, IS movement has been seen as an evolutionary response within *R. solanacearum* towards antimicrobial plant allelochemicals called isothiocyanates (ICTs) (Alderley et al. 2022). Therefore, IS movement was explored to see if they play a role in pH and salinity stress adaptation. There was a lot more IS movement compared to genetic mutations in response to stress conditions, which was mainly due to the IS element IS1021 moving within the megaplasmid (figure 4.7) and is supported by other research on *R. solanacearum* that has found that IS are more frequently found on the megaplasmid compared to the chromosome (Greenrod et al. 2022, Alderley et al. 2022). Acid and alkaline stress adaptation was associated with IS movement within or near virulence genes, such as the endoglucanase precursor, type II secretion proteins and the PhoPQ pathogenicity protein. PhoPQ has also been shown to be activated under acidic/alkaline conditions in plant bacterial pathogen *Erwinia chrysanthemi* (Venkatesh et al. 2006), suggesting a role in extreme pH adaptation within bacteria. IS within genes involved in biofilm formation, aggregation and

motility were also associated with both acid and alkaline stress tolerance (cyclic di-GMP phosphodiesterase Gmr) and specialized alkaline stress tolerance (filamentous hemagglutinin gene). Multiple enzymes were also seen to have IS movement nearby (minor extracellular protease Epr precursor, tRNA3(Ser)-specific nuclease WapA precursor, thioesterase superfamily protein), potentially providing increased metabolism or tolerance to stresses within the environment. IS movement away from an integrase core domain protein was seen in acidic and alkalinity evolved clones, suggesting that integration of DNA into the genome could also be in response to environmental stresses. Furthermore, multiple IS elements nearby hypothetical proteins were found associated with stress tolerance, highlighting the lack of understanding we have of gene function within *R.* solanacearum. Previous research has discovered that acid adaptation in a closely related bacterial species, *R. pseudosolanacearum,* was caused by the loss of function of the global regulator *PhcA* gene (Liu et al. 2022). However, mutations or IS movement near this gene was not observed in this study. Overall, whole genome sequencing has revealed that IS movement is the main genetic mechanism underpinning stress adaptation within this experiment, potentially suggesting a role in gene disruption or gene regulation for stress adaptation in *R. solanacearum*.

Finally, to see if *R. solanacearum* diversifies into specialist niches within the environment, comparisons in abiotic stress tolerance (acidic, alkaline, and saline environments) were made across our evolved clones that were adapted to extreme pH, salinity, and stress-free conditions, along with the ancestor (figure 4.8). This revealed diversification of *R. solanacearum* in their ability to tolerate acidic and alkaline stress conditions (figure 4.8A), mirroring the *R. solanacearum* population data (figure 4.2B). Higher diversity among clones exposed to one of the stress conditions was also observed compared to the stress-free adapted clones (figure 4.8B). Together, this suggests that exposure to different stresses within the environment is responsible for *R. solanacearum* diversification with negative trait correlations between these traits being a potential driver.

In conclusion, a comparative study on a *R. solanacearum* population revealed high abiotic stress tolerance diversity, with diversification based on acidic or alkaline tolerance and overall stress tolerance. To investigate how *R.*

solanacearum stress tolerant specialists and generalists arise within the environment, this pathogen was exposed to a variety of stress conditions (pH 4.5, pH 8.5, 0.5% salinity and 1% salinity), and a combination of stresses (pH 4.5 and pH 8.5 with salinity), in an evolution experiment. Predictions were made that specialists would arise when exposed to one stress if positive trait correlations were not present (figure 4.1), however in the presence of more than one stress the bacteria would adapt to become generalists with intermediate fitness in either stress. This experiment discovered that specialist adaptation occurred to single stress conditions, with insertion sequence (IS) movement driving this adaptation. On the other hand, generalists did not arise, with no adaptation to either stress occurring in the combined stress conditions. This lack of adaptation in the combined salinity and pH stress conditions suggests that the cumulative stress was difficult for *R. solanacearum* to adapt to and caused extinctions events when extreme pH was combined with 1% NaCl. Furthermore, adapting to all stresses had great metabolic consequences, indicating a trade-off between stress tolerance and resource utilisation. Overall, this highlights that the presence of environmental stressors can play a large role in altering *R. solanacearum* diversity, potentially affecting pathogen distribution and prevalence in environmental reservoirs.

5 Chapter 5: General Discussion

5.1 Overview

The *Ralstonia solanacearum* species complex (RSSC) is a globally distributed bacterial plant pathogen and the causative agent of bacterial wilt disease. This pathogen is thought to infect over 175 species of plants, including many important crops, having huge impacts on agriculture. A review of the literature revealed that while genetic and phenotypic diversity of this pathogen species has been explored, especially when concerning pathogenicity, phenotypic diversity on a broad range of ecologically relevant phenotypes was lacking. A better understanding of RSSC ecological diversity can therefore improve knowledge of the epidemiology, such as the evolution and dispersal, of this plant pathogen species. Hence, the aim of this thesis was to first explore the ecological diversity among the RSSC, then determine potential causes of this trait variation, and finally link phenotypic variation with genomic data, revealing genetic mechanisms underpinning certain traits. This chapter provides an overview of the results of this thesis and discusses them in the context of the broader knowledge of the field.

This research has progressed the understanding of phenotypic diversity among the bacterial plant pathogen RSSC at the global level (chapter 2), population level (chapter 3) and causally within the lab (chapter 4) (figure 5.1). My results suggest that the local environment selects for similar ecological differences within both RSSC species investigated within this thesis, *R. pseudosolanacearum* and *R. solanacearum*, aligning with the fact that both species share remarkably similar life cycles (chapter 2). Furthermore, exploring the UK population of *Ralstonia solanacearum* (chapter 3) highlighted a lack of genetic variation despite significant phenotypic differences observed (figure 5.1). However, diversification was observed over time, indicative of a relatively recent introduction of this population to the UK. Finally, evolving a single UK isolate, positioned in the 'middle' of the trait space of the whole RSSC collection (figure 5.1), to different abiotic environmental stress conditions within the laboratory revealed that exposure to stresses increases RSSC diversity. This is likely driven by the negative trait correlations observed between the different stresses and between stress adaptation and metabolism (chapter 4).,

resulting in contrasting evolutionary trajectories Overall, the RSSC is extremely phenotypically diverse, and this thesis provides a thorough investigation of the ecological diversity of this plant pathogen.

This study has also increased knowledge of the genetic mechanisms underpinning traits within the RSSC. Accessory genome variation was shown to explain ecotype differences within the RSSC in chapter 2, potentially suggesting that horizontal gene transfer within the local environment drives RSSC diversity. Furthermore, a gene and three gene regions were found associated with cold tolerance and rifampicin resistance within the RSSC respectively (chapter 2). Genomic analysis of the UK *R. solanacearum* population revealed few genetic differences despite three decades of sampling, suggesting a role for other genetic mechanisms not captured within this chapter involved in initial RSSC adaptation (chapter 3). Evolving one UK isolate within the laboratory revealed that insertion sequence (IS) movement in the *R. solanacearum* genome can drive adaptation to environmental stress conditions (chapter 4), potentially explaining the lack of genetic variation observed at the SNP level within the UK population in chapter 3. Therefore, this thesis provides a deeper understanding of the genetic mechanisms underpinning important phenotypic traits and how adaptation of RSSC strains initially occurs within the environment.

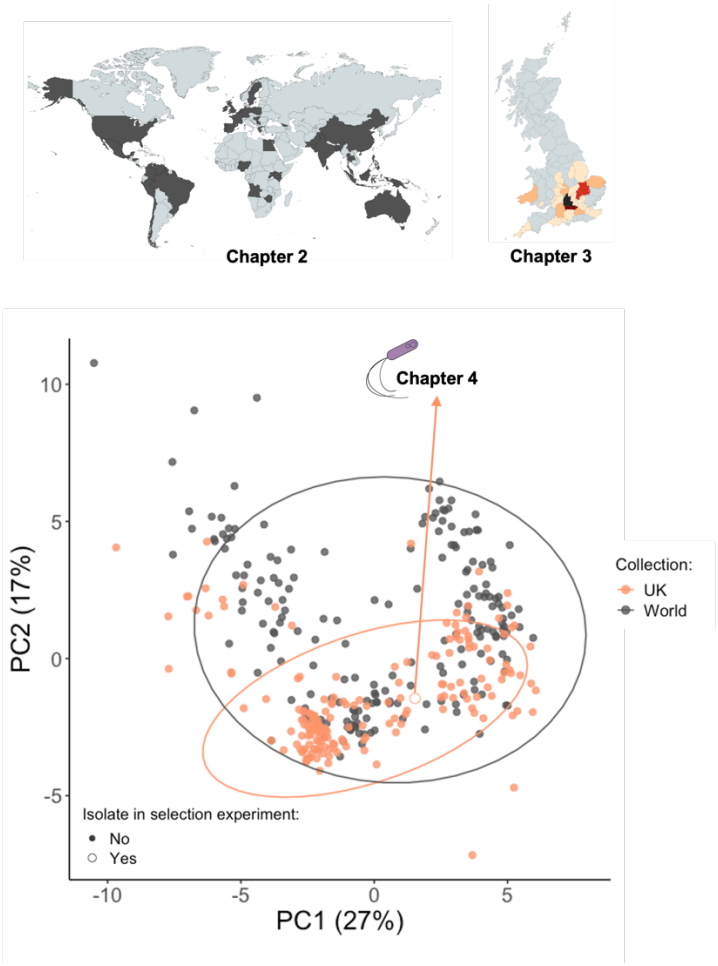


Figure 5.1: Overview of phenotypic diversity among both Ralstonia solanacearum species complex (RSSC) collections; world and UK collections. All 376 isolates from both world (n=194) and UK (n=182) collections phenotypic trait values (n=46) were converted into a principal component analysis (PCA) to visualise on a 2D graph. Each point represents one RSSC isolate, colour indicating which collection they are from (world collection in chapter 2 or UK collection in chapter 3) and shape indicates if it was the chosen isolate conducted in the evolutionary experiment in chapter 4. Ellipse show a 90% confidence interval around the centroid for each bacterial collection.

5.2 How phenotypically diverse is the *Ralstonia solanacearum* species complex (RSSC)?

The *Ralstonia solanacearum* species complex (RSSC) isolates are phenotypically and genetically diverse (Fegan et al. 1998), comprising of three separate species (Safni et al. 2014) and consisting of strains that can infect a wide range of host plants (Hayward 1991). Previous research has tried to determine the genetic (Fegan and Prior 2005, Castillo and Greenberg 2007, Geng et al. 2022, Sharma et al. 2022) and phenotypic (Cellier and Prior 2010) diversity of this pathogen species, especially when in relation to pathogenicity (Bocsanczy et al.

2017) and metabolism (Stevens and Van Elsas 2010, Cruz et al. 2012, Caruso et al. 2017). However, knowledge of RSSC phenotypic diversity using a broad range of ecologically relevant phenotypes is lacking.

The first two data chapters (chapter 2 and 3) within this thesis makes use of a collection of bacterial plant pathogens collated by FERA science ltd, Sand Hutton, York (n=379). High throughput phenotyping of 46 ecologically relevant phenotypic traits, spanning virulence, stress tolerance, and metabolism, was conducted on this bacterial collection to explore the ecological diversity of the RSSC. Chapter 2 explores the phenotypic diversity of RSSC across the world, spanning two separate species, *R. pseudosolanacearum* and *R*. solanacearum, revealing high ecological diversity among this bacterial pathogen. High diversity is usually correlated with increased survival, due to three factors: a greater tolerance within a changing environment, decreased competition among individuals, via the ability to occupy new niches, and increased evolvability, by increasing the chances of fixing advantageous genotypes (Horner-Devine et al. 2004, Cordero and Polz 2014, Bajic and Sanchez 2020). Therefore, the high ecological diversity discovered in this thesis reflects the worldwide distribution and complex life cycle RSSC strains have (Hayward 1991, Genin 2010, Bragard et al. 2019, EPPO 2022).

RSSC pathology has been shown to not correspond fully with its genetic separation into species or phylotypes (Lebeau et al. 2011) and this research highlights that RSSC ecological diversity also does not adhere to these phylogenetic boundaries. Chapter 2 found that five phenotypically distinct 'ecotypes' exist within the RSSC which do not adhere to the species separation. This suggests that the shared life cycle within both species is selecting for the same trait differences within both RSSC species. The five distinct phenotypic groups or 'ecotypes' differed in their trait specificity, either being oligotrophic, heat tolerant, cold tolerant, antibiotic-resistant/biofilm producer, and metabolically efficient. This reveals important habitat differentiators that are driving RSSC trait diversity and suggests that ecotype separation among both RSSC species could be causing this large diversity of RSSC strains rather than high diversity across all traits.

Next, chapter 3 explored the phenotypic diversity of a single population of one of the RSSC species, *Ralstonia solanacearum,* within the UK. This chapter used a

collection recently introduced UK *R. solanacearum* strain spanning three decades from the first recorded outbreak in 1992 to 2019. This revealed three distinct groups that differ in their growth in nutrient-limited conditions, and overall trait generalism and specialism. Two traits, antibiotic resistance and biofilm production, were also seen to be increasing within this population with time. This reveals important selective forces within the UK environment driving *R. solanacearum* adaptation. This thesis has also shown that UK *R. solanacearum* is less diverse than the worldwide collection (figure 5.1). This is expected as we are comparing a population to a worldwide collection. Furthermore, high diversity of bacteria has been linked to a larger habitat variability (McArthur et al. 1988) and the UK population of *R. solanacearum* should exist in a lot less variable habitat compared to the world collection of RSSC. UK *R. solanacearum* belong to the cold-tolerant strain of *Ralstonia solanacearum* (race 3 biovar 2 or phylotype IIB sequevar 1), the causative agent of potato brown rot in Europe (Fegan and Prior 2005, Safni et al. 2014). Therefore, it is thought to have a smaller host range and has been recently disseminated worldwide by the potato trade (Caruso et al. 2017) which would also result in reduced diversity due to the recent population expansion. However, a surprisingly large amount of phenotypic diversity was observed within this single population (figure 5.1) which can be explained by RSSC generalist nature (Hayward 1991, Genin and Boucher 2002). Furthermore, strains within the RSSC have large genomes with complex regulatory systems allowing large phenotypic differences despite high genetic similarities (Schell 2000, Cellier and Prior 2010, Genin and Denny 2012). Overall, this research has progressed the understanding of RSSC ecological diversity at the global and population level.

5.3 What is driving trait diversity among the *Ralstonia solanacearum* species complex (RSSC)?

Large ecological diversity was observed among the RSSC plant pathogen, and this thesis also tried to explore potential drivers of this trait diversity. Comparative analysis revealed that small amounts of trait diversity can be explained by the host, location, and time from which isolates were sampled, suggesting that

life history characteristics of bacteria isolates can drive diversification within the RSSC. Chapter 2 revealed that RSSC phenotypic diversity can be partly explained by the host and continent they were sampled from. However, chapter 3 revealed that at the population level the isolation source (environmental and crop host) and location cannot explain trait diversity. The time in which an isolate was sampled does drive some trait variation within a single population of *R. solanacearum* with increased diversification over time. This is to be expected as the UK population of *R. solanacearum* was recently introduced to the country (Elphinstone and Matthews-Berry 2017) and microorganisms increase in diversity when new niche opportunities open (Rainey and Travisano 1998). This suggests that different populations of RSSC within different countries and host plants may have distinctive selective forces driving RSSC trait differences. Overall, this suggests that at smaller scales (population level) increased diversification is occurring with time and host or location effects cannot be observed. However, at a global scale, the continent and host from which isolates were collected can explain trait diversification but these only explain small amounts of variation suggesting that other factors are also at play.

Trait correlations were also explored in chapters 3 and 4 as an explanation of the phenotypic diversification of *Ralstonia solanacearum* isolates. In chapter 3 I discovered that trait correlations differed between the three phenotypically distinct ecotypes identified within this study. Therefore, trait correlations were explored in more detail by conducting an evolution experiment on one UK *R. solanacearum* isolate in chapter 4. Experimental evolution is a powerful technique used to explore how natural populations may evolve and adapt to environmental conditions (Phillips and Burke 2021). This revealed that adapting to environmental stresses is limited by trait correlations and that negative trait correlations exist between extreme pH and salinity stress conditions. However, extreme pH and salinity were seen positively correlated among UK *R. solanacearum* within all three ecotypes in chapter 3. Together these results suggest that the two stresses, extreme pH and salinity, are both present within the same environment in the UK, causing a positive trait correlation between the two traits as observed in chapter 3. In support of this, salinity has been seen to coincide with high pH stress within the soil and rivers

(Sardinha et al. 2003, Jiang et al. 2022) with sewage and agricultural run-off causing increased pH and saline levels within these environments (Rowbury 1997, Shrivastava and Kumar 2015, Zhang et al. 2021). Furthermore, no adaptation to the combined stress conditions (pH and salinity) was observed in chapter 4, however, UK isolates were seen to be equally adapted to both conditions in chapter 3. This implies that *R. solanacearum* can adapt to both stress conditions but that this just did not occur within the lab. A few explanations for these contradicting results could also be that the 40 days within the laboratory experiment in chapter 4 may have not been enough time to observe this adaptation. Another explanation could be that fluctuations within the environment, which I did not explore in my evolutionary experiment, can select for adaptation to both stress conditions. Within fluctuating environments, rather than having to adapt to two stresses simultaneously which can be too stressful for survival, bacteria can adapt to one stress before the environment switches and then accumulate stress responses to the other environment separately. Additionally, trade-offs were observed between stress tolerance traits and metabolism within chapter 4 when minimal media was used, while rich media was used for stress tolerance traits in the high-throughput phenotyping experiment in chapter 3. Therefore, these trade-offs between traits may be only visible when resources are not abundant as stress tolerance is linked with metabolic ability. Previous research has also suggested that genetic correlations among traits will often be positive when resources are abundant, while negative correlations reflecting trade-offs may only be apparent when fitness is measured in a resource-poor environment (Sgrò and Hoffmann 2004). The evolutionary experiment in chapter 4 also highlighted a larger diversification of clones adapted to stress conditions compared to the stress-free adapted clones, suggesting that environmental stresses and trait correlations can drive increased phenotypic diversity. Overall, metadata or life-history characteristics explain surprisingly little phenotypic variation. However, trait correlations are abundant and dynamic, potentially explaining trait differences observed between RSSC isolates.

5.4 What genetic mechanisms are underpinning trait variation within the *Ralstonia solanacearum* species complex (RSSC)?

This research has also linked phenotypes with genetic information within the *Ralstonia solanacearum* species complex (RSSC). RSSC isolates also have large bipartite genomes with a great number of hypothetical proteins and therefore, a better understanding of how these complex genetic mechanisms link to trait characteristics is needed to provide useful annotations of these genes. This is especially important during this current omics era where vast quantities of genomic data are now available. RSSC are genetically diverse, with many regulatory pathways that respond to internal and environmental cues allowing for high phenotypic plasticity (Schell 2000, Cellier and Prior 2010, Genin and Denny 2012, Perrier et al. 2019, Chen et al. 2022, Yan et al. 2022). Furthermore, genome movements (such as prophages, insertion sequences, etc.) (Gonçalves et al. 2021, Greenrod et al. 2022), and epigenetic differences, for example methylation patterns (Erill et al. 2017), are also highly diverse among the RSSC. Therefore, this could explain why large amounts of phenotypic variation was not linked to genetic differences observed within this thesis. Despite this some key findings were discovered.

Within this thesis, chapter 2 highlighted that accessory genome variation is potentially driving ecotype differences among the two species studied within the RSSC. Previous research on RSSC genomics has also found that each RSSC strain appears to exhibit a highly diverse genetic content (Ailloud et al. 2011), suggesting that horizontal gene transfer (HGT) plays a large role in RSSC adaptability and genome content. Previous research on genomic analysis of RSSC strains suggested that HGT played a great role in shaping the genomic plasticity and genetic diversity of RSSC genomes (Geng et al. 2022) and this thesis has begun to link this with ecological trait differences. The transfer of plasmids between bacterial species is also known to help bacteria adapt within the environment, such as antimicrobial resistance plasmids (Dimitriu 2022). However, while small plasmids have been detected within RSSC strains, this had only been the case so far within three strains, CMR15, PS107 and T78 (Remenant et al. 2010, Genin and Denny 2012, Cho et al. 2019). Further research may discover more strains within the RSSC that contain

small plasmids, or this could suggest that plasmids are easily lost within RSSC strains. Overall, research has shown that recombination plays a major role in RSSC genome evolution, as large numbers of genomic islands surrounded by mobile elements are present within the genome (Remenant et al. 2010, Peeters et al. 2013, Geng et al. 2022). This supports my findings that gene acquisition is responsible for ecotype differences at a global scale

Chapter 3 showed that *R. solanacearum* isolates within the UK are extremely genetically similar, despite 30 years of sampling. UK *R. solanacearum* belongs to the cold tolerant strain of *Ralstonia solanacearum* (race 3 biovar 2 or phylotype IIB sequevar 1), the causative agent of potato brown rot in Europe (Fegan and Prior 2005, Safni et al. 2014) and has only recently been introduced to the country (Elphinstone and Matthews-Berry 2017). Therefore, this lack of genetic variation could be due to a recent invasion and expansion of the population. However, the pathogen has been present in the UK for at least 30 years and some genome differences should be expected to have been occurred within that time. Another explanation could be that UK *R. solanacearum* is at an evolutionary fitness peak, as cold tolerance, or other adaptations to the UK environment, limit adaptation due to trade-offs. Furthermore, annual bottlenecking at winter could explain low genetic diversity. Each winter *R. solanacearum* densities fall to undetectable levels within rivers (Elphinstone and Matthews-Berry 2017). It is thought that *R. solanacearum*'s ability to survive the cool winter temperatures in UK rivers is primarily due to the presence of *Solanum dulcamara* along the river beds with bacteria colonising the roots over winter (Genin and Boucher 2002, Champoiseau et al. 2009). Therefore, population densities should be expected to be a lot lower at winter than summer, suggesting an annual bottlenecking event. However, while UK *R. solanacearum* were genetically very similar, there was still large variation in phenotype, suggesting that differences could be due to other genetic information, such as epigenetics, not captured in this study.

This thesis has revealed that the high genetic similarity seen among UK *R. solanacearum* isolates in their core genome and their gene presence/absence matrix (chapter 3) could be due to insertion sequence (IS) movement instead of core genome differences. Chapter 4 revealed that insertion sequence (IS)

movement had a large role in the initial adaptation of *R. solanacearum* to environmental stress conditions within an evolution experiment. Therefore, this could be the case within the UK environment also. Specific IS movement, particularly IS1021 loss within the megaplasmid, was linked to abiotic stress tolerance adaptation within the laboratory. While IS movement among *Ralstonia solanacearum* has been explored (Greenrod et al. 2022), linking specific IS movements with isolate's phenotypic variation is still required to determine IS movement functionality. Therefore, future research could focus on comparing different trait measurements with UK *R. solanacearum's* IS movement profile.

Overall, linking genetic data with RSSC phenotypic variation has revealed novel insights into how this plant pathogen adapts. Accessory genome variation drives RSSC local adaptation to the environment, explaining differences in ecotypes discovered in chapter 2. While initial adaptation to the UK environment can be driven by IS movement within *R. solanacearum,* explaining the absence of SNP genetic mutations within this population. Surprisingly, little phenotypic differences among RSSC isolates can be explained by their core genome variation, justifying the need to classify this pathogen as a species complex.

5.5 Wider implications of this research

Due to the danger RSSC poses to our food supply, it has been classed as a quarantine pest in many countries, including the EU (EPPO 2022), and is classed as a bioterrorism select agent in the US (USDA 2020) (Cellier and Prior 2010). However, the high diversity and plastic nature of RSSC means that it is difficult to control. This is because RSSC can infect a large range of host plants, including many weed species, such as *Solanum dulcamara* in Europe (Wenneker et al. 1999) and can persist in non-host environments, such as within water sources (and potential irrigation sources), within soil communities, and on farm equipment for a debated period of time (Bragard et al. 2019). This research highlights the ecological diversity of RSSC and reveals that this is not linked to RSSC phylogeny (chapter 2). This can create problems when designing control strategies, highlighting the need to test potential control strategies on a wide range of isolates. This research also suggests that the surrounding microbial community has a large impact on RSSC diversity,

with accessory genome variation, most likely caused by horizontal gene transfer (HGT), resulting in adaptation to environmental niches (chapter 2). Therefore, when thinking about the control and limiting the spread of pathogenic bacteria, like RSSC, plant pathologists should consider the surrounding microbial community. However, this research also highlights why RSSC is called a species complex as few species-specific differences were found. Therefore, this research suggests that when designing control methods we can consider all RSSC species as the same pathogen.

An increase in antibiotic resistance over time was seen in this thesis within UK *R. solanacearum* (chapter 3) which can have implications for the control of this pathogen. For example, reducing the efficacy of antibiotic-producing biocontrol agents (Zhou et al. 2012, Singh and Kumar Yadav 2016). Antagonistic bacteria, like *Bacillus amyloliquefaciens* (Singh and Kumar Yadav 2016) or *Pseudomonas brassicacearum* J12 strain, produce antimicrobial compounds, such as 4-diacetylphloroglucinol (2,4-DAPG) (Zhou et al. 2012), significantly reducing bacterial wilt disease symptoms in planta (Zhou et al. 2012, Singh and Kumar Yadav 2016). However, natural antimicrobial resistance within the population may mean that some biocontrol agents will be or will quickly become ineffective. This increase in antimicrobial resistance within environmental isolates of *R. solanacearum* also suggests that they can act as reservoirs for antimicrobial resistance genes also having implications for antibiotic efficacy in agricultural and even clinical settings. Resistance genes are usually associated with mobile genetic elements, which can be transferred between distantly related bacteria within the environment (Wellington et al. 2013). Therefore, if transferred to other pathogenic bacteria this could have implications in human medicine. However, the risk of this should be low due to the lack of plasmids found among RSSC strains (Remenant et al. 2010, Genin and Denny 2012, Cho et al. 2019).

Furthermore, chapter 4 revealed that mutations were not the main genetic driver of *R. solanacearum* adaptation within the laboratory. Insertion sequence movement was discovered as being highly abundant and was linked to environmental stress evolution. This highlights the importance of looking at multiple genetic mechanisms, such as IS movement, prophages, methylation patterns and so on, when looking at genetic causes of adaptation after evolutionary

experiments for RSSC and other microorganisms. Bioinformatic pipelines are constantly being developed to improve the detection of these variants and more long-read sequencing (such as Nanopore and PacBio) and should help to discover different variants more easily in the future.

5.6 Future developments and limitations

While this thesis has revealed more about the diversity of the *Ralstonia solanacearum* species complex (RSSC), there is still a lot still left unexplored. The aim of this thesis was to explore the phenotypic diversity of this plant pathogen and one method to conduct this was by measuring a broad range of ecologically relevant traits (n=46) on a collection of RSSC isolates (chapter 2 and 3). However, while this is the most extensive study on RSSC ecological diversity to date, these chosen traits will not have captured all the phenotypic diversity present among the RSSC and more traits, such as host range, motility, or growth on solid media to name a few, could separate these isolates even further. Furthermore, conducting this comparative analysis on more RSSC isolates would most likely reveal higher phenotypic diversity within the RSSC. For example, this thesis does not include one of the three RSSC species, *R. syzygii* (Wicker et al. 2012), a highly diverse bacterial species which has therefore been subdivided into three separate subspecies (Safni et al. 2014, 2018). Additionally, within the RSSC bacterial collection used in this thesis there are gaps in the metadata due to limitations on the samples and information available at FERA science ltd.

Furthermore, within chapter 3, UK *R. solanacearum* diversification over time (1992-2019) was explored. However, certain years had no *R. solanacearum* samples representing them. This was due to the high detection limit, two viable pathogen cells per ml of river water (Elphinstone and Matthews-Berry 2017), needed to isolate *R. solanacearum* from river water in the UK. Therefore, within these years the pathogen never reached detectable levels and therefore no samples were collected. Within this thesis, I have discovered diversification of UK isolates over time, with antibiotic resistance and biofilm production traits increasing within the population with time. A continuation of sampling of UK *R. solanacearum* isolates can therefore reveal if this pattern continues and can also explore if these isolates

will remain clonal or if genetic diversification will begin. Overall, more phenotypic traits and isolates involved within this study could reveal more about the diversity of this plant pathogen. However, this would be a huge undertaking and there are limitations, such as which isolates are detectable within the environment, which were out of our control.

Within this thesis, I used comparative analysis on environmental isolates to explore the phenotypic diversity of a natural bacterial population. However, some of the trait variation observed could be due to lab artefacts. This was controlled for as much as possible through multiple methods. First, each isolate's stress traits were divided by their growth in the absence of that stressor, within the same batch, to attempt to control for batch effects and media preferences between isolates. Secondly, three repeats were also conducted for each isolate and trait combination to reduce random measurement error effects. Furthermore, among each inoculation plate used for the high-throughput phenotyping were two well-characterised RSSC reference strains, the *R. pseudosolanacearum* type strain GMI1000 from French Guyana, and *R. solanacearum* type strain K60 from USA. These strains were repeated three/four times per plate and were placed strategically across the microplates to detect and account for both batch and plate position effects. A linear model was then conducted using batch, plate position and technical replicate as random effects was then conducted on these two strains to determine that only a small proportion of variation was explained by these variables (7%, 6% and 0% for batch, plate position and technical replicate respectively). Overall, multiple control methods were used to reduce lab artifacts within this experiment. However, despite these efforts some variation between isolates may be lab artifacts and relating these trait measurements to natural environments should be taken with a touch of caution.

This thesis also highlights the potential use of genome wide association study (GWAS) techniques on exploring genetic causes of phenotypic traits within microorganisms. Chapter 2 showed that GWAS techniques can be used to link phenotypes to genotypes within the RSSC, where a type II secretion system was found to be associated with cold tolerance and three novel genome regions associated with rifampicin resistance. Bacterial GWAS have three main confounding

factors to take into consideration. First, bacteria have higher linkage between genes via population sweeps, or bottlenecks, which removes much of the genetic diversity within a population. They are also asexual and lack the ability to reduce linkage disequilibrium (LD) as quickly as sexually reproducing organisms (Read and Massey 2014), therefore causative genes should ideally be acquired multiple times by different lineages, as this improves the evidence for the gene causing the trait (Read and Massey 2014). Finally, they have a higher population stratification due to lack of recombination. Population structure confounds GWAS as the non-random distribution of alleles within subpopulations can cause significant associations between genotypes and phenotypes to be detected which are due to bacteria being related rather than the genome region causing the phenotype (Read and Massey 2014, Chen and Shapiro 2015). The problem of population stratification is particularly acute in highly clonal (rarely recombining) bacteria, and in those with separate geographic or host-associated subpopulations (Chen and Shapiro 2015). New refinements on GWAS approaches for microbes are constantly being developed, each using different methods to control for these confounding factors. Some software use clustering methods to control for population structure while others use phylogenetic trees (Collins and Didelot 2018) to take into account relatedness between clonal lineages. This research used the pyseer GWAS software for microbial GWAS (Lees et al. 2018) with distances from a phylogenetic tree used to correct for population structure. This software can also utilise unitigs and gene presence/absence as the input genetic variable as well as mutations within the core genome, such as single-nucleotide polymorphisms (SNPs). This is important as my research in chapter 2 suggests that accessory genome variation is important for RSSC trait variation. However, despite correcting for these confounding factors there are still limitations in the software. For example, my results highlight that those traits with strong signals, such as antibiotic resistance, deliver more promising results compared to more subtle traits, such as growth on different carbon sources, due to population structure confounding results agreeing with past studies (Power et al. 2016). Overall, this thesis has shown that GWAS can be used within bacterial species to discover novel genes associated with traits, however it is important to note that these are associations and may not be the causative gene/mutation.

5.7 Concluding remarks

This thesis has given an in-depth investigation into the phenotypic diversity of the bacterial plant pathogen *Ralstonia solanacearum* species complex (RSSC). It has considered diversity at the global and population level, as well as within the laboratory (figure 5.1). While also linking some of this trait variation to genetic data, discovering the causative genetic mechanisms behind certain traits. The RSSC is highly diverse phenotypically and a large overlap in trait diversity exists between two RSSC species (*R. pseudosolanacearum* and *R. solanacearum*), supporting the classification that this plant pathogen is a species complex and suggests that differences between members cannot be determined by core genome differences alone. I have identified that niche specialisation drives high diversity among this plant pathogen species, both at the species and population level. However, while ecological differences were shown to be driven by accessory genome variation at the global level, insertion sequence (IS) movement may be driving initial adaptation to the environment within a single UK population of *R. solanacearum*. I also showed that environmental stresses can drive increased *R. solanacearum* diversity, caused by negative trait-correlations between different environmental stresses and metabolic traits. Finally, a genome-wide association study (GWAS) revealed a type II secretion system associated with cold tolerance and three novel genome regions associated with rifampicin resistance within the RSSC. Overall, this thesis has shown that comparative analysis combined with evolution experiments can build knowledge on the adaptation of microorganisms in natural environments and provides insights into RSSC ecological diversity improving our knowledge of the epidemiology of this plant pathogen.

Appendices

Appendix A: Bacterial collections and supplemental methods

Appendix Table A.1: Global *Ralstonia solanacearum* species complex (RSSC) collection. Table listing the world RSSC isolate collection, used within chapter 2, with identification names, year, host, and country isolated from as well as the assigned species and phylotype based on whole genome sequencing analysis.

York Number (collection: University of York)	**Protect Number** (collection: FERA Science ltd.)	**NCPPB Number** (collection: FERA Science ltd.)	**Other Identifier** (collection: FERA Science ltd.)	**Host**	**Year**	**Continent**	**Country**	**Species**	**Phylotype**
YO001	7357	NCPPB 1018		Potato	1961	Africa	Angola	*Ralstonia pseudosolanacearum*	III
YO002	6825	NCPPB 1483		Potato	1963	Oceania	Australia	*Ralstonia solanacearum*	II
YO003		NCPPB 2245		Stylosanthes humulis	1969	Oceania	Australia	*Ralstonia pseudosolanacearum*	I
YO004		NCPPB 3980		Potato	1997	Oceania	Australia	*Ralstonia solanacearum*	II
YO005		NCPPB 3992		Tobacco		Oceania	Australia	*Ralstonia pseudosolanacearum*	I
YO006	4364	NCPPB 4001		Ginger		Oceania	Australia	*Ralstonia pseudosolanacearum*	I
YO007	3615			Potato		Oceania	Australia	*Ralstonia pseudosolanacearum*	I
YO008			21203224	Potato	2012	Asia	Bangladesh	*Ralstonia solanacearum*	II
YO009			21204041	Potato	2012	Asia	Bangladesh	*Ralstonia pseudosolanacearum*	I
YO010			21208789	Potato	2012	Asia	Bangladesh	*Ralstonia solanacearum*	II
YO011			21404030	Potato	2014	Asia	Bangladesh	*Ralstonia solanacearum*	II
YO012	3603			Potato	1993	Europe	Belgium	*Ralstonia solanacearum*	II
YO013		NCPPB 0613		Potato	1958	South America	Brazil	*Ralstonia solanacearum*	II
YO014	5683			Tobacco	1968	South America	Brazil	*Ralstonia solanacearum*	II
YO015		NCPPB 3649		Banana	1979	South America	Brazil	*Ralstonia solanacearum*	II
YO016		NCPPB 3650		Banana	1979	South America	Brazil	*Ralstonia solanacearum*	II
YO017	1193	NCPPB 3868		Potato	1991	South America	Brazil	*Ralstonia solanacearum*	II

YO018	1195	NCPPB 3864		Chili	1991	South America	Brazil	*Ralstonia pseudosolanacearum*	I
YO019	1199, 6282	NCPPB 3862		Chili	1991	South America	Brazil	*Ralstonia solanacearum*	II
YO020	1198	NCPPB 3863		Tomato	1991	South America	Brazil	*Ralstonia pseudosolanacearum*	I
YO021	1202, 7360	NCPPB 3866		Potato	1993	South America	Brazil	*Ralstonia pseudosolanacearum*	I
YO022		NCPPB 3982		Potato		South America	Chile	*Ralstonia solanacearum*	II
YO023	4369	NCPPB 4006		Olive		Asia	China	*Ralstonia pseudosolanacearum*	I
YO024	4370	NCPPB 4007		Mulberry		Asia	China	*Ralstonia pseudosolanacearum*	I
YO025	4375, 5787	NCPPB 4012		Mulberry		Asia	China	*Ralstonia pseudosolanacearum*	I
YO026	4374, 5786, 6289	NCPPB 4011, NCPPB 3850		Mulberry		Asia	China	*Ralstonia solanacearum*	II
YO027		NCPPB 3994		Olive		Asia	China	*Ralstonia pseudosolanacearum*	I
YO028		NCPPB 3998		Ginger		Asia	China	*Ralstonia pseudosolanacearum*	I
YO029	4366	NCPPB 4003		Ginger		Asia	China	*Ralstonia pseudosolanacearum*	I
YO030	4371	NCPPB 4008		Peanut		Asia	China	*Ralstonia pseudosolanacearum*	I
YO031	6822	NCPPB 282		Potato	1950	South America	Colombia	*Ralstonia solanacearum*	II
YO032	5770	NCPPB 3594		Heliconia caribaea	1960	South America	Colombia	*Ralstonia solanacearum*	II
YO033	5680			Tobacco	1966	South America	Colombia	*Ralstonia solanacearum*	II
YO034	5771	NCPPB 2154		Heliconia sp.	1958	North America	Costa Rica	*Ralstonia pseudosolanacearum*	I
YO035	8026	NCPPB 787		Banana	1959	North America	Costa Rica	*Ralstonia solanacearum*	II
YO036		NCPPB 0788		Banana	1959	North America	Costa Rica	*Ralstonia solanacearum*	II
YO037		NCPPB 0790		Nightshade	1959	North America	Costa Rica	*Ralstonia pseudosolanacearum*	I
YO038		NCPPB 0791		False daisy	1959	North America	Costa Rica	*Ralstonia pseudosolanacearum*	I

YO039	5688			Potato	1972	North America	Costa Rica	*Ralstonia solanacearum*	II
YO040	4367	NCPPB 4004		Ginger		North America	Costa Rica	*Ralstonia pseudosolanacearum*	I
YO041	5681	NCPPB 3977		M. perfoliatum		North America	Costa Rica	*Ralstonia solanacearum*	II
YO042		NCPPB 3993		Pepper		North America	Costa Rica	*Ralstonia pseudosolanacearum*	I
YO043	3612			Potato		North America	Costa Rica	*Ralstonia solanacearum*	II
YO044	6823	NCPPB 643		Potato	1959	Europe	Cyprus	*Ralstonia solanacearum*	II
YO045	6281	NCPPB 1584		Potato	1963	Europe	Cyprus	*Ralstonia solanacearum*	II
YO046	6824	NCPPB 909		Potato	1961	Africa	Egypt	*Ralstonia solanacearum*	II
YO047		NCPPB 1115		Potato	1961	Africa	Egypt	*Ralstonia solanacearum*	II
YO048	6827	NCPPB 1824		Potato	1966	Africa	Egypt	*Ralstonia solanacearum*	II
YO049	803			Potato	1991	Africa	Egypt	*Ralstonia solanacearum*	II
YO050	804			Potato	1991	Africa	Egypt	*Ralstonia solanacearum*	II
YO051	805			Potato	1991	Africa	Egypt	*Ralstonia solanacearum*	II
YO052	819			Potato	1991	Africa	Egypt	*Ralstonia solanacearum*	II
YO053	1328			Potato	1994	Africa	Egypt	*Ralstonia solanacearum*	II
YO054	1538			Potato	1995	Africa	Egypt	*Ralstonia solanacearum*	II
YO055	1539			Potato	1995	Africa	Egypt	*Ralstonia solanacearum*	II
YO056	1540			Potato	1995	Africa	Egypt	*Ralstonia solanacearum*	II
YO057	1541			Potato	1995	Africa	Egypt	*Ralstonia solanacearum*	II
YO058	1542			Potato	1995	Africa	Egypt	*Ralstonia solanacearum*	II
YO059	1543			Potato	1995	Africa	Egypt	*Ralstonia solanacearum*	II
YO060	1329			Potato	1995	Africa	Egypt	*Ralstonia solanacearum*	II
YO061	1544			Potato	1995	Africa	Egypt	*Ralstonia solanacearum*	II
YO062	1330			Potato	1996	Africa	Egypt	*Ralstonia solanacearum*	II
YO063	1331			Potato	1997	Africa	Egypt	*Ralstonia solanacearum*	II
YO064	1332			Potato	1998	Africa	Egypt	*Ralstonia solanacearum*	II
YO065	3122			Potato	1998	Africa	Egypt	*Ralstonia solanacearum*	II

YO066	4433, 7361	NCPPB 4153		Potato	1998	Africa	Egypt	*Ralstonia solanacearum*	II
YO067	3123			Potato	1998	Africa	Egypt	*Ralstonia solanacearum*	II
YO068	4567			Water	2002	Africa	Egypt	*Ralstonia solanacearum*	II
YO069	4573			Water	2002	Africa	Egypt	*Ralstonia solanacearum*	II
YO070	4578			Potato	2002	Africa	Egypt	*Ralstonia solanacearum*	II
YO071	4584			Soil	2002	Africa	Egypt	*Ralstonia solanacearum*	II
YO072	4587				2002	Africa	Egypt	*Ralstonia solanacearum*	II
YO073	4592				2002	Africa	Egypt	*Ralstonia solanacearum*	II
YO074		NCPPB 1500		Potato	1963	Oceania	Fiji	*Ralstonia pseudosolanacearum*	I
YO075		NCPPB 1702		Potato	1964	Oceania	Fiji	*Ralstonia pseudosolanacearum*	I
YO076		NCPPB 1029		Pelargonium capitatum	1961	Africa	France	*Ralstonia pseudosolanacearum*	III
YO077		NCPPB 2200		Tomato	1966	North America	France	*Ralstonia solanacearum*	II
YO078	4437, 7362	NCPPB 4157		Potato	1995	Europe	France	*Ralstonia solanacearum*	II
YO079			21108193	Tomato	2011	Europe	France	*Ralstonia pseudosolanacearum*	I
YO080		NCPPB 2204		Tomato	1968	South America	French Guyana	*Ralstonia pseudosolanacearum*	I
YO081	2015			Tomato		South America	French Guyana	*Ralstonia pseudosolanacearum*	I
YO082		NCPPB 3181		Nightshade	1978	Africa	Gambia	*Ralstonia pseudosolanacearum*	III
YO083	7790			Tomato	2011	Asia	Georgia	*Ralstonia pseudosolanacearum*	I
YO084			21112409	Tomato	2011	Asia	Georgia	*Ralstonia pseudosolanacearum*	I
YO085	4441, 3604, 5689	NCPPB 4161		Potato	1996	Europe	Germany	*Ralstonia solanacearum*	II
YO086		NCPPB 1789		Potato	1965	Europe	Greece	*Ralstonia solanacearum*	II
YO087	6828	NCPPB 2015		Potato	1967	Europe	Greece	*Ralstonia solanacearum*	II
YO088	8027	NCPPB 3205		Banana	1979	South America	Guyana	*Ralstonia solanacearum*	II

YO089		NCPPB 0789		Banana	1959	North America	Honduras	*Ralstonia solanacearum*	II
YO090			21123931	Potato	2012	Europe	Hungary	*Ralstonia solanacearum*	II
YO091			21123932	Potato	2012	Europe	Hungary	*Ralstonia solanacearum*	II
YO092			21123933	Potato	2012	Europe	Hungary	*Ralstonia solanacearum*	II
YO093			21123934	Potato	2012	Europe	Hungary	*Ralstonia solanacearum*	II
YO094			21123935	Potato	2012	Europe	Hungary	*Ralstonia solanacearum*	II
YO095			21123936	Potato	2012	Europe	Hungary	*Ralstonia solanacearum*	II
YO096		NCPPB 1331		Potato	1962	Asia	India	*Ralstonia solanacearum*	II
YO097		NCPPB 1333		Potato	1962	Asia	India	*Ralstonia solanacearum*	II
YO098	8025	NCPPB 3214		Banana	1980	Asia	India	*Ralstonia solanacearum*	II
YO101	147	NCPPB 3793		Potato	1985	Asia	Indonesia	*Ralstonia pseudosolanacearum*	I
YO109	143			Clove	1987	Asia	Indonesia	*Ralstonia solanacearum*	II
YO111	148	NCPPB 3794		Clove	1987	Asia	Indonesia	*Ralstonia solanacearum*	II
YO112	7233		20720807	Water	2007	Europe	Ireland	*Ralstonia solanacearum*	II
YO113	7234		20720808	Water	2007	Europe	Ireland	*Ralstonia solanacearum*	II
YO114			20719104	Potato	2007	Europe	Ireland	*Ralstonia solanacearum*	II
YO115	7235		20720809		2007	Europe	Ireland	*Ralstonia solanacearum*	II
YO116			20722587	Tomato	2007	Europe	Ireland	*Ralstonia solanacearum*	II
YO117			20722588	Tomato	2007	Europe	Ireland	*Ralstonia solanacearum*	II
YO119	140	NCPPB 3445		Clove	1983	Asia	Java	*Ralstonia syzygii*	IV
YO120	6577	NCPPB 173		Potato	1945	Africa	Kenya	*Ralstonia solanacearum*	II
YO121	7358	NCPPB 1028		Potato	1961	Africa	Kenya	*Ralstonia solanacearum*	II
YO122		NCPPB 1045		Eggplant	1961	Africa	Kenya	*Ralstonia pseudosolanacearum*	I
YO123		NCPPB 1049		Tomato	1961	Africa	Kenya	*Ralstonia solanacearum*	II
YO124	4142, 4448	NCPPB 4215		Water	2001	Africa	Kenya	*Ralstonia pseudosolanacearum*	I
YO125	4134, 4444	NCPPB 4211		Pelargonium hortorum	2001	Africa	Kenya	*Ralstonia solanacearum*	II
YO126	4135, 4445	NCPPB 4212		Pelargonium hortorum	2001	Africa	Kenya	*Ralstonia solanacearum*	II

YO127	4138, 4446	NCPPB 4213		Water	2001	Africa	Kenya	*Ralstonia solanacearum*	II
YO128	4140, 4447	NCPPB 4214		Soil	2001	Africa	Kenya	*Ralstonia pseudosolanacearum*	III
YO129	6826	NCPPB 1489		Potato	1963	Africa	Madeira	*Ralstonia solanacearum*	II
YO130		NCPPB 0792		Teak	1960	Asia	Malaysia	*Ralstonia pseudosolanacearum*	I
YO131		NCPPB 1052		Ginger	1961	Asia	Malaysia	*Ralstonia pseudosolanacearum*	I
YO132		NCPPB 1614		Potato	1964	Asia	Malaysia	*Ralstonia solanacearum*	II
YO133		NCPPB 3190		Tomato	1978	Asia	Malaysia	*Ralstonia pseudosolanacearum*	I
YO134		NCPPB 2199		Eggplant	1965	North America	Martinique	*Ralstonia solanacearum*	II
YO135	876	NCPPB 253		Pine tree	1949	Africa	Mauritius	*Ralstonia pseudosolanacearum*	I
YO136		NCPPB 0500		Broad bean	1956	Africa	Mauritius	*Ralstonia pseudosolanacearum*	I
YO137		NCPPB 0501		Cabbage	1956	Africa	Mauritius	*Ralstonia pseudosolanacearum*	I
YO138		NCPPB 0503		Dahlia sp.	1956	Africa	Mauritius	*Ralstonia pseudosolanacearum*	I
YO139		NCPPB 1621		Potato	1960	Africa	Mauritius	*Ralstonia pseudosolanacearum*	I
YO140		NCPPB 1484		Strelitzia reginae	1963	Africa	Mauritius	*Ralstonia pseudosolanacearum*	I
YO141		NCPPB 1485		Common bean	1963	Africa	Mauritius	*Ralstonia pseudosolanacearum*	I
YO142	5682	NCPPB 3974		Tomato		North America	Mexico	*Ralstonia solanacearum*	II
YO143	6830	NCPPB 3238		Potato	1982	Europe	Netherlands	*Ralstonia solanacearum*	II
YO144	4436, 3605, 5690	NCPPB 4156		Potato	1995	Europe	Netherlands	*Ralstonia solanacearum*	II
YO145		NCPPB 1703		Potato	1965	Europe	Nigeria	*Ralstonia pseudosolanacearum*	III
YO146		NCPPB 2088		Potato	1968	Africa	Nigeria	*Ralstonia solanacearum*	II
YO147			21106713	Potato	2011	Asia	Pakistan	*Ralstonia solanacearum*	II
YO148			21108524	Potato	2011	Asia	Pakistan	*Ralstonia solanacearum*	II

YO149	5781	NCPPB 1123		Tomato	1961	Oceania	Papua New Guinea	*Ralstonia pseudosolanacearum*	I
YO150	5782	NCPPB 1140		Tomato	1961	Oceania	Papua New Guinea	*Ralstonia pseudosolanacearum*	I
YO151		NCPPB 2937		Potato	1975	Oceania	Papua New Guinea	*Ralstonia pseudosolanacearum*	I
YO152	7028, 6285, 3602	NCPPB 3985		Eggplant	1987	South America	Peru	*Ralstonia solanacearum*	II
YO153	5772	NCPPB 3990		Potato	1989	South America	Peru	*Ralstonia solanacearum*	II
YO154	3613			Tomato		South America	Peru	*Ralstonia pseudosolanacearum*	I
YO156	5785, 6287	NCPPB 3996		Tomato		South America	Peru	*Ralstonia pseudosolanacearum*	I
YO157	8024	NCPPB 2315		Banana		South America	Peru	*Ralstonia solanacearum*	II
YO158		NCPPB 3986		Potato		South America	Peru	*Ralstonia solanacearum*	II
YO159	5776	NCPPB 3970		Banana	1992	Asia	Philippines	*Ralstonia solanacearum*	II
YO160	5778	NCPPB 3971		Banana		Asia	Philippines	*Ralstonia solanacearum*	II
YO161	6288, 4368, 3614	NCPPB 4005		Ginger		Asia	Philippines	*Ralstonia pseudosolanacearum*	I
YO162			21422327		2014	Europe	Poland	*Ralstonia solanacearum*	II
YO163			21622099	Rose	2016	Europe	Poland	*Ralstonia pseudosolanacearum*	I
YO164			21622100	Rose	2016	Europe	Poland	*Ralstonia pseudosolanacearum*	I
YO165		NCPPB 1019		Tomato	1960	Europe	Portugal	*Ralstonia solanacearum*	II
YO166	3608			Potato	1995	Europe	Portugal	*Ralstonia solanacearum*	II
YO167	4438	NCPPB 4158		Potato	1995	Europe	Portugal	*Ralstonia solanacearum*	II
YO168	7421	NCPPB 1225		Tomato	1958	North America	Puerto Rico	*Ralstonia solanacearum*	II
YO169		NCPPB 1226		Tomato	1958	North America	Puerto Rico	*Ralstonia solanacearum*	II
YO170			21123062 (1)	Potato	2012	Europe	Serbia	*Ralstonia solanacearum*	II
YO171			21123062 (2)	Potato	2012	Europe	Serbia	*Ralstonia solanacearum*	II
YO172			21123063 (1)	Potato	2012	Europe	Serbia	*Ralstonia solanacearum*	II
YO173			21123063 (2)	Potato	2012	Europe	Serbia	*Ralstonia solanacearum*	II

YO174			21025506/2		2013	Europe	Serbia	*Ralstonia solanacearum*	II
YO175			21025507/5		2013	Europe	Serbia	*Ralstonia solanacearum*	II
YO176			21025509/10		2013	Europe	Serbia	*Ralstonia solanacearum*	II
YO177			21025510/11		2013	Europe	Serbia	*Ralstonia solanacearum*	II
YO178		NCPPB 1763		Tomato	1965	Africa	Seychelles	*Ralstonia pseudosolanacearum*	I
YO180	4440, 3606	NCPPB 4160		Potato	1996	Europe	Spain	*Ralstonia solanacearum*	II
YO181		NCPPB 1323		Potato	1962	Asia	Sri Lanka	*Ralstonia solanacearum*	II
YO182		NCPPB 3217		Tumeric	1980	Asia	Sri Lanka	*Ralstonia pseudosolanacearum*	I
YO183	6293, 3607	NCPPB 2505		Potato	1972	Europe	Sweden	*Ralstonia solanacearum*	II
YO184	6578	NCPPB 2797		Solanum dulcamara	1974	Europe	Sweden	*Ralstonia solanacearum*	II
YO185		NCPPB 2796		Solanum dulcamara	1975	Europe	Sweden	*Ralstonia solanacearum*	II
YO187	4363	NCPPB 4000		Ginger		Asia	Thailand	*Ralstonia pseudosolanacearum*	I
YO188	5684			Tomato	1957	South America	Trinidad	*Ralstonia solanacearum*	II
YO189		NCPPB 0446		Banana	1957	South America	Trinidad	*Ralstonia solanacearum*	II
YO190		NCPPB 0616		Tomato	1957	South America	Trinidad	*Ralstonia solanacearum*	II
YO191	8023	NCPPB 2198		Banana	1968	South America	Trinidad	*Ralstonia pseudosolanacearum*	I
YO192		NCPPB 2201		Tomato	1968	South America	Trinidad	*Ralstonia solanacearum*	II
YO193		NCPPB 1486		Peanut	1963	Africa	Uganda	*Ralstonia pseudosolanacearum*	I
YO194		NCPPB 2484		Peanut	1969	Africa	Uganda	*Ralstonia pseudosolanacearum*	I
YO195			21517093	Potato	2015	Africa	Uganda	*Ralstonia solanacearum*	II
YO199	6594	NCPPB 3854		Potato	1992	Europe	UK	*Ralstonia solanacearum*	II
YO372	5788, 5687, 1532, 3611	NCPPB 325, NCPPB 3973		Tomato	1953	North America	USA	*Ralstonia solanacearum*	II
YO373		NCPPB 0337		Tobacco	1954	North America	USA	*Ralstonia solanacearum*	II
YO374		NCPPB 0338		Tobacco	1954	North America	USA	*Ralstonia solanacearum*	II

YO375	4372, 5685	NCPPB 4009		Tobacco	1955	North America	USA	*Ralstonia pseudosolanacearum*	I
YO376	7422	NCPPB 1580		Tomato	1959	North America	USA	*Ralstonia pseudosolanacearum*	I
YO378		NCPPB 1579		Ginger	1961	North America	USA	*Ralstonia pseudosolanacearum*	I
YO379		NCPPB 1581		Strelitzia reginae	1961	North America	USA	*Ralstonia pseudosolanacearum*	I
YO380		NCPPB 4362		Pelargonium hortorum	2003	North America	USA	*Ralstonia solanacearum*	II
YO381		NCPPB 3969		Banana		South America	Venezuela	*Ralstonia solanacearum*	II
YO382		NCPPB 0283		Solanum panduraforme	1950	Africa	Zimbabwe	*Ralstonia pseudosolanacearum*	III
YO383		NCPPB 0332		Potato	1954	Africa	Zimbabwe	*Ralstonia pseudosolanacearum*	III
YO384		NCPPB 0505		Comfrey	1956	Africa	Zimbabwe	*Ralstonia pseudosolanacearum*	III

Appendix Table A.2: UK *R. solanacearum* collection. Table listing the UK *R. solanacearum* isolate collection, used within chapter 3 and 4, with identification names, year, host, and county isolated from. YO336, the isolate chosen for the evolution experiment within chapter 4, is highlighted in grey.

York Number (collection: University of York)	**Protect Number** (collection: FERA Science ltd.)	**NCPPB Number** (collection: FERA Science ltd.)	**Other Identifier** (collection: FERA Science ltd.)	**Host**	**Year**	**Continent**	**Country**	**County**
YO196	1143			Potato	1992	Europe	UK	Oxfordshire
YO197	1123, 2797			Potato	1992	Europe	UK	Oxfordshire
YO198	1140			Potato	1992	Europe	UK	Oxfordshire
YO199	6594	NCPPB 3854		Potato	1992	Europe	UK	Oxfordshire
YO200	6595	NCPPB 3855		Potato	1992	Europe	UK	Oxfordshire
YO201	6596	NCPPB 3856		Potato	1992	Europe	UK	Oxfordshire
YO202	6598	NCPPB 3858		Potato	1992	Europe	UK	Oxfordshire
YO203		NCPPB 3815		Potato	1992	Europe	UK	Oxfordshire
YO204	6003			Solanum dulcamara	1993	Europe	UK	Oxfordshire
YO205	6011			Solanum dulcamara	1993	Europe	UK	Oxfordshire
YO206	6057			Solanum dulcamara	1993	Europe	UK	Oxfordshire
YO207	6058			Solanum dulcamara	1993	Europe	UK	Oxfordshire
YO208	6059			Solanum dulcamara	1993	Europe	UK	Oxfordshire
YO209	6060			Solanum dulcamara	1993	Europe	UK	Oxfordshire
YO210	6084			Solanum dulcamara	1993	Europe	UK	Oxfordshire
YO211	6087			Solanum dulcamara	1993	Europe	UK	Oxfordshire
YO212	6089			Solanum dulcamara	1993	Europe	UK	Oxfordshire
YO213	6091			Solanum dulcamara	1993	Europe	UK	Oxfordshire
YO214	6093			Solanum dulcamara	1993	Europe	UK	Oxfordshire
YO215	6095			Solanum dulcamara	1993	Europe	UK	Oxfordshire
YO216	6100			Solanum dulcamara	1993	Europe	UK	Oxfordshire
YO217	6634			Solanum dulcamara	1993	Europe	UK	Oxfordshire
YO218	6635			Solanum dulcamara	1993	Europe	UK	Oxfordshire
YO219	6097			Solanum dulcamara	1993	Europe	UK	Oxfordshire
YO220	6001			Solanum dulcamara	1994	Europe	UK	Wiltshire
YO221	6044			Solanum dulcamara	1994	Europe	UK	Berkshire

YO222	6045			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO223	6046			Solanum dulcamara	1994	Europe	UK	Berkshire
YO224	6047			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO225	6048			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO226	6049			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO227	6050			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO228	6051			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO229	6052			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO230	6053			Solanum dulcamara	1994	Europe	UK	Berkshire
YO231	6054			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO232	6055			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO233	6056			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO234	6061			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO235	6064			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO236	6068			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO237	6070			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO238	6071			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO239	6074			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO240	6075			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO241	6076			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO242	6077			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO243	6078			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO244	6079			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO245	6080			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO246	6081			Solanum dulcamara	1994	Europe	UK	Berkshire
YO247	6082			Solanum dulcamara	1994	Europe	UK	Berkshire
YO248	6083			Solanum dulcamara	1994	Europe	UK	Oxfordshire
YO249	5981			Water	1995	Europe	UK	Hertfordshire
YO250	5984			Water	1995	Europe	UK	Hertfordshire

YO251	5989			Water	1995	Europe	UK	Hertfordshire
YO252	6012			Water	1995	Europe	UK	Wiltshire
YO253	6013			Water	1995	Europe	UK	Wiltshire
YO254	6014			Water	1995	Europe	UK	Wiltshire
YO255	6015			Water	1995	Europe	UK	Wiltshire
YO256	6016			Water	1995	Europe	UK	Wiltshire
YO257	6636			Solanum dulcamara	1995	Europe	UK	Oxfordshire
YO258	5986			Water	1995	Europe	UK	Hertfordshire
YO259	6002			Water	1996	Europe	UK	Oxfordshire
YO260	6004			Solanum dulcamara	1996	Europe	UK	Greater London
YO261	6005			Solanum dulcamara	1996	Europe	UK	Buckinghamshire
YO262	6006			Solanum dulcamara	1996	Europe	UK	Surrey
YO263	6007			Solanum dulcamara	1996	Europe	UK	Surrey
YO264	6008			Solanum dulcamara	1996	Europe	UK	Surrey
YO265	6009			Solanum dulcamara	1996	Europe	UK	Berkshire
YO266	6062			Solanum dulcamara	1996	Europe	UK	Greater London
YO267	6102			Potato	1996	Europe	UK	Berkshire
YO268	6109			Water	1996	Europe	UK	Greater London
YO269	2490			Water	1997	Europe	UK	Surrey
YO270	2487			Water	1997	Europe	UK	Berkshire
YO271	2488			Water	1997	Europe	UK	Hampshire
YO272	2489			Water	1997	Europe	UK	Surrey
YO273	6017			Tomato	1997	Europe	UK	Bedfordshire
YO274	6018			Tomato	1997	Europe	UK	Bedfordshire
YO275	6019			Water	1997	Europe	UK	Bedfordshire
YO276	6020			Water	1997	Europe	UK	Bedfordshire
YO277	6105			Water	1997	Europe	UK	Bedfordshire
YO278	6106			Tomato	1997	Europe	UK	Bedfordshire
YO279	6107			Tomato	1997	Europe	UK	Bedfordshire
YO280	6108			Water	1997	Europe	UK	Bedfordshire
YO281	3325			Water	1998	Europe	UK	Cambridgeshire
YO282	6063			Water	1998	Europe	UK	Greater London
YO283	3326			Water	1998	Europe	UK	Bedfordshire
YO284	3597			Water	1999	Europe	UK	Northamptonshire
YO285	3596			Water	1999	Europe	UK	Northamptonshire
YO286	3729			Potato	1999	Europe	UK	Northamptonshire
YO287	3792			Water	2000	Europe	UK	Kent

YO288	5518			Water	2000	Europe	UK	Scotland
YO289	5822			Water	2005	Europe	UK	Bedfordshire
YO290	5823			Water	2005	Europe	UK	Warwickshire
YO291	5824			Water	2005	Europe	UK	Gloucestershire
YO292	5825			Water	2005	Europe	UK	Gloucestershire
YO293	5826			Water	2005	Europe	UK	Gloucestershire
YO294	5827			Water	2005	Europe	UK	Worcestershire
YO295	5828			Water	2005	Europe	UK	Worcestershire
YO296	5870			Water	2005	Europe	UK	Gloucestershire
YO297	5871			Water	2005	Europe	UK	Bedfordshire
YO298	5872			Water	2005	Europe	UK	Cambridgeshire
YO299	5873			Water	2005	Europe	UK	Bedfordshire
YO300	5874			Water	2005	Europe	UK	Cambridgeshire
YO301	5956			Potato	2005	Europe	UK	Nottinghamshire
YO302	6418			Water	2006	Europe	UK	Lincolnshire
YO303	6419			Water	2006	Europe	UK	Carmarthenshire
YO304	6420			Water	2006	Europe	UK	Warwickshire
YO305	6421			Water	2006	Europe	UK	Mid Glamorgan
YO306	6423			Water	2006	Europe	UK	Cambridgeshire
YO307	6424			Water	2006	Europe	UK	Cambridgeshire
YO308	6425			Water	2006	Europe	UK	Dorset
YO309	6426			Water	2006	Europe	UK	Dorset
YO310	6427			Water	2006	Europe	UK	Carmarthenshire
YO311	6428			Water	2006	Europe	UK	Carmarthenshire
YO312	6429			Water	2006	Europe	UK	West Midlands
YO313	6430			Water	2006	Europe	UK	West Midlands
YO314	6431			Water	2006	Europe	UK	Warwickshire
YO315	6432		W06/265	Water	2006	Europe	UK	Warwickshire
YO316	6436			Water	2006	Europe	UK	Carmarthenshire
YO317	6437			Water	2006	Europe	UK	Carmarthenshire
YO318	6438			Water	2006	Europe	UK	Dorset
YO319	6439			Water	2006	Europe	UK	Mid Glamorgan
YO320	6440			Water	2006	Europe	UK	Bedfordshire
YO321	6441		W06/427	Water	2006	Europe	UK	Warwickshire
YO322	6442		W06/414	Water	2006	Europe	UK	Staffordshire
YO323	6460		W06/614	Water	2006	Europe	UK	Carmarthenshire
YO324	6461			Water	2006	Europe	UK	Bedfordshire
YO325	6462			Water	2006	Europe	UK	Warwickshire
YO326	6463			Water	2006	Europe	UK	Bedfordshire
YO327	6464			Water	2006	Europe	UK	Dorset
YO328	6465			Water	2006	Europe	UK	Carmarthenshire
YO329	6466			Water	2006	Europe	UK	Mid Glamorgan

YO330	6470			Water	2006	Europe	UK	Warwickshire
YO331	6471			Water	2006	Europe	UK	Carmarthenshire
YO332	6472			Water	2006	Europe	UK	Cambridgeshire
YO333	6477		W06/614	Water	2006	Europe	UK	Carmarthenshire
YO334	6478		W06/818	Water	2006	Europe	UK	Cambridgeshire
YO335	6940			Water	2007	Europe	UK	Hampshire
YO336	6941			Water	2007	Europe	UK	Cambridgeshire
YO337	6942			Water	2007	Europe	UK	Cambridgeshire
YO338	6943			Water	2007	Europe	UK	Surrey
YO339	6944			Water	2007	Europe	UK	Hampshire
YO340	6945		W07/440	Water	2007	Europe	UK	Berkshire
YO341	6946		W07/439	Water	2007	Europe	UK	Berkshire
YO342	6947		W07/435	Water	2007	Europe	UK	Berkshire
YO343	6948		W07/438	Water	2007	Europe	UK	Berkshire
YO344	7114			Water	2008	Europe	UK	Oxfordshire
YO345	7115			Solanum dulcamara	2008	Europe	UK	Oxfordshire
YO346	7116			Solanum dulcamara	2008	Europe	UK	Oxfordshire
YO347	7117			Solanum dulcamara	2008	Europe	UK	Oxfordshire
YO348	7118			Water	2008	Europe	UK	Oxfordshire
YO349	7482			Potato	2009	Europe	UK	Cornwall
YO350			21314705	Water	2013	Europe	UK	Berkshire
YO351			21314706	Water	2013	Europe	UK	Berkshire
YO352			21415687	Water	2014	Europe	UK	Berkshire
YO353			21415697	Water	2014	Europe	UK	Berkshire
YO354			21517183	Water	2015	Europe	UK	Berkshire
YO355			21517184	Water	2015	Europe	UK	Berkshire
YO356			21620088	Water	2016	Europe	UK	Cambridgeshire
YO357			21620089	Water	2016	Europe	UK	Cambridgeshire
YO358			21620094	Water	2016	Europe	UK	Norfolk
YO359			21620744	Water	2016	Europe	UK	Norfolk
YO360			21713910	Water	2017	Europe	UK	Norfolk
YO361			21713920	Water	2017	Europe	UK	Norfolk
YO362			21714435	Water	2017	Europe	UK	Cambridgeshire
YO363			21714855	Water	2017	Europe	UK	Norfolk
YO364			21802385	Water	2018	Europe	UK	Norfolk
YO365			21812347	Water	2018	Europe	UK	Cambridgeshire
YO366			21812382	Water	2018	Europe	UK	Norfolk
YO367			21813021	Water	2018	Europe	UK	Cambridgeshire
YO368			21813608	Water	2018	Europe	UK	Norfolk
YO369			21813731	Water	2018	Europe	UK	Cambridgeshire

YO370	5519			Water		Europe	UK	
YO371	5517			Water		Europe	UK	
YO385				Solanum dulcamara	2019	Europe	UK	
YO386				Water	2019	Europe	UK	
YO387				Solanum dulcamara	2019	Europe	UK	
YO388				Water	2019	Europe	UK	
YO389				Solanum dulcamara	2019	Europe	UK	
YO390				Water	2019	Europe	UK	

Appendix Table A.3: Modified OS Media (OSG). Minimal medium, containing salts required for bacterial growth. Prepared from various stock solutions to avoid co-precipitation of the salts. Used within chapter 2, 3 and 4.

Solution Name	Add	Amount
OSA	Na_2HPO_4	35.05g (or $Na_2HPO_4 \times H_2O$, 44.5g or $Na_2HPO_4 \times 7H_2O$, 66.1g)
	KH_2PO_4	34.0g
	H_2O dist.	500ml
	Autoclave	
OSB	$MgSO_4 \times 7H_2O$	11.9g
	H_2O dist.	100ml
	Autoclave	
OSC	$(NH_4)_2SO_4$	10.0g
	H_2O dist.	100ml
	Autoclave	
OSD_1	$CaCl_2 \times 2H_2O$	2.2g
	H_2O dist.	50ml
	Autoclave	
OSD_2	$FeSO_4 \times 7H_2O$ dist.	50mg
	H_2O dist.	50ml
	Sterile Filter	
OSD_3	$(NH_4)_6Mo_7O_{24} \times 4H_2O$	5mg
	H_2O dist.	50ml
	Autoclave	

	(Prepare a 50mg/50ml solution then dilute 10x (ie: 5ml added to 45ml H_2O))	
OSE	EDTA (Na-salt)	250mg
	$ZnSO_4$ x $7H_2O$	1095mg
	$FeSO_4$ x $7H_2O$	500mg
	$MnSO_4$ x H_2O	154mg (or $Mn(II)Cl_2$ x $4H_2O$, 180mg)
	$CuSO_4$ x $5H_2O$	39mg
	$Co(NO_3)_2$ x $6H_2O$	25mg
	$Na_2B_4O_7$ x $10H_2O$	18mg
	$NiCl_2$ x $6H_2O$	130mg
	H_2O dist.	100ml
	Add 2 droplets of 1 N HCl	
	Sterile Filter	
	(Protect this solution from light – alu foil)	
OSF	Carbon of choice!	g to make 100mM solution
	H_2O dist.	50ml
	Autoclave	
All solutions can be stored at room temperature.		
OSG Medium	**Media:**	**Volume:**
	H_2O dist. sterile	848ml
	OSD_1	2ml

	OSD_2	2ml
	OSD_3	2ml
	OSA	100ml
	OSB	10ml
	OSC	10ml
	OSE	1ml
	water	5ml
Make sure pH is at around 7 before filter sterilising the OSG media.		
Mix by constant stirring with a magnet stirrer, sterile conditions		
Sterile filter (0.22µl) and aliquot in sterile 50ml falcon tubes. Seal with parafilm and store at 4°C.		
The carbon source can be added before starting the experiment. Add 5ml at 100mM to 45ml of OSG media to make 10mM overall.		

Appendix Table A.4: DNA extraction protocol. DNA extraction protocol of *Ralstonia solanacearum* species complex (RSSC) isolates, used in chapter 2, 3 and 4, using the Qiagen DNeasy Blood and Tissue Kit.

Preparation:	
Buffer AW1 and Buffer AW2	Buffer AW1 and Buffer AW2 are supplied as concentrates. Before using for the first time, add the appropriate volume of ethanol (96–100%) as indicated on the bottle and shake thoroughly. Buffer AW1 and Buffer AW2 are stable for at least 1 year after the addition of ethanol when stored closed at room temperature (15–25°C).
Proteinase K	Store in fridge for over a year. Or freeze for longer storage. The activity of proteinase K is 600 mAU/ml solution (or 40 mAU/mg protein), as has shown optimal results.
RNase A	RNase treatment was conducted. The concentration of RNase A used is 50 mg/ml.
1. Bacterial pellet preparation:	
1.1 Grow up bacteria	Place inoculation loop into cryopreserve freezer stock of strain and place onto SPA media plates. Grow at 28°C for 48 hours.
1.2 Place in Eppendorf	Take a single colony and place in an Eppendorf with 200µl of distilled water in.
1.3 Freeze pellets	Centrifuge Eppendorf's and remove excess water. Freeze pellets at -80°C.
2. Cell lysis:	
2.1 Resuspend pellet in 180 µl Buffer ATL.	Allowed precipitate in ATL to dissolve before use (place ATL bottle on 37°C hot block for 10 minutes or so).
2.2 Add 20 µl proteinase K.	Mix thoroughly by vortexing.
2.3 Incubate at 56°C for 2 hours	Place in a water bath shaking overnight and incubate until the tissue is completely lysed. After incubation the lysate may appear viscous, but should not be gelatinous as it may clog the DNeasy Mini spin column. If the lysate appears very gelatinous, incubate for longer or add more proteinase K.
3. DNA Purification:	
3.1 Add 2µl of RNase A	Add 2µl of RNase A (50 mg/ml), mix by vortexing, and incubate for 15 minutes at room temperature.
3.2 Vortex for 15 seconds	
3.3 Add 200 µl Buffer AL	Mix thoroughly by vortexing.
3.4 Add 200 µl ethanol (96–100%)	Mix again thoroughly by vortexing. (A white precipitate may form on addition of Buffer AL and ethanol. This precipitate does not interfere with the DNeasy procedure).

3.5 Pulse spin	Do a 1-3 pulse spin to ensure all liquid from lid id in Eppendorf.
3.6 Pipet the mixture into the DNeasy Mini spin column	All sample including precipitate placed into spin column. Ensure spin column is also in a 2 ml collection tube.
3.7 Incubate at room temperature for 5 minutes	
3.8 Centrifuge at full speed (17 G) for 3 minutes	
3.9 Discard flow-through but not collection tube	Tap collection tube onto blue paper to make sure you get rid of any excess liquid. Then place collection tube back on.
3.10 Centrifuge at full speed (17 G) for 1 minute	
3.11 Discard flow-through but not collection tube	Tap collection tube onto blue paper to make sure you get rid of any excess liquid. Then place collection tube back on.
4. Wash steps:	
4.1 Wash 1: Add 500 µl Buffer AW1	Invert tube to ensure was buffer has contacted all surfaces.
4.2 Incubate at room temperature for 5 minutes	
4.3 Centrifuge at full speed (17 G) for 3 minutes	Centrifuge for 3 minutes to ensure all liquid has passed through.
4.4 Discard flow-through but not collection tube	Tap collection tube onto blue paper to make sure you get rid of any excess liquid. Then place collection tube back on.
4.5 Centrifuge at full speed (17 G) for 1 minute	
4.6 Discard flow-through but not collection tube	Tap collection tube onto blue paper to make sure you get rid of any excess liquid. Then place collection tube back on.
4.7 Wash 2: Add 500µl Buffer AW2	Invert tube to ensure was buffer has contacted all surfaces.
4.8 Incubate at room temperature for 5 minutes	
4.9 Centrifuge at full speed (17 G) for 3 minutes	Centrifuge for 3 minutes to ensure all liquid has passed through.
4.10 Discard flow-through but not collection tube	Tap collection tube onto blue paper to make sure you get rid of any excess liquid. Then place collection tube back on.
4.11 Centrifuge at full speed (17 G) for 1 minute	
4.12 Discard flow-through but not collection tube	Tap collection tube onto blue paper to make sure you get rid of any excess liquid. Then place collection tube back on.

4.13 Wash 2 (Repeat): Add 500µl Buffer AW2	Invert tube to ensure was buffer has contacted all surfaces.
4.14 Incubate at room temperature for 5 minutes	
4.15 Centrifuge at full speed (17 G) for 3 minutes	Centrifuge for 3 minutes to ensure all liquid has passed through.
4.16 Discard flow-through but not collection tube	Tap collection tube onto blue paper to make sure you get rid of any excess liquid. Then place collection tube back on.
4.17 Centrifuge at full speed (17 G) for 3 minutes	Centrifuge for 3 minutes to ensure all liquid has passed through.
4.18 Place the DNeasy Mini spin column in a new 2 ml collection tube	
4.19 Incubate at room temperature for 15 minutes with lids open on spin column	It is important to dry the membrane of the DNeasy Mini spin column since residual ethanol may interfere with subsequent reactions.
5. Elution:	
5.1 Elution 1: Add 50µl of Tris-HCl [10mM], pH 8.4	Leave for 3 minutes
5.2 Centrifuge for 1 minute (17G)	
5.3 Elution 2: Add 50µl of Tris-HCl	Leave for 3 minutes
5.4 Centrifuge for 1 minute (17G)	
5.5 Transfer eluted liquid from collection tube into Eppendorf	Place 20µl in a separate one for quantification. Freeze at -80°C and avoid all unnecessary freeze thawing.

Appendix B: Chapter 2

Appendix Table B.1: List of phenotypic traits. 46 phenotypic traits were collected across a collection of RSSC (n=194). Each trait is ecologically relevant to RSSC lifecycle, either involved in metabolism, stress tolerance (abiotic and biotic) and virulence. Each trait is also grouped into 13 phenotypic groups for easier comparisons.

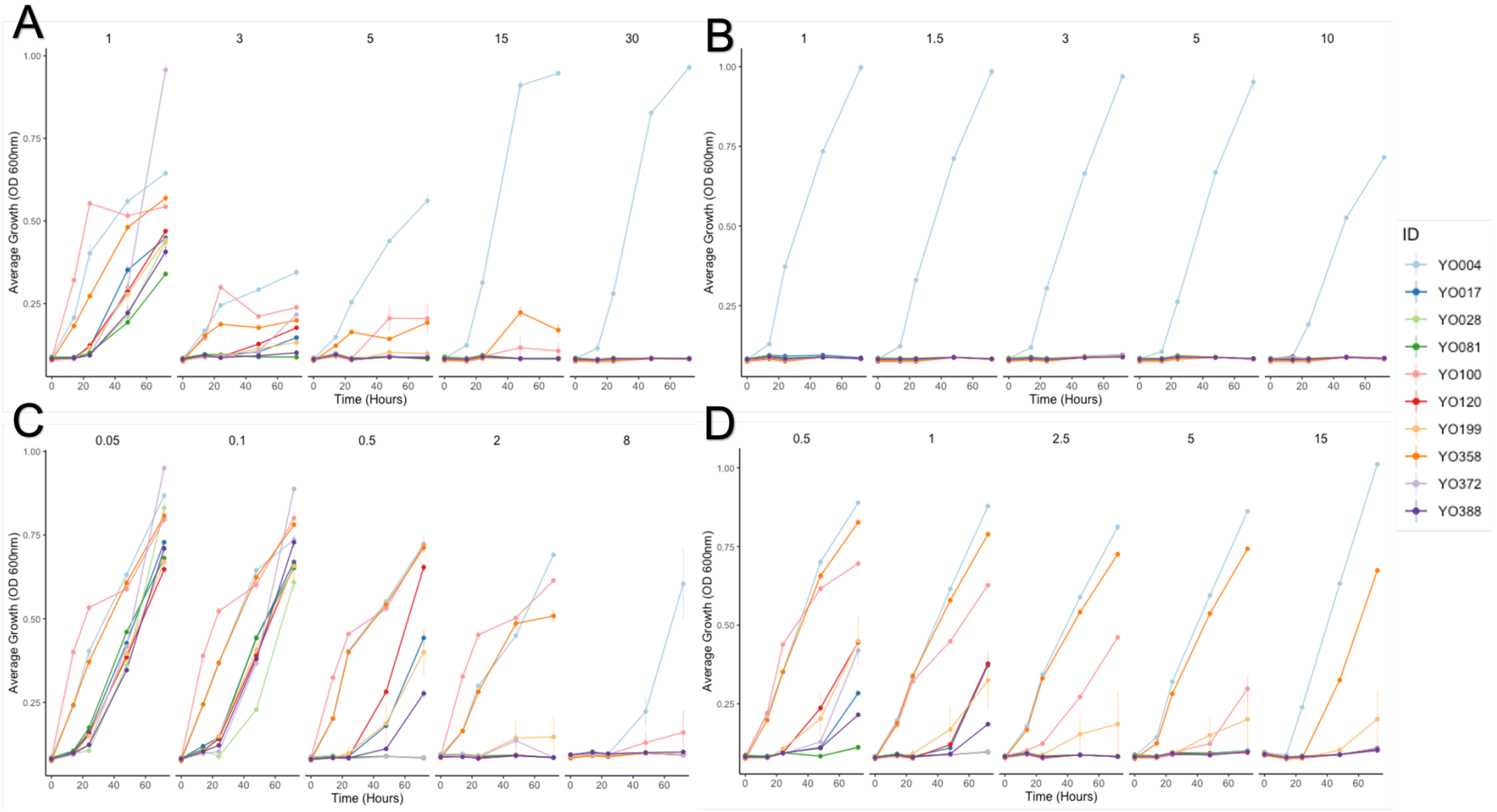
Theme	Phenotype Group	Trait Name	More Information
Metabolism	Complex Media	SP	Area under the growth curve in SP media with 24 hour reads up to 5 days of growth
Metabolism	Complex Media	CS	Area under the growth curve in CS media with 24 hour reads up to 5 days of growth
Metabolism	Complex Media	CPG	Area under the growth curve in CPG media with 24 hour reads up to 5 days of growth
Metabolism	Complex Media	NB	Area under the growth curve in NB media with 24 hour reads up to 5 days of growth
Metabolism	Single Carbons	Arabinose	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Asparagine	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	CitricAcid	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Glucose	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Glutamine	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.

Metabolism	Single Carbons	Glycine	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Histidine	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Maltose	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	MalicAcid	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Nicotinamide	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Proline	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Serine	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Sorbitol	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	SuccinicAcid	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Sucrose	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Xylose	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.

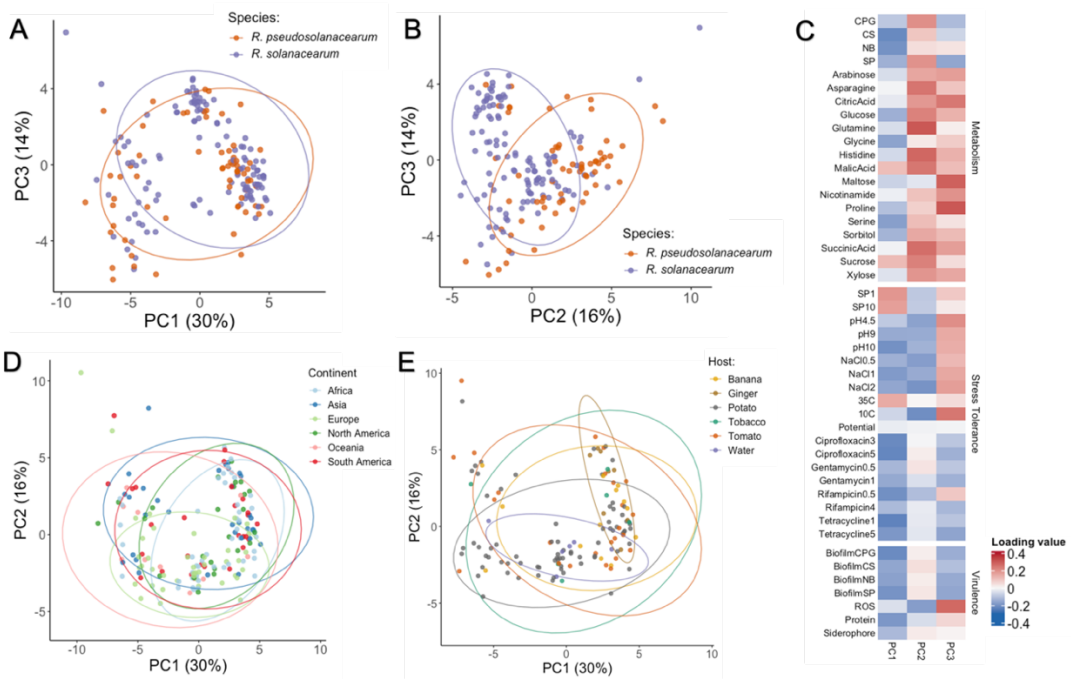
Abiotic Stress Tolerance	Nutrient Limited	SP1	Relative growth in nutrient limited conditions. Area under the curve (auc) in SP media diluted with distilled water 100-fold (1% SP) divided by auc in 100% SP. Auc calculated for 5 days of growth with OD reads every 24 hours.
Abiotic Stress Tolerance	Nutrient Limited	SP10	Relative growth in nutrient limited conditions. Area under the curve (auc) in SP media diluted with distilled water 10-fold (10% SP) divided by auc in 100% SP. Auc calculated for 5 days of growth with OD reads every 24 hours.
Abiotic Stress Tolerance	pH	pH4.5	Relative growth in pH 4.5 compared to pH 7, calculated by dividing the area under the growth curve (across 5 days) in the stress condition by the no stress condition. SP media was used.
Abiotic Stress Tolerance	pH	pH9	Relative growth in pH 9 compared to pH 7, calculated by dividing the area under the growth curve (across 5 days) in the stress condition by the no stress condition. SP media was used.
Abiotic Stress Tolerance	pH	pH10	Relative growth in pH 10 compared to pH 7, calculated by dividing the area under the growth curve (across 5 days) in the stress condition by the no stress condition. SP media was used.
Abiotic Stress Tolerance	Salinity	NaCl0.5	Relative growth in 0.5% NaCl compared to 0% NaCl, calculated by dividing the area under the growth curve (across 5 days) in the stress condition by the no stress condition. SP media was used.
Abiotic Stress Tolerance	Salinity	NaCl1	Relative growth in 1% NaCl compared to 0% NaCl, calculated by dividing the area under the growth curve (across 5 days) in the stress condition by the no stress condition. SP media was used.
Abiotic Stress Tolerance	Salinity	NaCl2	Relative growth in 2% NaCl compared to 0% NaCl, calculated by dividing the area under the growth curve (across 5 days) in the stress condition by the no stress condition. SP media was used.
Abiotic Stress Tolerance	Heat Tolerance	35C	Relative growth at 35C compared to 28C, calculated by dividing the area under the growth curve (across 5 days) in the stress condition by the no stress condition. SP media was used.
Abiotic Stress Tolerance	Cold Tolerance	10C	Relative growth at 10C compared to 28C, calculated by dividing the area under the growth curve (across 5 days) in the stress condition by the no stress condition. SP media was used.

Abiotic Stress Tolerance	Potential	Potential	Water potential tolerance. Relative growth in 15% polyethylene glycol (PEG)-4000 compared to 0% polyethylene glycol, calculated by dividing the area under the growth curve up to 5 days of growth of stress condition by no stress condition. SP media was used.
Biotic Stress Tolerance	Antibiotic Resistance	Gentamycin0.5	Relative growth in 0.5 ug/ml of Gentamycin antibiotic, calculated by dividing the OD at 600nm after 48 hours of growth in stress condition by no stress condition. SP media was used.
Biotic Stress Tolerance	Antibiotic Resistance	Gentamycin1	Relative growth in 1 ug/ml of Gentamycin antibiotic, calculated by dividing the OD at 600nm after 48 hours of growth in stress condition by no stress condition. SP media was used.
Biotic Stress Tolerance	Antibiotic Resistance	Tetracycline1	Relative growth in 1 ug/ml of Tetracycline antibiotic, calculated by dividing the OD at 600nm after 48 hours of growth in stress condition by no stress condition. SP media was used.
Biotic Stress Tolerance	Antibiotic Resistance	Tetracycline5	Relative growth in 5 ug/ml of Tetracycline antibiotic, calculated by dividing the OD at 600nm after 48 hours of growth in stress condition by no stress condition. SP media was used.
Biotic Stress Tolerance	Antibiotic Resistance	Rifampicin0.5	Relative growth in 0.5 ug/ml of Rifampicin antibiotic, calculated by dividing the OD at 600nm after 48 hours of growth in stress condition by no stress condition. SP media was used.
Biotic Stress Tolerance	Antibiotic Resistance	Rifampicin4	Relative growth in 4 ug/ml of Rifampicin antibiotic, calculated by dividing the OD at 600nm after 48 hours of growth in stress condition by no stress condition. SP media was used.
Biotic Stress Tolerance	Antibiotic Resistance	Ciprofloxacin3	Relative growth in 3 ug/ml of Ciprofloxacin antibiotic, calculated by dividing the OD at 600nm after 48 hours of growth in stress condition by no stress condition. SP media was used.
Biotic Stress Tolerance	Antibiotic Resistance	Ciprofloxacin5	Relative growth in 5 ug/ml of Ciprofloxacin antibiotic, calculated by dividing the OD at 600nm after 48 hours of growth in stress condition by no stress condition. SP media was used.
Virulence	Biofilm	BiofilmCPG	Biofilm quantification after growth in CPG media for 7 days. Crystal violet assay used.
Virulence	Biofilm	BiofilmSP	Biofilm quantification after growth in SP media for 7 days. Crystal violet assay used.

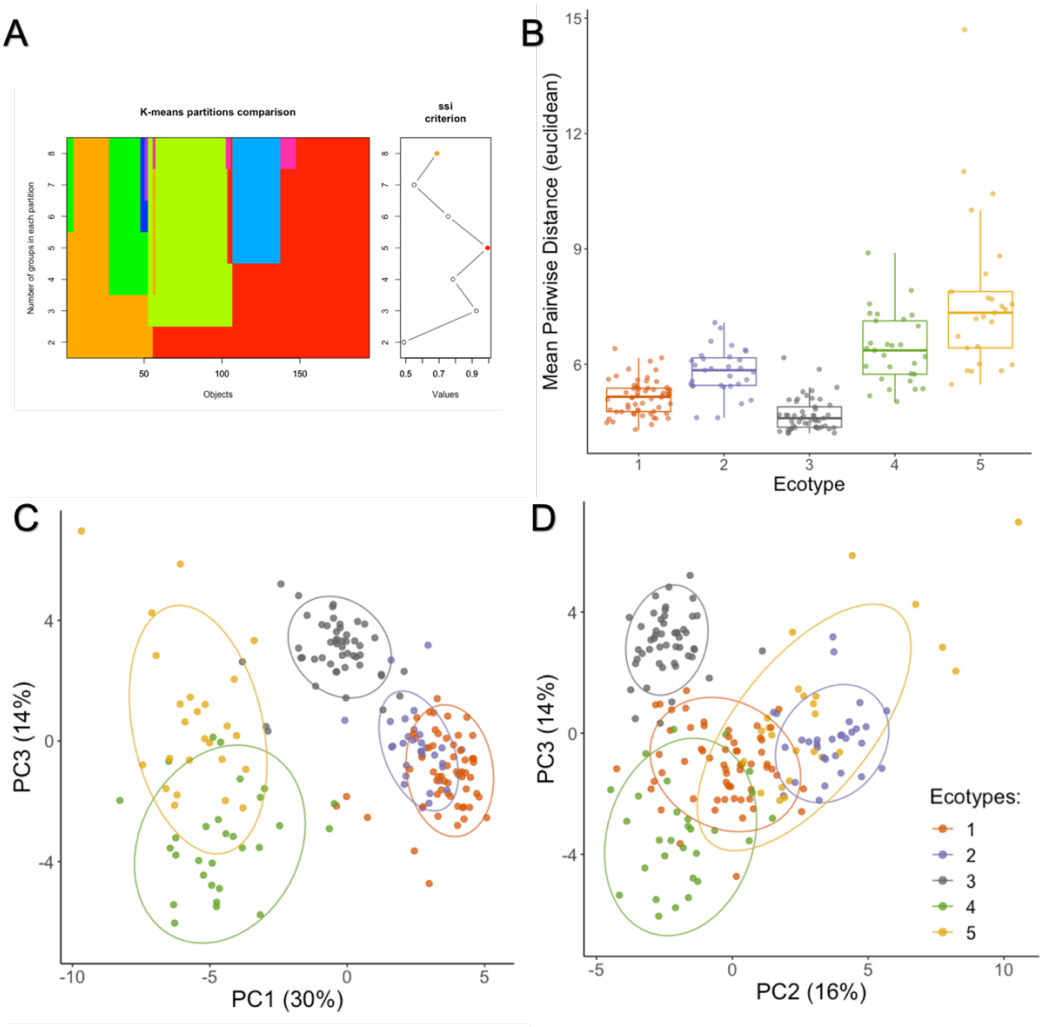
Virulence	Biofilm	BiofilmCS	Biofilm quantification after growth in CS media for 7 days. Crystal violet assay used.
Virulence	Biofilm	BiofilmNB	Biofilm quantification after growth in NB media for 7 days. Crystal violet assay used.
Virulence	ROS	ROS	Oxidative stress tolerance from reactive oxygen species (ROS). Relative growth in 100mM H_2O_2 compared to 0mM, calculated by dividing the area under the growth curve up to 5 days of growth of stress condition by no stress condition. SP media was used.
Virulence	Siderophore	Siderophore	Siderophore production (psu) per cell. Siderophore production was quantified using the CAS assay protocol using bacteria supernatant after 3 days of growth in SP media. Psu per cell was then calculated by dividing psu by growth (OD at 600nm).
Virulence	Protein	Protein	Extracellular protein produced per cell. Protein in each isolate's supernatant was quantified using Bradford assay technique after 72 hours of growth in SP media. Protein amount (ug/ml) was then divided by the growth (OD 600nm) to determine protein produced per cell.


A
1
3
5
15
30
B
1
1.5
3
5
10
C
0.05
0.1
0.5
2
8
D
0.5
1
2.5
5
15
Average Growth (OD 600nm)
Time (Hours)
ID
YO004
YO017
YO028
YO081
YO100
YO120
YO199
YO358
YO372
YO388

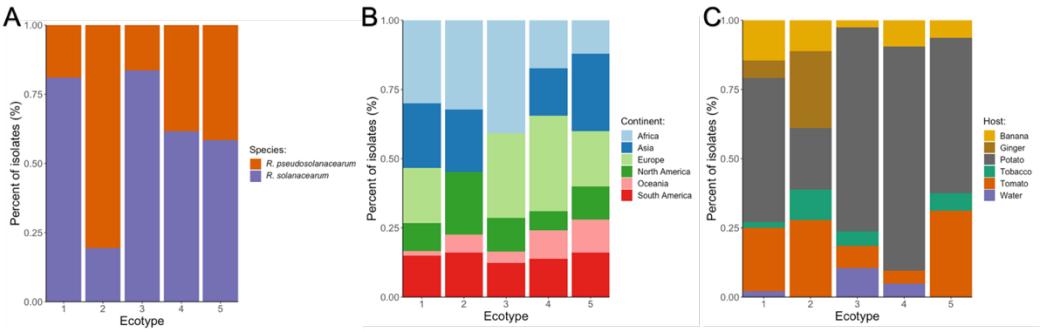
Appendix Figure B.1: Preliminary work on a subset of 10 RSSC isolates was conducted to choose 'low' and 'high' antibiotic concentrations for phenotyping. Numbers at the top indicate antibiotic concentrations in μg/ml for (A) ciprofloxacin (B) gentamycin (C) rifampicin and (D) tetracycline. Concentrations chosen for the phenotyping were as follows; 3 and 5μg/ml for ciprofloxacin, 0.5 and 1 μg/ml for gentamycin, 0.5 and 4μg/ml for rifampicin, and 1 and 5μg/ml for tetracycline. Error bars show one standard error from the mean (N=6).



Appendix Figure B.2: *Ralstonia solanacearum* species complex (RSSC) phenotypic diversity cannot be explained fully by species, continent sampled from or host isolated from. (A) PCA plot with PC1 (29.6%) and PC3 (13.7%). Each point represents a isolate in the RSSC collection, coloured by species assigned by whole genome sequencing (n=185). (B) PCA plot with PC2 (16.5%) and PC3 (13.7%). (C) Heatmap of loadings, contribution each trait has on the three PCs. (D) PCA plot with PC1 (29.6%) and PC2 (16.5%). Each point represents a isolate in the RSSC collection, coloured by continent sampled from (n=185). (E) PCA plot with PC1 (29.6%) and PC2 (16.5%). Each point represents a isolate in the RSSC collection, coloured by host isolated from (n=133). Eclipses show 90% confidence intervals around the centroid of each group.



Appendix Figure B.3: *Ralstonia solanacearum* species complex (RSSC) clusters into five separate ecotypes. (A) ssi criterion showing optimal number of clusters as 5 for k-means clustering, providing evidence for 5 ecotypes. (B) Mean pairwise distance (Euclidean distance) from all isolates grouped by the five ecotypes. Boxplots lines show the median per ecotype of their mean pairwise distances with the box indicating the interquartile range and whiskers showing the 95% quantile range. (C) PCA plot with PC1 (29.6%) and PC3 (13.7%). Each point represents a isolate in the RSSC collection (n=194), coloured by ecotype assigned by k-means clustering (k=5). (D) PCA plot with PC2 (16.5%) and PC3 (13.7%). Each point represents a isolate in the RSSC collection (n=194), coloured by ecotype assigned by k-means clustering (k=5). Eclipses show 90% confidence intervals around the centroid of each group.



Appendix Figure B.4: Continent, host and species all vary slightly in their distribution across the five ecotypes. (A) The distribution of isolates within each ecotype belonging to the two separate RSSC species within this study. The distribution of isolates per ecotype isolated from different continents (B) and hosts (C) respectively.

Supplementary Table B.2: Cold tolerance traits significant genetic associations. Associations were determined using a GWAS technique with gene presence/absence as the input genetic variable. Phenotypic genetic variables were binary_cold_survival, where growth (1) or no growth (0) was determined on a CPG agar plate after five days at 4°C. Also, binary_low_temperature was determined by growth (1), optical density at 600nm (OD_{600}) over 0.13, and no growth (0), OD_{600} below 0.13, after 5 days of incubation at 10°C.

binary_cold_survival					
COGs					
Gene Name	**Filter p value**	**lrt p value**	**Allele frequency**	**Beta**	**Variant h2**
hxcR_2~~~epsE	2.28E-07	2.56E-05	8.42E-01	3.68E-01	2.21E-01
binary_low_temp					
COGs					
Gene Name	**Filter p value**	**lrt p value**	**Allele frequency**	**Beta**	**Variant h2**
hxcR_2~~~epsE	1.27E-08	9.68E-06	8.42E-01	4.11E-01	2.32E-01
group_4570	2.94E-08	2.22E-05	8.39E-01	3.93E-01	2.23E-01
epsF_3~~~epsF_4	6.61E-08	4.67E-05	8.37E-01	3.78E-01	2.14E-01
group_4620	2.69E-08	1.98E-05	8.34E-01	4.05E-01	2.24E-01

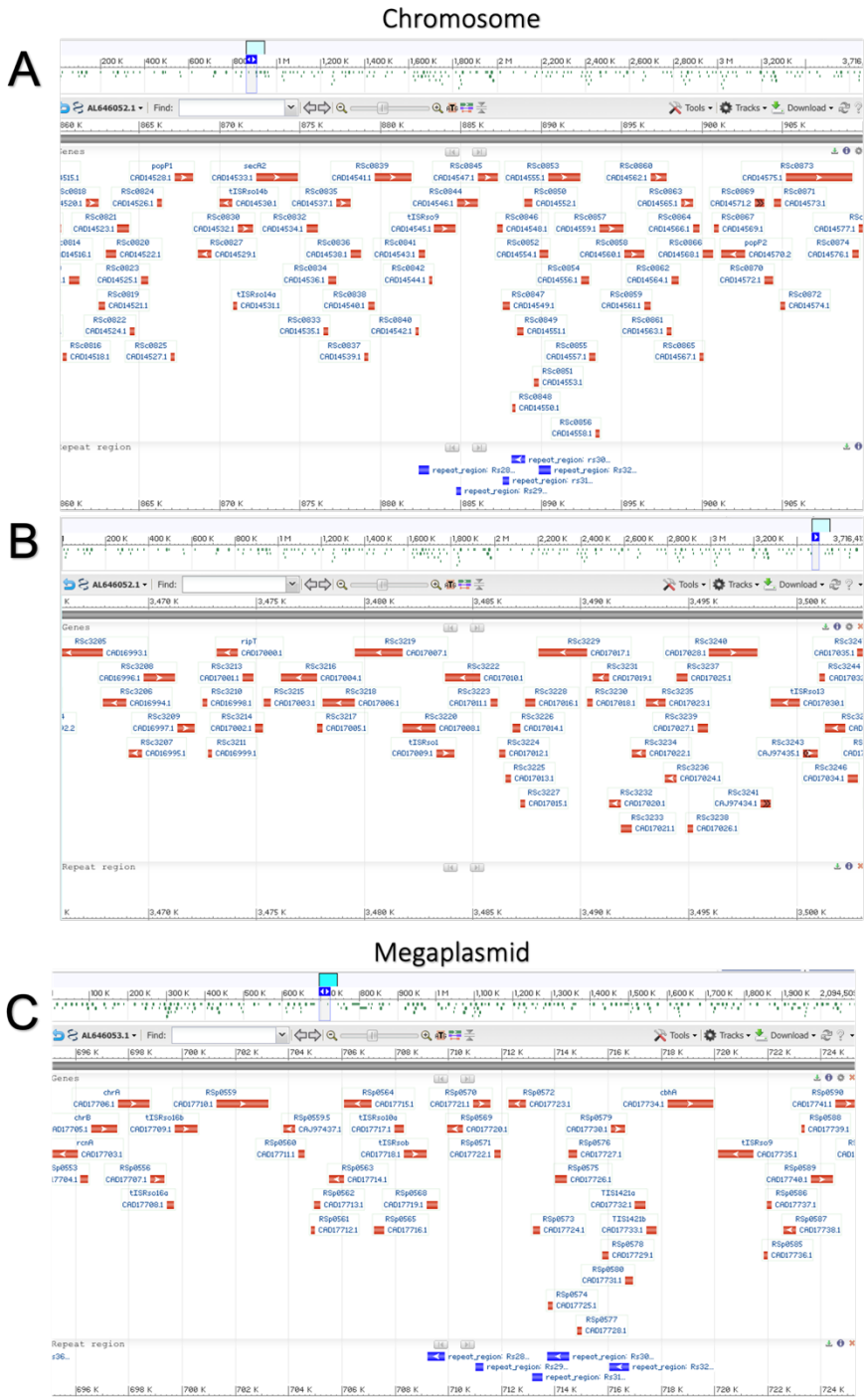
Appendix Table B.3: Rifampicin resistance at 0.5µg/ml (binary trait) significant hits determined using a GWAS technique with COG and unitigs as the input genetic variables.

binary_rifampicin_0.5					
COGs					
Gene Name	**Filter p value**	**lrt p value**	**Allele frequency**	**Beta**	**Variant h2**
group_6769	0.0000	0.0000	0.7770	0.6850	0.2810
pcpR_6~~~pcpR_5~~~pcpR_4	0.0000	0.0000	0.7630	0.7440	0.2630
gltR_2	0.0000	0.0000	0.7630	0.7440	0.2630
group_1663	0.0000	0.0000	0.7610	0.5950	0.2200
group_1875	0.0000	0.0000	0.7300	0.7020	0.2800
group_8240	0.0000	0.0000	0.7210	0.8810	0.2970
group_6641	0.0000	0.0000	0.2820	-0.4880	0.2970
group_6653	0.0000	0.0000	0.2650	-0.4570	0.2880
group_4170	0.0000	0.0000	0.2820	-0.6020	0.3350
group_3539	0.0000	0.0000	0.2650	-0.3450	0.2240
group_3648	0.0000	0.0000	0.2480	-0.3760	0.2560
group_3899	0.0000	0.0000	0.2390	-0.4140	0.2880
group_6634	0.0000	0.0000	0.2230	-0.3560	0.2470
group_6627	0.0000	0.0000	0.2170	-0.3310	0.2310
group_7058	0.0000	0.0000	0.2510	-0.3960	0.2410
group_5647	0.0000	0.0000	0.2060	-0.3520	0.2560
group_4715	0.0000	0.0000	0.2310	-0.4010	0.2880
group_4091	0.0000	0.0000	0.2340	-0.3140	0.2270
group_4260	0.0000	0.0000	0.2000	-0.2990	0.2190
group_3946	0.0000	0.0000	0.2340	-0.3480	0.2570
group_4139	0.0000	0.0000	0.2420	-0.3610	0.2360
group_2885	0.0000	0.0000	0.2140	-0.4760	0.3140
group_4277	0.0000	0.0000	0.2170	-0.3940	0.2860
group_6655	0.0000	0.0000	0.2170	-0.4770	0.3130
group_4241	0.0000	0.0000	0.2000	-0.4180	0.2910
group_4662	0.0000	0.0000	0.2200	-0.4580	0.2940
group_7251	0.0000	0.0000	0.1770	-0.3960	0.2620
group_6923	0.0000	0.0000	0.2060	-0.4200	0.2700
fitB	0.0000	0.0000	0.1940	-0.3660	0.2490
group_6722	0.0000	0.0000	0.1720	-0.3550	0.2550
group_7366	0.0000	0.0000	0.1630	-0.4480	0.3040
group_6925	0.0000	0.0000	0.1410	-0.4800	0.2990
group_3326	0.0000	0.0000	0.1580	-0.4220	0.2830
group_7354	0.0000	0.0000	0.1830	-0.3490	0.2240
group_3802	0.0000	0.0000	0.1720	-0.3150	0.2230
group_6912	0.0000	0.0000	0.1630	-0.3320	0.2280

group_4962	0.0000	0.0000	0.1770	-0.4140	0.2820
group_4121	0.0000	0.0000	0.1490	-0.3450	0.2530
group_3664	0.0000	0.0000	0.1070	-0.4100	0.2610
UNITIGS					
Gene Name	**Hits**	**Maximum -log10(p value)**	**Average allele frequency**	**Average minor allele frequency**	**Average beta**
cds-WP_039562627.1	97	16.0590	0.0710	0.0710	0.6347
cds-WP_039562510.1	94	15.0381	0.0791	0.0791	0.6103
cds-WP_039562639.1	69	15.5258	0.0841	0.0841	0.5666
cds-WP_039563187.1	50	13.9872	0.0690	0.0690	0.5948
cds-WP_039559476.1	40	9.2403	0.0763	0.0763	0.5301
cds850	37	15.4711	0.1329	0.1329	0.5689
cds-WP_039562505.1	36	14.2396	0.0620	0.0620	0.6343
cds-WP_039562644.1	35	12.2388	0.0677	0.0677	0.5766
cds-WP_039563184.1	34	12.8894	0.0769	0.0769	0.5699
cds854	33	15.4815	0.1075	0.1075	0.5671
cds-WP_039562646.1	33	15.5436	0.0764	0.0764	0.6163
cds-WP_039562516.1	30	15.5719	0.0697	0.0697	0.6063
cds858	26	13.5560	0.1464	0.1464	0.5172
cds-WP_039563186.1	26	10.0119	0.0798	0.0798	0.5710
cds-WP_039562624.1	25	11.1530	0.0587	0.0587	0.5649
cds-WP_039562504.1	25	17.4425	0.0765	0.0765	0.6916
cds874	23	16.5969	0.1118	0.1118	0.5888
cds-WP_039562499.1	23	14.1481	0.0657	0.0657	0.6323
cds867	22	15.7570	0.0951	0.0951	0.6048
cds-WP_039562664.1	20	9.7670	0.0736	0.0736	0.5152
cds-WP_039562526.1	20	15.3233	0.0634	0.0634	0.6845
cds-WP_039562650.1	19	13.1938	0.0707	0.0707	0.5451
cds-WP_039562672.1	18	11.3958	0.0728	0.0728	0.5289
cds-WP_039562654.1	18	10.0195	0.0731	0.0731	0.5345
cds-WP_039562519.1	18	9.9586	0.0699	0.0699	0.5602
cds852	17	14.8539	0.1578	0.1578	0.5859
cds866	16	11.7352	0.1845	0.1845	0.5366
cds-WP_039562629.1	16	11.7235	0.0632	0.0632	0.6036
cds-WP_039562673.1	15	11.7033	0.0803	0.0803	0.5236
cds873	14	15.0809	0.1359	0.1359	0.5538
cds3215	14	11.7352	0.1597	0.1597	0.5354
cds-WP_039559430.1	14	11.1624	0.1079	0.1079	0.5114
cds-WP_039562631.1	13	10.2950	0.0756	0.0756	0.5129
cds880	11	9.4473	0.1098	0.1098	0.4958
cds3234	11	13.4214	0.1122	0.1122	0.5796

cds871	10	10.8928	0.0775	0.0775	0.5647
cds-RSUY_RS22605	10	8.4437	0.0676	0.0676	0.5317
cds-WP_039562621.1	10	13.2882	0.0822	0.0822	0.5393
cds-WP_039562645.1	10	15.0809	0.0874	0.0874	0.5220
cds-WP_039562514.1	10	14.6635	0.0583	0.0583	0.6751
cds888	9	7.9431	0.1031	0.1031	0.4998
cds856	9	12.0675	0.1424	0.1424	0.6057
cds872	9	12.0372	0.1523	0.1523	0.4932
cds-WP_039559478.1	9	10.8477	0.0779	0.0779	0.5473
cds-WP_039562638.1	9	11.9586	0.0632	0.0632	0.5627
cds4024	8	8.9547	0.1781	0.1781	0.4890
cds-WP_003278075.1	8	8.4584	0.0968	0.0968	0.5504
cds-WP_039562508.1	8	12.4584	0.0651	0.0651	0.5623
cds4022	7	12.0675	0.1662	0.1662	0.5884
cds4033|cds4034	7	11.0487	0.2297	0.2297	0.6310
cds4033	7	11.7352	0.1622	0.1622	0.5129
cds855	6	9.7773	0.0798	0.0798	0.5197
cds876	5	12.4101	0.1928	0.1928	0.5172
cds866|cds867	5	11.0487	0.2196	0.2196	0.6580
cds878	5	10.0150	0.1380	0.1380	0.5060
cds-WP_039562627.1|cds-WP_039562624.1	5	14.3665	0.0727	0.0727	0.7056
cds-WP_039563185.1|cds-WP_014615737.1	5	11.4935	0.0552	0.0552	0.7000
cds884	4	7.5986	0.1085	0.1085	0.5418
cds864	4	9.4908	0.1784	0.1784	0.5503
cds4031	4	9.4908	0.1784	0.1784	0.5503
cds847	4	9.3019	0.1550	0.1550	0.4440
cds859	4	9.8601	0.0528	0.0528	0.5875
cds868	4	14.1018	0.0697	0.0697	0.6383
cds-WP_003278076.1	4	7.6383	0.0957	0.0957	0.5233
cds-WP_039562672.1|cds-WP_039562664.1	4	8.9547	0.0571	0.0571	0.5435
cds-WP_039562631.1|cds-WP_039562629.1	4	10.1811	0.0662	0.0662	0.6423
cds870	3	11.5560	0.1277	0.1277	0.6480
cds849|cds850	3	12.5143	0.1672	0.1672	0.5550
cds889	3	7.9508	0.0901	0.0901	0.5613
cds849	3	16.3675	0.0985	0.0985	0.7403
cds-WP_039559427.1|cds-WP_039562672.1	3	9.4522	0.0582	0.0582	0.5310
cds-WP_039562623.1	3	9.9666	0.0761	0.0761	0.5807
cds-WP_039562502.1|cds-WP_039562504.1	3	13.7959	0.0620	0.0620	0.7540

cds-WP_039562523.1	3	9.6635	0.0516	0.0516	0.5940
cds4029	2	7.3788	0.0997	0.0997	0.5255
cds3218	2	8.8665	0.2610	0.2610	0.5345
cds877	2	7.8996	0.0774	0.0774	0.5045
cds851	2	8.5513	0.1154	0.1154	0.5060
malQ	2	7.2495	0.1028	0.1028	0.6315
cds855|cds856	2	11.7645	0.0564	0.0564	0.6740
cds853|cds854	2	8.4921	0.0718	0.0718	0.5125
cds869	2	11.4535	0.0577	0.0577	0.6545
cds875	2	8.9626	0.0831	0.0831	0.4950
cds-B7R79_RS16590	2	14.7773	0.0690	0.0690	0.7535
cds-WP_003265739.1	2	13.9872	0.0999	0.0999	0.7480
cds-WP_039562638.1|cds-WP_039562631.1	2	8.8633	0.0747	0.0747	0.5290
cds-RSUY_RS22855	2	13.7959	0.0620	0.0620	0.7335
cds-WP_039562508.1|cds-WP_039562510.1	2	10.3325	0.0747	0.0747	0.5730
cds-WP_039562505.1|cds-WP_039562508.1	2	12.8153	0.0690	0.0690	0.7425
cds-WP_039562502.1	2	14.1135	0.0606	0.0606	0.7855
cds862	1	7.0182	0.0535	0.0535	0.5260
cds891	1	7.7235	0.0704	0.0704	0.6170
cds886	1	8.5952	0.0704	0.0704	0.5430
cds2545|cds2546	1	6.9914	0.0535	0.0535	0.4430
cds874|cds875	1	8.9208	0.0873	0.0873	0.5440
cds890	1	7.2890	0.0507	0.0507	0.5710
cds892	1	7.1180	0.0761	0.0761	0.4820
cds3203	1	7.1180	0.0761	0.0761	0.4820
cds887	1	7.8356	0.0648	0.0648	0.5140
cds884|cds885	1	8.1624	0.0789	0.0789	0.5890
cds-WP_003265723.1	1	7.6946	0.0507	0.0507	0.5260
cds-WP_039562650.1|cds-WP_039562646.1	1	11.5607	0.0761	0.0761	0.6640
cds-WP_039562512.1|cds-WP_039562514.1	1	14.8861	0.0507	0.0507	0.8390
cds-WP_014615737.1	1	7.0301	0.0507	0.0507	0.5430
cds-WP_039562624.1|cds-WP_039562623.1	1	7.4283	0.0930	0.0930	0.5190
cds-WP_039562645.1|cds-WP_039562644.1	1	8.5331	0.0507	0.0507	0.6000
cds-WP_039563185.1	1	10.2823	0.0507	0.0507	0.7680
cds-WP_047942537.1	1	9.0570	0.0648	0.0648	0.6400



Appendix Figure B.5: Rifampicin resistance (0.5 g/ml) gene region hits identified in GWAS. Significantly associated gene regions, identified in the GWAS, was viewed in GMI1000 RSSC isolate (accession number: 000009125.1) on NCBI (https://www.ncbi.nlm.nih.gov/assembly/). These regions include between (A) 870 – 900 kb and (B) 3.47 – 3.5 Mb within the chromosome, along with between (C) 870 – 900 kb within the megaplasmid.

Appendix C: Chapter 3

Appendix Table C.1: List of phenotypic traits. 46 phenotypic traits were collected across a collection of *R. solanacearum* isolates (n=182). Each trait is ecologically relevant to the lifecycvle of *R. solanacearum*, either involved in metabolism, stress tolerance (abiotic and biotic) and virulence. Each trait is also grouped into 13 phenotypic groups for easier comparisons.

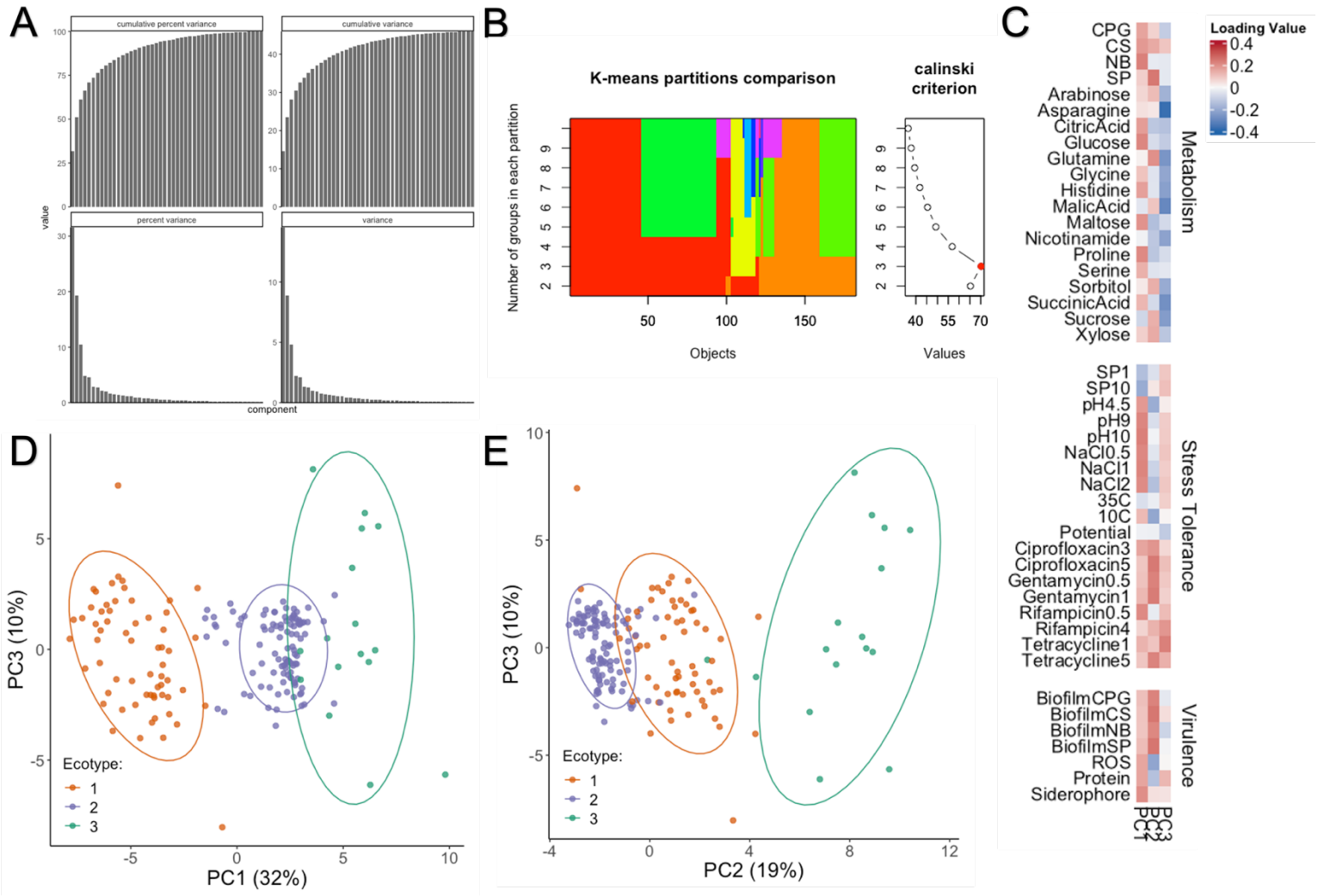
Theme	Phenotype Group	Trait Name	More Information
Metabolism	Complex Media	SP	Area under the growth curve in SP media with 24 hour reads up to 5 days of growth
Metabolism	Complex Media	CS	Area under the growth curve in CS media with 24 hour reads up to 5 days of growth
Metabolism	Complex Media	CPG	Area under the growth curve in CPG media with 24 hour reads up to 5 days of growth
Metabolism	Complex Media	NB	Area under the growth curve in NB media with 24 hour reads up to 5 days of growth
Metabolism	Single Carbons	Arabinose	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Asparagine	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	CitricAcid	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Glucose	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Glutamine	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.

Metabolism	Single Carbons	Glycine	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Histidine	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Maltose	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	MalicAcid	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Nicotinamide	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Proline	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Serine	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Sorbitol	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	SuccinicAcid	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Sucrose	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.
Metabolism	Single Carbons	Xylose	Growth in 10mM of single carbon resource. Growth measured by the area under the growth curve (auc), up to 5 days with 24 hour reads. auc was then divided by the auc of that isolate in no carbon media.

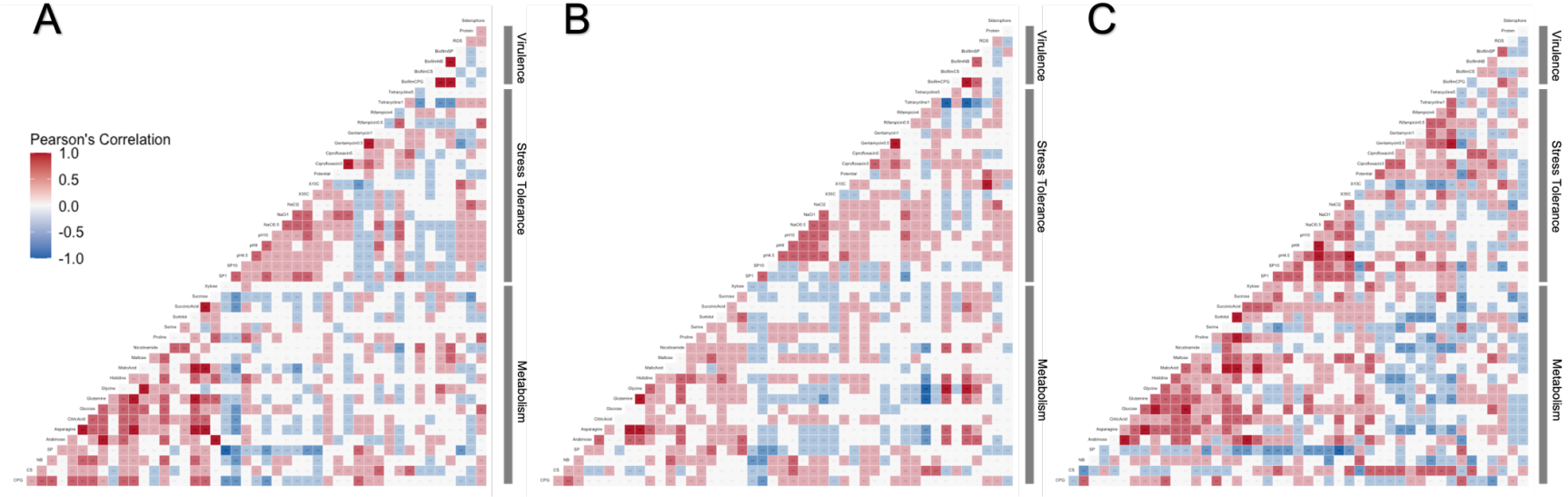
Abiotic Stress Tolerance	Nutrient Limited	SP1	Relative growth in nutrient limited conditions. Area under the curve (auc) in SP media diluted with distilled water 100-fold (1% SP) divided by auc in 100% SP. Auc calculated for 5 days of growth with OD reads every 24 hours.
Abiotic Stress Tolerance	Nutrient Limited	SP10	Relative growth in nutrient limited conditions. Area under the curve (auc) in SP media diluted with distilled water 10-fold (10% SP) divided by auc in 100% SP. Auc calculated for 5 days of growth with OD reads every 24 hours.
Abiotic Stress Tolerance	pH	pH4.5	Relative growth in pH 4.5 compared to pH 7, calculated by dividing the area under the growth curve (across 5 days) in the stress condition by the no stress condition. SP media was used.
Abiotic Stress Tolerance	pH	pH9	Relative growth in pH 9 compared to pH 7, calculated by dividing the area under the growth curve (across 5 days) in the stress condition by the no stress condition. SP media was used.
Abiotic Stress Tolerance	pH	pH10	Relative growth in pH 10 compared to pH 7, calculated by dividing the area under the growth curve (across 5 days) in the stress condition by the no stress condition. SP media was used.
Abiotic Stress Tolerance	Salinity	NaCl0.5	Relative growth in 0.5% NaCl compared to 0% NaCl, calculated by dividing the area under the growth curve (across 5 days) in the stress condition by the no stress condition. SP media was used.
Abiotic Stress Tolerance	Salinity	NaCl1	Relative growth in 1% NaCl compared to 0% NaCl, calculated by dividing the area under the growth curve (across 5 days) in the stress condition by the no stress condition. SP media was used.
Abiotic Stress Tolerance	Salinity	NaCl2	Relative growth in 2% NaCl compared to 0% NaCl, calculated by dividing the area under the growth curve (across 5 days) in the stress condition by the no stress condition. SP media was used.
Abiotic Stress Tolerance	Heat Tolerance	35C	Relative growth at 35C compared to 28C, calculated by dividing the area under the growth curve (across 5 days) in the stress condition by the no stress condition. SP media was used.
Abiotic Stress Tolerance	Cold Tolerance	10C	Relative growth at 10C compared to 28C, calculated by dividing the area under the growth curve (across 5 days) in the stress condition by the no stress condition. SP media was used.

Abiotic Stress Tolerance	Potential	Potential	Water potential tolerance. Relative growth in 15% polyethylene glycol (PEG)-4000 compared to 0% polyethylene glycol, calculated by dividing the area under the growth curve up to 5 days of growth of stress condition by no stress condition. SP media was used.
Biotic Stress Tolerance	Antibiotic Resistance	Gentamycin0.5	Relative growth in 0.5 ug/ml of Gentamycin antibiotic, calculated by dividing the OD at 600nm after 48 hours of growth in stress condition by no stress condition. SP media was used.
Biotic Stress Tolerance	Antibiotic Resistance	Gentamycin1	Relative growth in 1 ug/ml of Gentamycin antibiotic, calculated by dividing the OD at 600nm after 48 hours of growth in stress condition by no stress condition. SP media was used.
Biotic Stress Tolerance	Antibiotic Resistance	Tetracycline1	Relative growth in 1 ug/ml of Tetracycline antibiotic, calculated by dividing the OD at 600nm after 48 hours of growth in stress condition by no stress condition. SP media was used.
Biotic Stress Tolerance	Antibiotic Resistance	Tetracycline5	Relative growth in 5 ug/ml of Tetracycline antibiotic, calculated by dividing the OD at 600nm after 48 hours of growth in stress condition by no stress condition. SP media was used.
Biotic Stress Tolerance	Antibiotic Resistance	Rifampicin0.5	Relative growth in 0.5 ug/ml of Rifampicin antibiotic, calculated by dividing the OD at 600nm after 48 hours of growth in stress condition by no stress condition. SP media was used.
Biotic Stress Tolerance	Antibiotic Resistance	Rifampicin4	Relative growth in 4 ug/ml of Rifampicin antibiotic, calculated by dividing the OD at 600nm after 48 hours of growth in stress condition by no stress condition. SP media was used.
Biotic Stress Tolerance	Antibiotic Resistance	Ciprofloxacin3	Relative growth in 3 ug/ml of Ciprofloxacin antibiotic, calculated by dividing the OD at 600nm after 48 hours of growth in stress condition by no stress condition. SP media was used.
Biotic Stress Tolerance	Antibiotic Resistance	Ciprofloxacin5	Relative growth in 5 ug/ml of Ciprofloxacin antibiotic, calculated by dividing the OD at 600nm after 48 hours of growth in stress condition by no stress condition. SP media was used.
Virulence	Biofilm	BiofilmCPG	Biofilm quantification after growth in CPG media for 7 days. Crystal violet assay used.
Virulence	Biofilm	BiofilmSP	Biofilm quantification after growth in SP media for 7 days. Crystal violet assay used.

Virulence	Biofilm	BiofilmCS	Biofilm quantification after growth in CS media for 7 days. Crystal violet assay used.
Virulence	Biofilm	BiofilmNB	Biofilm quantification after growth in NB media for 7 days. Crystal violet assay used.
Virulence	ROS	ROS	Oxidative stress tolerance from reactive oxygen species (ROS). Relative growth in 100mM H_2O_2 compared to 0mM, calculated by dividing the area under the growth curve up to 5 days of growth of stress condition by no stress condition. SP media was used.
Virulence	Siderophore	Siderophore	Siderophore production (psu) per cell. Siderophore production was quantified using the CAS assay protocol using bacteria supernatant after 3 days of growth in SP media. Psu per cell was then calculated by dividing psu by growth (OD at 600nm).
Virulence	Protein	Protein	Extracellular protein produced per cell. Protein in each isolate's supernatant was quantified using Bradford assay technique after 72 hours of growth in SP media. Protein amount (ug/ml) was then divided by the growth (OD 600nm) to determine protein produced per cell.


A
cumulative percent variance
cumulative variance
percent variance
variance
value
component
B
K-means partitions comparison
calinski criterion
Number of groups in each partition
Objects
Values
C
Loading Value
0.4
0.2
0
-0.2
-0.4
CPG
CS
NB
SP
Arabinose
Asparagine
CitricAcid
Glucose
Glutamine
Glycine
Histidine
MalicAcid
Maltose
Nicotinamide
Proline
Serine
Sorbitol
SuccinicAcid
Sucrose
Xylose
Metabolism
SP1
SP10
pH4.5
pH9
pH10
NaCl0.5
NaCl1
NaCl2
35C
10C
Potential
Ciprofloxacin3
Ciprofloxacin5
Gentamycin0.5
Gentamycin1
Rifampicin0.5
Rifampicin4
Tetracycline1
Tetracycline5
Stress Tolerance
BiofilmCPG
BiofilmCS
BiofilmNB
BiofilmSP
ROS
Protein
Siderophore
Virulence
PC3
PC2
PC1
D
PC3 (10%)
PC1 (32%)
Ecotype:
1
2
3
E
PC3 (10%)
PC2 (19%)
Ecotype:
1
2
3

Appendix Figure C.1: UK *R. solanacearum* clusters into three phenotypic groups or 'ecotypes'. (A) variance explained by each principal component (PC) in the principal component analysis (PCA) conducted on the 46 individual traits of the 182 UK isolates. (B) calinski criterion determines that the optimal cluster of phenotypic groups within the UK population to be 3. (C) Loadings of PC 1, 2 and 3 for the PCA plots. (D) PC1 and 3 plotted for each isolate (n=182) coloured by cluster assigned by k-means clustering. PC1 explains 32% of the total variation in the dataset and PC3 10%. (E) PC2 and 3 plotted for each isolate (n=182) coloured by cluster assigned by k-means clustering. PC2 explains 19% of the total variation in the dataset and PC3 10%.

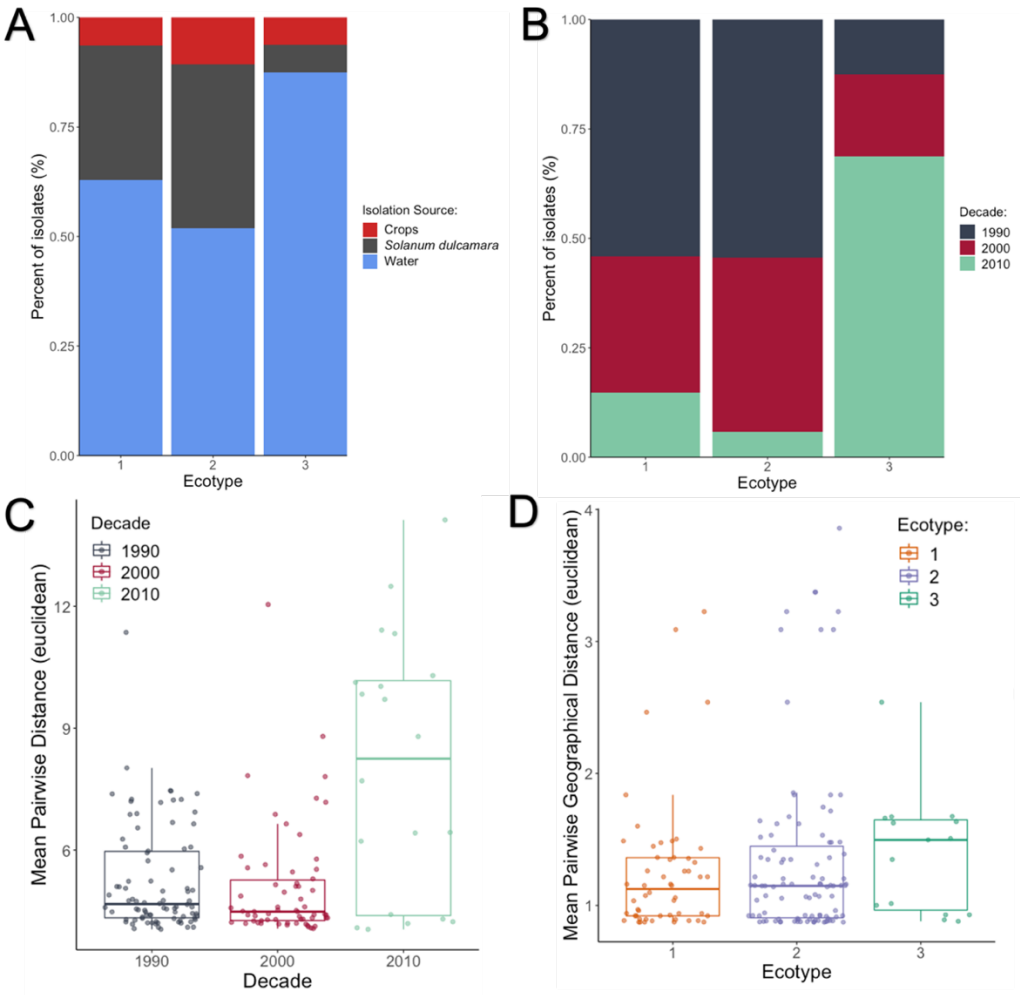


Appendix Figure C.2: The three ecotypes differ in their trait correlation patterns. Pearson's paired correlations of the 46 traits. This is done for each ecotype, assigned by k-means clustering, with 62 isolates in ecotype 1 (A), 104 isolates in 2 (B) and 16 in ecotype 3 (C).

Appendix Table C.2: Table of the strong significant trait correlations per ecotype. Significance was determined by having a p value less than 0.05, strongest correlations are those with a Pearson's correlation coefficient (r) above 0.4 for positive trait correlations or below -0.4 for negative trait correlations.

Ecotype 1			
Trait 1	**Trait 2**	**r**	**p**
ROS	ColdTolerance	0.6731	0.0000
NutrientLimited	HeatTolerance	0.5505	0.0000
Salinity	HeatTolerance	0.5419	0.0000
Salinity	Protein	0.5083	0.0000
pH	NutrientLimited	0.4819	0.0001
ComplexMedia	AntibioticResistance	0.4429	0.0003
Salinity	pH	0.4409	0.0003
SingleCarbons	ComplexMedia	0.4342	0.0004
ROS	pH	0.4276	0.0005
pH	HeatTolerance	0.4178	0.0007
HeatTolerance	ColdTolerance	0.4051	0.0011
SingleCarbons	NutrientLimited	-0.4220	0.0006
ComplexMedia	ColdTolerance	-0.4950	0.0000
NutrientLimited	ComplexMedia	-0.6838	0.0000
Ecotype 2			
Trait 1	**Trait 2**	**r**	**p**
Salinity	pH	0.7387	0.0000
ROS	ColdTolerance	0.7164	0.0000
SingleCarbons	Biofilm	0.5915	0.0000
Salinity	Protein	0.4598	0.0000
Salinity	AntibioticResistance	0.4548	0.0000
ComplexMedia	AntibioticResistance	0.4541	0.0000
pH	ColdTolerance	0.4476	0.0000
ROS	Potential	0.4355	0.0000
NutrientLimited	ComplexMedia	-0.6939	0.0000
Ecotype 3			
Trait 1	**Trait 2**	**r**	**p**
Salinity	pH	0.8036	0.0002
Salinity	NutrientLimited	0.7056	0.0023
pH	HeatTolerance	0.6760	0.0040
ROS	ColdTolerance	0.6158	0.0111
pH	NutrientLimited	0.5688	0.0215
Salinity	HeatTolerance	0.5659	0.0223
Protein	ComplexMedia	0.5584	0.0246
NutrientLimited	HeatTolerance	0.5505	0.0271
ColdTolerance	AntibioticResistance	-0.5212	0.0384

NutrientLimited	ColdTolerance	-0.5463	0.0285



Appendix Figure C.3: Causes of ecotype separation/diversification within UK *R. solanacearum* population due to time of isolation, not isolation source or location. (A) Isolates collected from different sources are evenly distributed across the 3 clusters. (B) Isolates from the three separate decades are not evenly represented across the three clusters. (C) Mean pairwise distance, as measured using Euclidean distances (dist() function in R stats package) from all other isolates. This shows that isolated from most recent decade, 2010s, are significantly more distant from other isolates and therefore are more diverse phenotypically. (D) Mean pairwise geographical distance (using longitude and latitude rather than phenotypes) from all other isolates are the same across the three ecotypes.

Appendix Table C.3: Phenotype unitig GWAS unitig hits summary table. Input phenotype as trait column. Gene name as significant unitig annotation. Hits is the number of significant unitigs annotated as that gene. Maximum -log 10 (p value) across all hits, adjusted p value was adjusted for multiple testing (N=54).

binary_ciprofloxacin_3	cds-ALF88548.1	1	6.5229	4.79050624	0.0833	0.0833	0.429
binary_ciprofloxacin_3	pcaR_1	2	6.5229	4.79050624	0.4912	0.4912	0.384
binary_ciprofloxacin_3	cds-ALF87858.1	2	7.2457	5.51330624	0.4912	0.4912	0.416
binary_rifampicin_4	tycC_3	31	4.4698	2.73740624	0.1145	0.1145	0.1489
binary_rifampicin_4	tycC_2	8	4.4698	2.73740624	0.116	0.116	0.147
binary_rifampicin_4	cds-ALF90441.1	1	4.3546	2.62220624	0.946	0.054	0.21
binary_tetracycline_5	tycC_3	91	6.7773	5.04490624	0.1122	0.1122	0.2825
binary_tetracycline_5	tycC_2	64	6.7773	5.04490624	0.1113	0.1113	0.2733
binary_tetracycline_5	czcA_3	1	4.1379	2.40550624	0.935	0.065	0.356
binary_tetracycline_5	tycC_1	1	4.9136	3.18120624	0.101	0.101	0.281
binary_tetracycline_5	dhbF_3	1	4.9136	3.18120624	0.101	0.101	0.281
binary_tetracycline_5	cds-ALF90441.1|cds-ALF90442.1|eriC|cds-ALF90444.1|cds-ALF90445.1|cds-ALF90446.1|cds-ALF90447.1|cds-ALF90448.1|cds-ALF90449.1|cds-ALF90450.1|bbsG|yfdE|cysO|mec|moeZ_2|cds-ALF90457.1|cds-ALF90458.1|recD|glpG|cds-ALF90462.1|cds-ALF90463.1|cds-ALF90464.1|fadK|cds-ALF90460.1	1	6.6216	4.88920624	0.929	0.071	0.438
binary_tetracycline_5	cds-ALF90441.1	1	6.0605	4.32810624	0.946	0.054	0.484
binary_tetracycline_5	cds-ALF90702.1|cds-ALF90703.1|cdiA2_3	1	4.1379	2.40550624	0.935	0.065	0.356
biofilmCPG	czcA_3	1	4.4179	2.68550624	0.935	0.065	0.875
biofilmCPG	cds-ALF90441.1|cds-ALF90442.1|eriC|cds-ALF90444.1|cds-ALF90445.1|cds-ALF90446.1|cds-ALF90447.1|cds-ALF90448.1|cds-ALF90449.1|cds-ALF90450.1|bbsG|yfdE|cysO|mec|moeZ_2|cds-ALF90457.1|cds-ALF90458.1|recD|glpG|cds-ALF90462.1|cds-ALF90463.1|cds-ALF90464.1|fadK|cds-ALF90460.1	1	5.6459	3.91350624	0.929	0.071	0.957
biofilmCPG	cds-ALF90441.1	1	5.0846	3.35220624	0.946	0.054	1.04

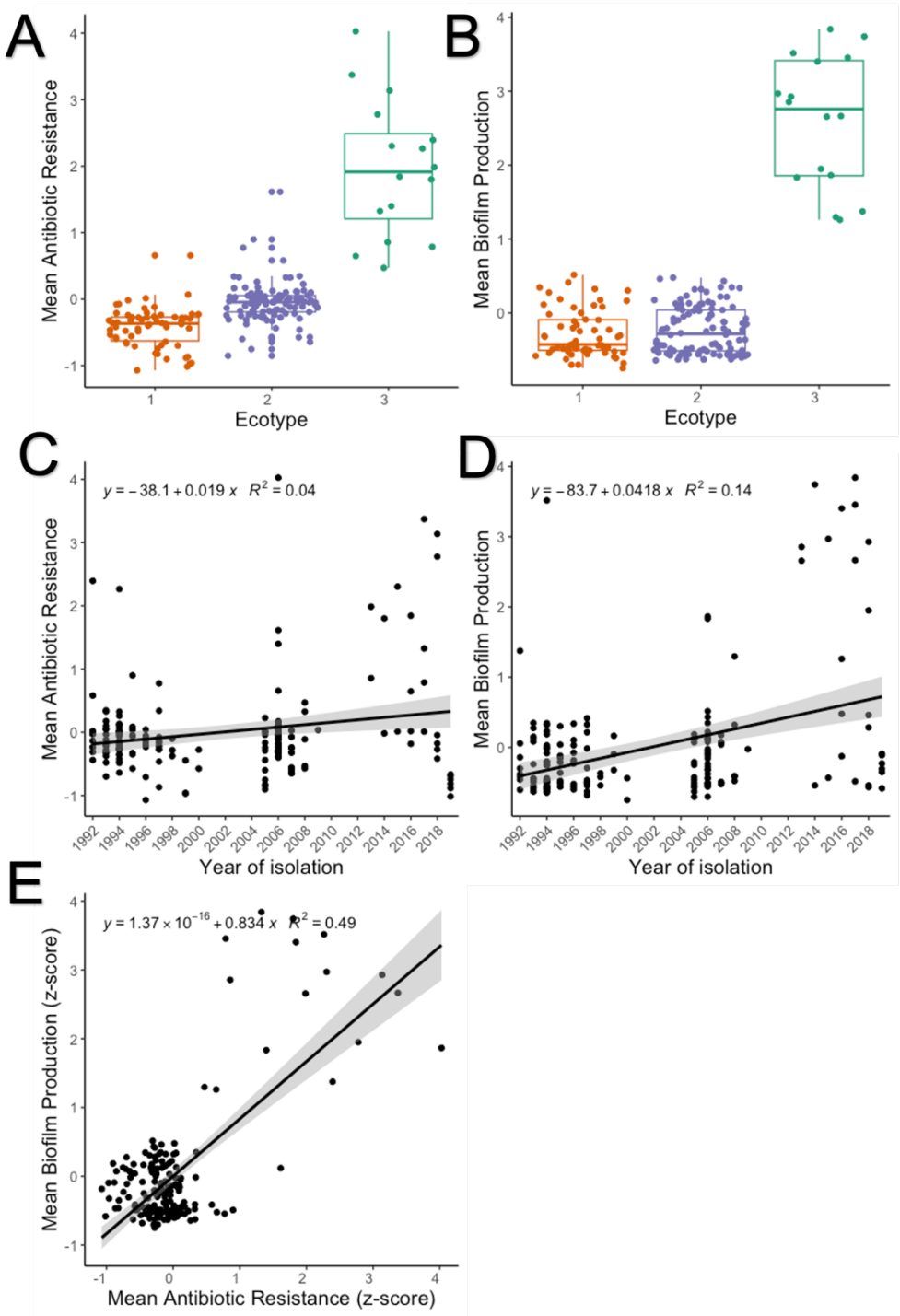
biofilmCPG	cds-ALF90702.1|cds-ALF90703.1|cdiA2_3	1	4.4179	2.68550624	0.935	0.065	0.875
biofilmCS	tycC_2	29	4.9431	3.21070624	0.1019	0.1019	2.4503
biofilmCS	tycC_3	20	4.9431	3.21070624	0.1059	0.1059	2.434
biofilmCS	cds-ALF90441.1|cds-ALF90442.1|eriC|cds-ALF90444.1|cds-ALF90445.1|cds-ALF90446.1|cds-ALF90447.1|cds-ALF90448.1|cds-ALF90449.1|cds-ALF90450.1|bbsG|yfdE|cysO|mec|moeZ_2|cds-ALF90457.1|cds-ALF90458.1|recD|glpG|cds-ALF90462.1|cds-ALF90463.1|cds-ALF90464.1|fadK|cds-ALF90460.1	1	7.0357	5.30330624	0.929	0.071	5.89
biofilmCS	cds-ALF90441.1	1	6.1871	4.45470624	0.946	0.054	5.55
biofilmNB	czcA_3	1	4.2434	2.51100624	0.935	0.065	0.772
biofilmNB	tycC_2	4	4.5243	2.79190624	0.0893	0.0893	0.661
biofilmNB	tycC_3	4	4.5243	2.79190624	0.0893	0.0893	0.661
biofilmNB	cds-ALF90441.1|cds-ALF90442.1|eriC|cds-ALF90444.1|cds-ALF90445.1|cds-ALF90446.1|cds-ALF90447.1|cds-ALF90448.1|cds-ALF90449.1|cds-ALF90450.1|bbsG|yfdE|cysO|mec|moeZ_2|cds-ALF90457.1|cds-ALF90458.1|recD|glpG|cds-ALF90462.1|cds-ALF90463.1|cds-ALF90464.1|fadK|cds-ALF90460.1	1	7.1561	5.42370624	0.929	0.071	0.971
biofilmNB	cds-ALF90702.1|cds-ALF90703.1|cdiA2_3	1	4.2434	2.51100624	0.935	0.065	0.772
biofilmSP	cds-ALF90441.1	1	4.1662	2.43380624	0.946	0.054	1.57
ciprofloxacin_3	czcA_3	1	3.9431	2.21070624	0.935	0.065	0.22
ciprofloxacin_3	cds-ALF90441.1|cds-ALF90442.1|eriC|cds-ALF90444.1|cds-ALF90445.1|cds-ALF90446.1|cds-ALF90447.1|cds-ALF90448.1|cds-ALF90449.1|cds-ALF90450.1|bbsG|yfdE|cysO|mec|moeZ_2|cds-ALF90457.1|cds-ALF90458.1|recD|glpG|cds-ALF90462.1|cds-ALF90463.1|cds-ALF90464.1|fadK|cds-ALF90460.1	1	5.7447	4.01230624	0.929	0.071	0.258
ciprofloxacin_3	cds-ALF90441.1	1	5.3675	3.63510624	0.946	0.054	0.285

ciprofloxacin_3	cds-ALF90702.1|cds-ALF90703.1|cdiA2_3	1	3.9431	2.21070624	0.935	0.065	0.22
ciprofloxacin_5	czcA_3	1	3.9431	2.21070624	0.935	0.065	0.236
ciprofloxacin_5	cds-ALF90441.1|cds-ALF90442.1|eriC|cds-ALF90444.1|cds-ALF90445.1|cds-ALF90446.1|cds-ALF90447.1|cds-ALF90448.1|cds-ALF90449.1|cds-ALF90450.1|bbsG|yfdE|cysO|mec|moeZ_2|cds-ALF90457.1|cds-ALF90458.1|recD|glpG|cds-ALF90462.1|cds-ALF90463.1|cds-ALF90464.1|fadK|cds-ALF90460.1	1	5.3936	3.66120624	0.929	0.071	0.267
ciprofloxacin_5	cds-ALF90441.1	1	5.6198	3.88740624	0.946	0.054	0.315
ciprofloxacin_5	cds-ALF90702.1|cds-ALF90703.1|cdiA2_3	1	3.9431	2.21070624	0.935	0.065	0.236
CS	tycC_3	67	5.3747	3.64230624	0.1164	0.1164	18.4045
CS	tycC_2	50	5.3747	3.64230624	0.1127	0.1127	19.158
CS	czcA_3	1	5.821	4.08860624	0.935	0.065	26.6
CS	cds-ALF90441.1|cds-ALF90442.1|eriC|cds-ALF90444.1|cds-ALF90445.1|cds-ALF90446.1|cds-ALF90447.1|cds-ALF90448.1|cds-ALF90449.1|cds-ALF90450.1|bbsG|yfdE|cysO|mec|moeZ_2|cds-ALF90457.1|cds-ALF90458.1|recD|glpG|cds-ALF90462.1|cds-ALF90463.1|cds-ALF90464.1|fadK|cds-ALF90460.1	1	7.2366	5.50420624	0.929	0.071	28.6
CS	cds-ALF90441.1	1	8.0778	6.34540624	0.946	0.054	34.5
CS	cds-ALF90702.1|cds-ALF90703.1|cdiA2_3	1	5.821	4.08860624	0.935	0.065	26.6
siderophore	tycC_3	75	4.5952	2.86280624	0.1139	0.1139	22.5693
siderophore	tycC_2	55	4.5952	2.86280624	0.1103	0.1103	22.5582
siderophore	czcA_3	1	6.1152	4.38280624	0.935	0.065	35.3

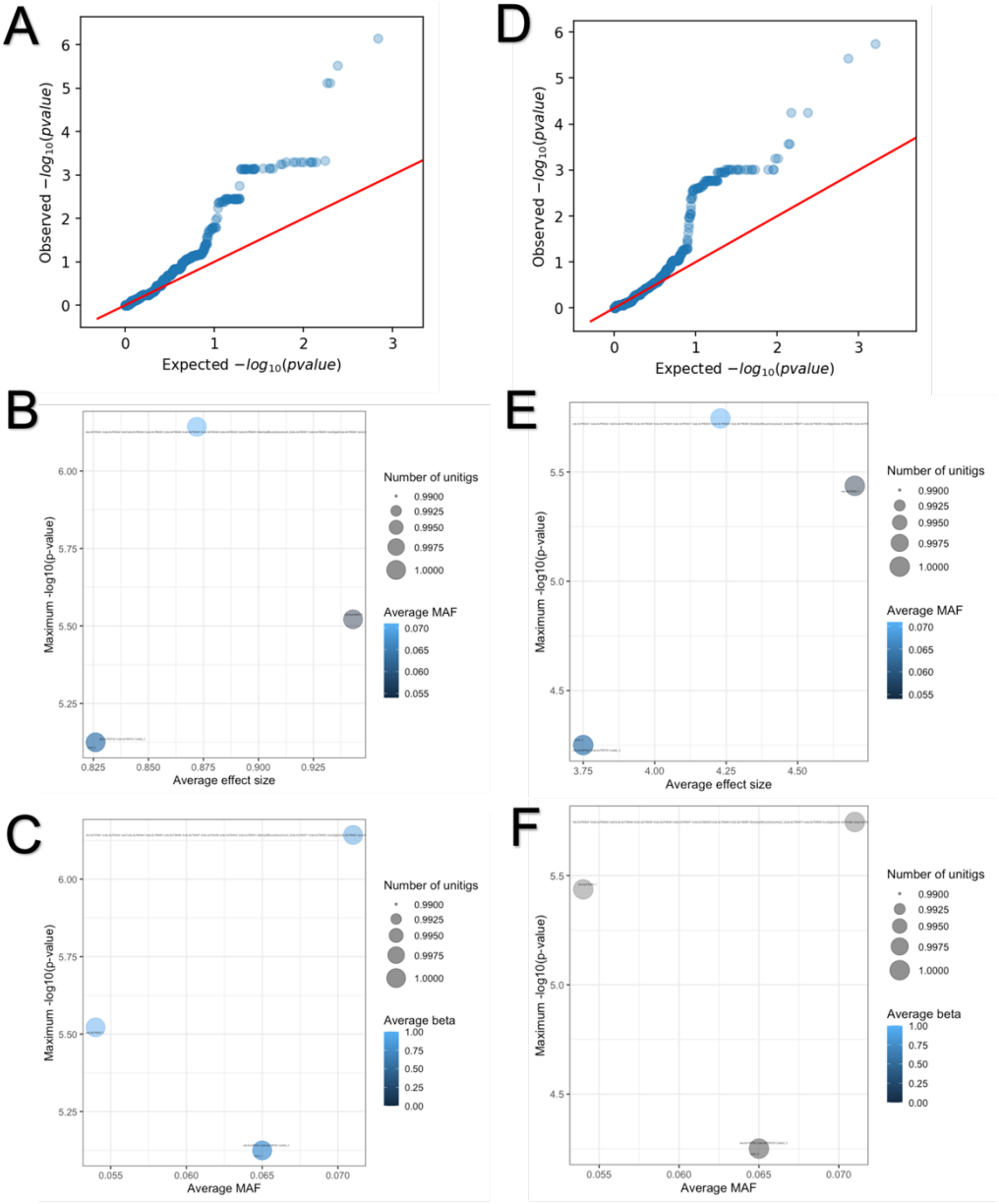
siderophore	cds-ALF90441.1|cds-ALF90442.1|eriC|cds-ALF90444.1|cds-ALF90445.1|cds-ALF90446.1|cds-ALF90447.1|cds-ALF90448.1|cds-ALF90449.1|cds-ALF90450.1|bbsG|yfdE|cysO|mec|moeZ_2|cds-ALF90457.1|cds-ALF90458.1|recD|glpG|cds-ALF90462.1|cds-ALF90463.1|cds-ALF90464.1|fadK|cds-ALF90460.1	1	7.2782	5.54580624	0.929	0.071	36.9
siderophore	cds-ALF90441.1	1	6.0128	4.28040624	0.946	0.054	38.4
siderophore	cds-ALF90702.1|cds-ALF90703.1|cdiA2_3	1	6.1152	4.38280624	0.935	0.065	35.3
gentamycin_0.5	tycC_3	28	4.4989	2.76650624	0.1107	0.1107	0.1294
gentamycin_0.5	tycC_2	9	4.4989	2.76650624	0.1031	0.1031	0.135
gentamycin_0.5	czcA_3	1	6.0287	4.29630624	0.935	0.065	0.222
gentamycin_0.5	cds-ALF90441.1|cds-ALF90442.1|eriC|cds-ALF90444.1|cds-ALF90445.1|cds-ALF90446.1|cds-ALF90447.1|cds-ALF90448.1|cds-ALF90449.1|cds-ALF90450.1|bbsG|yfdE|cysO|mec|moeZ_2|cds-ALF90457.1|cds-ALF90458.1|recD|glpG|cds-ALF90462.1|cds-ALF90463.1|cds-ALF90464.1|fadK|cds-ALF90460.1	1	7.3391	5.60670624	0.929	0.071	0.234
gentamycin_0.5	cds-ALF90441.1	1	6.3279	4.59550624	0.946	0.054	0.252
gentamycin_0.5	cds-ALF90702.1|cds-ALF90703.1|cdiA2_3	1	6.0287	4.29630624	0.935	0.065	0.222
gentamycin_1	tycC_3	57	6.1209	4.38850624	0.12	0.12	0.1579
gentamycin_1	tycC_2	30	6.1209	4.38850624	0.1203	0.1203	0.1518
gentamycin_1	czcA_3	1	5.9031	4.17070624	0.935	0.065	0.213
gentamycin_1	cds-ALF90441.1|cds-ALF90442.1|eriC|cds-ALF90444.1|cds-ALF90445.1|cds-ALF90446.1|cds-ALF90447.1|cds-ALF90448.1|cds-ALF90449.1|cds-ALF90450.1|bbsG|yfdE|cysO|mec|moeZ_2|cds-ALF90457.1|cds-ALF90458.1|recD|glpG|cds-ALF90462.1|cds-ALF90463.1|cds-ALF90464.1|fadK|cds-ALF90460.1	1	8.2823	6.54990624	0.929	0.071	0.243
gentamycin_1	cds-ALF90441.1	1	6.9872	5.25480624	0.946	0.054	0.256

gentamycin_1	cds-ALF90702.1|cds-ALF90703.1|cdiA2_3	1	5.9031	4.17070624	0.935	0.065	0.213
nicotinamide	tycC_3	55	6.2418	4.50940624	0.1041	0.1041	0.1007
nicotinamide	tycC_2	19	6.2418	4.50940624	0.0987	0.0987	0.1069
nicotinamide	tycC_1	1	4.0255	2.29310624	0.101	0.101	0.0932
nicotinamide	dhbF_3	1	4.0255	2.29310624	0.101	0.101	0.0932
rifampicin_4	tycC_3	50	6.5702	4.83780624	0.1251	0.1251	0.0518
rifampicin_4	tycC_2	22	6.5702	4.83780624	0.1332	0.1332	0.0483
rifampicin_4	cds-ALF90441.1	1	4.8633	3.13090624	0.946	0.054	0.0688
salinity_0.5	czcA_3	1	4.4134	2.68100624	0.935	0.065	0.204
salinity_0.5	cds-ALF90441.1|cds-ALF90442.1|eriC|cds-ALF90444.1|cds-ALF90445.1|cds-ALF90446.1|cds-ALF90447.1|cds-ALF90448.1|cds-ALF90449.1|cds-ALF90450.1|bbsG|yfdE|cysO|mec|moeZ_2|cds-ALF90457.1|cds-ALF90458.1|recD|glpG|cds-ALF90462.1|cds-ALF90463.1|cds-ALF90464.1|fadK|cds-ALF90460.1	1	4.8182	3.08580624	0.929	0.071	0.205
salinity_0.5	cds-ALF90441.1	1	5.2277	3.49530624	0.946	0.054	0.245
salinity_0.5	cds-ALF90702.1|cds-ALF90703.1|cdiA2_3	1	4.4134	2.68100624	0.935	0.065	0.204
SP	cds-ALF90441.1|cds-ALF90442.1|eriC|cds-ALF90444.1|cds-ALF90445.1|cds-ALF90446.1|cds-ALF90447.1|cds-ALF90448.1|cds-ALF90449.1|cds-ALF90450.1|bbsG|yfdE|cysO|mec|moeZ_2|cds-ALF90457.1|cds-ALF90458.1|recD|glpG|cds-ALF90462.1|cds-ALF90463.1|cds-ALF90464.1|fadK|cds-ALF90460.1	1	4.3778	2.64540624	0.929	0.071	9.61
SP	cds-ALF90441.1	1	4.6946	2.96220624	0.946	0.054	11.4
tetracycline_1	cds-ALF90441.1	1	4.699	2.96660624	0.946	0.054	0.379
tetracycline_5	tycC_3	95	9.0173	7.28490624	0.1125	0.1125	0.2064
tetracycline_5	tycC_2	67	9.0173	7.28490624	0.1119	0.1119	0.1952
tetracycline_5	czcA_3	1	5.6478	3.91540624	0.935	0.065	0.235
tetracycline_5	dhbF_3	1	5.3904	3.65800624	0.101	0.101	0.188
tetracycline_5	tycC_1	1	5.3904	3.65800624	0.101	0.101	0.188

tetracycline_5	cds-ALF90441.1|cds-ALF90442.1|eriC|cds-ALF90444.1|cds-ALF90445.1|cds-ALF90446.1|cds-ALF90447.1|cds-ALF90448.1|cds-ALF90449.1|cds-ALF90450.1|bbsG|yfdE|cysO|mec|moeZ_2|cds-ALF90457.1|cds-ALF90458.1|recD|glpG|cds-ALF90462.1|cds-ALF90463.1|cds-ALF90464.1|fadK|cds-ALF90460.1	1	6.4067	4.67430624	0.929	0.071	0.241
tetracycline_5	cds-ALF90441.1	1	7.2941	5.56170624	0.946	0.054	0.295
tetracycline_5	cds-ALF90702.1|cds-ALF90703.1|cdiA2_3	1	5.6478	3.91540624	0.935	0.065	0.235
xylose	czcA_3	1	3.9706	2.23820624	0.935	0.065	0.386
xylose	cds-ALF90702.1|cds-ALF90703.1|cdiA2_3	1	3.9706	2.23820624	0.935	0.065	0.386



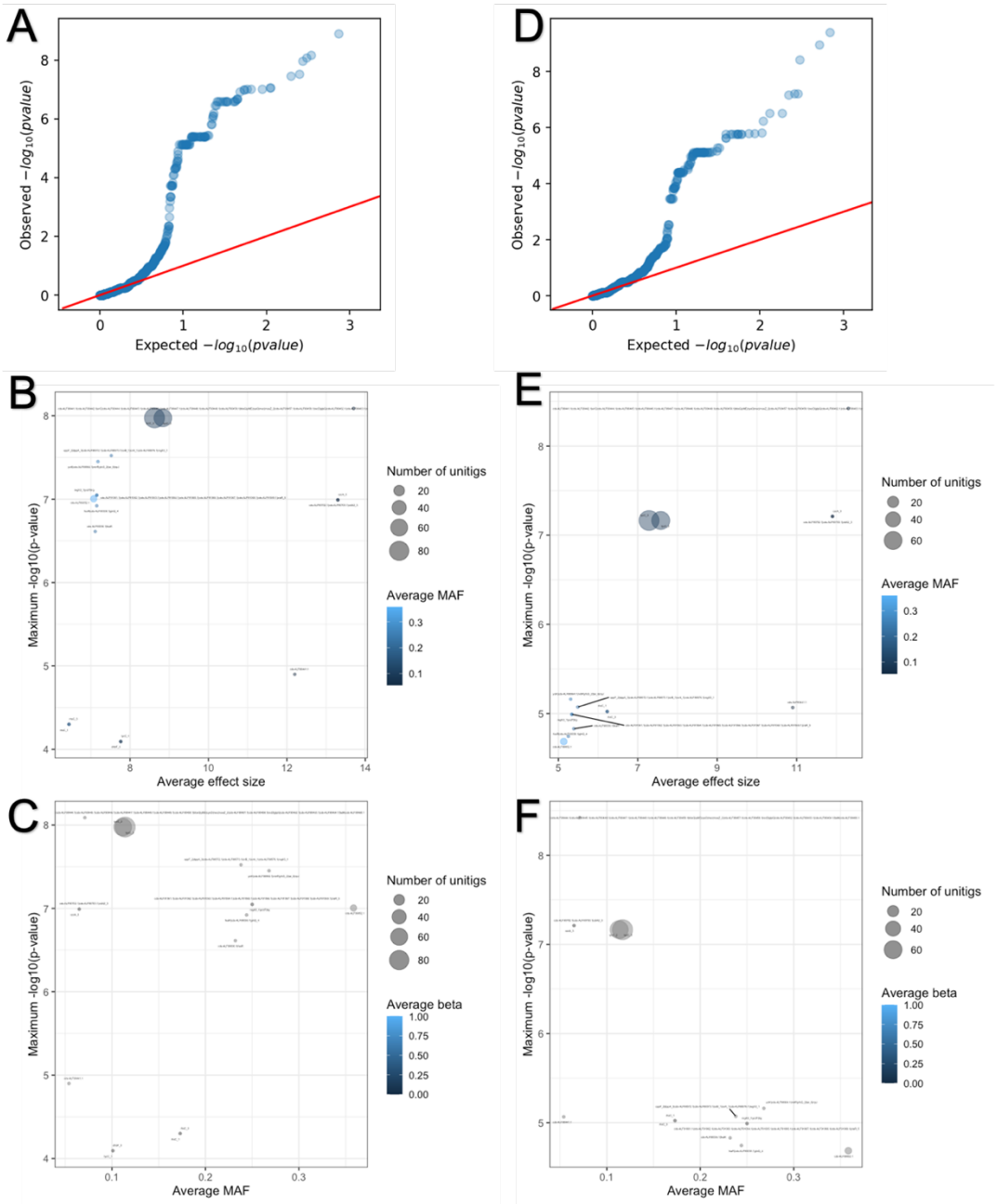
Appendix Figure C.4: Antibiotic resistance and biofilm production traits. Taking mean standardised value (z-score) across all antibiotic resistance traits (n=8) and biofilm production traits (n=4) per isolate. Shows that ecotype 3 has higher antibiotic resistance and biofilm production compared to the other two ecotypes (A-B). Antibiotic resistance and biofilm production are increasing over time (C-D). Biofilm production and antibiotic resistance are correlated with one another (E).



Appendix Figure C.5:Ecotype GWAS unitig hits. Unitig hits when the input continuous phenotype was (A, B and C) ecotype or (D, E and F) PC2 values from figure 1.2C. (A) and (D) show QQ plots of all unitigs for ecotype and PC2 values respectively. (B, C) and (E, F) show significant unitig hits for ecotype and PC2 values respectively.

Appendix Table C.4: Ecotype GWAS significant hits summary.

SNPs					
Gene Name	**Filter p value**	**lrt p value**	**Allele frequency**	**Beta**	**Variant h2**
Ecotype					
CP012687_293375_A_G	0.0318	0.0073	0.1370	0.3630	0.2060
PC2					
CP012687_293375_A_G	0.0571	0.0060	0.1370	1.8600	0.2110
UNITIGS					
Gene Name	**Hits**	**Maximum -log10(p value)**	**Avgerage allele frequency**	**Average minor allele frequency**	**Average beta**
Ecotype					
czcA_3	1	5.1244	0.9350	0.0650	0.8260
cds-ALF90441.1|cds-ALF90442.1|eriC|cds-ALF90444.1|cds-ALF90445.1|cds-ALF90446.1|cds-ALF90447.1|cds-ALF90448.1|cds-ALF90449.1|cds-ALF90450.1|bbsG|yfdE|cysO|mec|moeZ_2|cds-ALF90457.1|cds-ALF90458.1|recD|glpG|cds-ALF90462.1|cds-ALF90463.1|cds-ALF90464.1|fadK|cds-ALF90460.1	1	6.1427	0.9290	0.0710	0.8720
cds-ALF90441.1	1	5.5214	0.9460	0.0540	0.9430
cds-ALF90702.1|cds-ALF90703.1|cdiA2_3	1	5.1244	0.9350	0.0650	0.8260
PC2					
czcA_3	1	4.2503	0.9350	0.0650	3.7500
cds-ALF90441.1|cds-ALF90442.1|eriC|cds-ALF90444.1|cds-ALF90445.1|cds-ALF90446.1|cds-ALF90447.1|cds-ALF90448.1|cds-ALF90449.1|cds-ALF90450.1|bbsG|yfdE|cysO|mec|moeZ_2|cds-ALF90457.1|cds-ALF90458.1|recD|glpG|cds-ALF90462.1|cds-ALF90463.1|cds-ALF90464.1|fadK|cds-ALF90460.1	1	5.7447	0.9290	0.0710	4.2300
cds-ALF90441.1	1	5.4365	0.9460	0.0540	4.7000
cds-ALF90702.1|cds-ALF90703.1|cdiA2_3	1	4.2503	0.9350	0.0650	3.7500



Appendix Figure C.6: Temporal GWAS unitig hits. Unitig hits when the input continuous phenotype was (A, B and C) year or (D, E and F) decade. (A) and (D) show QQ plots of all unitigs for year and decade respectively. (B, C) and (E, F) show significant unitig hits for year and decade respectively.

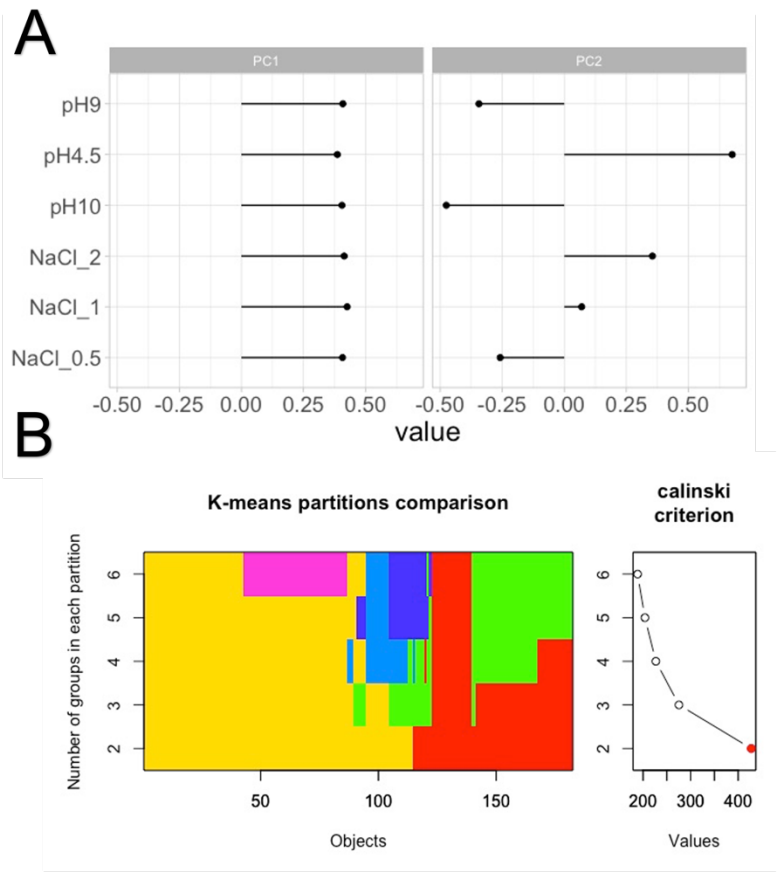
Appendix Table C.5: Temporal GWAS significant associations summary table.

SNPs					
Gene Name	**Filter p value**	**lrt p value**	**Allele frequency**	**Beta**	**Variant h2**
Year					
CP012687_293375_A_G	0.0003	0.0000	0.1370	9.2800	0.4130
Decade					
CP012687_293375_A_G	0.0002	0.0000	0.1370	8.3000	0.4130
COGs					
Gene Name	**Filter p value**	**lrt p value**	**Allele frequency**	**Beta**	**Variant h2**
Year					
group_2086	0.0079	0.0000	0.8870	-7.9800	0.3120
group_1877	0.0000	0.0000	0.2320	-6.4100	0.3510
group_1933	0.0000	0.0000	0.2140	-7.3300	0.3910
group_1905	0.0000	0.0000	0.1960	-7.0300	0.3630
Decade					
group_2086	0.0059	0.0000	0.8870	-7.1100	0.3510
group_1877	0.0000	0.0001	0.2320	-4.9000	0.3000
group_1933	0.0000	0.0000	0.2140	-5.5900	0.3330
group_1905	0.0000	0.0001	0.1960	-5.3200	0.1060
UNITIGS					
Gene Name	**Hits**	**Maximum -log10(p value)**	**Avgerage allele frequency**	**Average minor allele frequency**	**Average beta**
Year					
tycC_3	87	7.9747	0.1140	0.1140	8.6238
tycC_2	66	7.9747	0.1115	0.1115	8.8370
czcA_3	1	6.9914	0.9350	0.0650	13.3000
cds-ALF90036.1|feaR	1	6.6126	0.7680	0.2320	7.1100
feaR|cds-ALF90038.1|glnQ_4	1	6.9208	0.7560	0.2440	7.1500
tycC_1	1	4.0931	0.1010	0.1010	7.7600
dhbF_3	1	4.0931	0.1010	0.1010	7.7600
yciK|cds-ALF88964.1|mtrR|yfcG_2|tar_6|ripJ	1	7.4510	0.7320	0.2680	7.1800

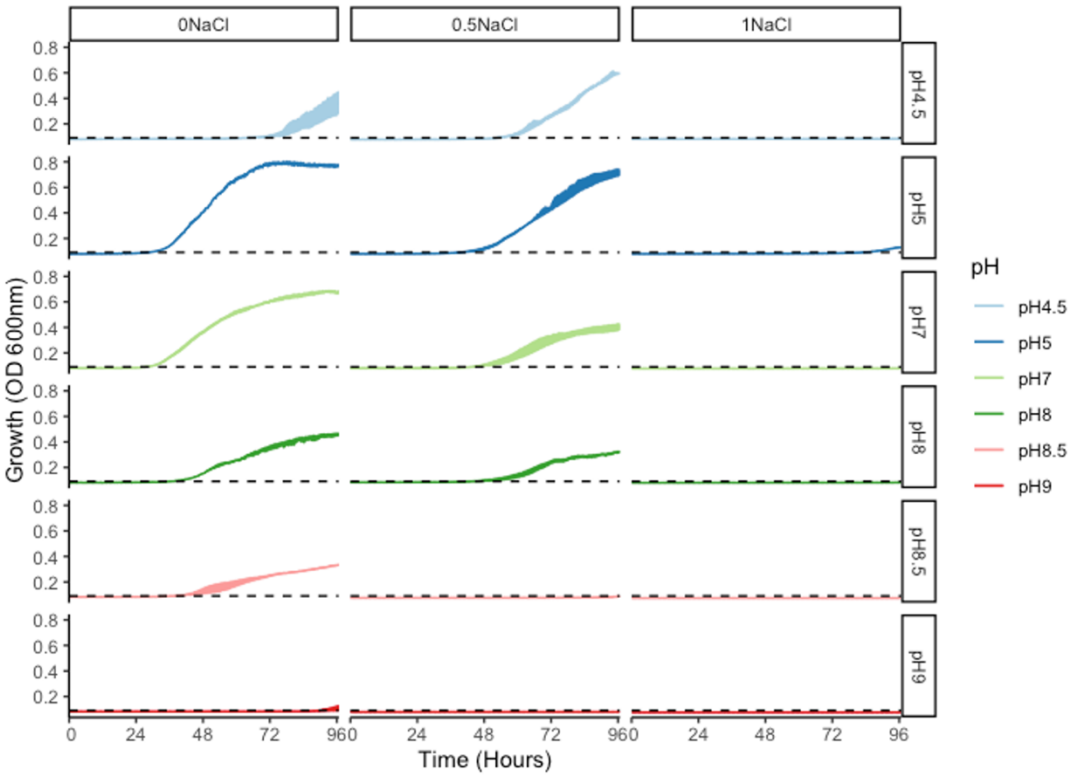
cds-ALF90441.1|cds-ALF90442.1|eriC|cds-ALF90444.1|cds-ALF90445.1|cds-ALF90446.1|cds-ALF90447.1|cds-ALF90448.1|cds-ALF90449.1|cds-ALF90450.1|bbsG|yfdE|cysO|mec|moeZ_2|cds-ALF90457.1|cds-ALF90458.1|recD|glpG|cds-ALF90462.1|cds-ALF90463.1|cds-ALF90464.1|fadK|cds-ALF90460.1	1	8.0883	0.9290	0.0710	13.7000
oppF_2|dppA_3|cds-ALF88572.1|cds-ALF88573.1|iorB_1|iorA_1|cds-ALF88576.1|regX3_1	1	7.5229	0.7620	0.2380	7.5200
cds-ALF89052.1	5	7.0048	0.6416	0.3584	7.0720
cds-ALF91061.1|cds-ALF91062.1|cds-ALF91063.1|cds-ALF91064.1|cds-ALF91065.1|cds-ALF91066.1|cds-ALF91067.1|cds-ALF91068.1|cds-ALF91069.1|zraR_5	1	7.0477	0.7500	0.2500	7.1500
rhsC_1	1	4.3002	0.8270	0.1730	6.4400
rhsC_3	1	4.3002	0.8270	0.1730	6.4400
cds-ALF90441.1	1	4.8996	0.9460	0.0540	12.2000
cds-ALF90702.1|cds-ALF90703.1|cdiA2_3	1	6.9914	0.9350	0.0650	13.3000
regX3_1|ycdT|trg	1	7.0477	0.7500	0.2500	7.1500
Decade					
tycC_3	77	7.1649	0.1169	0.1169	7.2822
tycC_2	61	7.1649	0.1133	0.1133	7.5785
czcA_3	1	7.2111	0.9350	0.0650	11.9000
cds-ALF90036.1|feaR	1	4.8297	0.7680	0.2320	5.3900
feaR|cds-ALF90038.1|glnQ_4	1	4.7447	0.7560	0.2440	5.2500
yciK|cds-ALF88964.1|mtrR|yfcG_2|tar_6|ripJ	1	5.1618	0.7320	0.2680	5.3100

cds-ALF90441.1|cds-ALF90442.1|eriC|cds-ALF90444.1|cds-ALF90445.1|cds-ALF90446.1|cds-ALF90447.1|cds-ALF90448.1|cds-ALF90449.1|cds-ALF90450.1|bbsG|yfdE|cysO|mec|moeZ_2|cds-ALF90457.1|cds-ALF90458.1|recD|glpG|cds-ALF90462.1|cds-ALF90463.1|cds-ALF90464.1|fadK|cds-ALF90460.1	1	8.4214	0.9290	0.0710	12.3000
oppF_2|dppA_3|cds-ALF88572.1|cds-ALF88573.1|iorB_1|iorA_1|cds-ALF88576.1|regX3_1	1	5.0726	0.7620	0.2380	5.4900
cds-ALF89052.1	5	4.6861	0.6416	0.3584	5.1360
cds-ALF91061.1|cds-ALF91062.1|cds-ALF91063.1|cds-ALF91064.1|cds-ALF91065.1|cds-ALF91066.1|cds-ALF91067.1|cds-ALF91068.1|cds-ALF91069.1|zraR_5	1	4.9914	0.7500	0.2500	5.3400
rhsC_3	1	5.0232	0.8270	0.1730	6.2300
rhsC_1	1	5.0232	0.8270	0.1730	6.2300
cds-ALF90441.1	1	5.0670	0.9460	0.0540	10.9000
cds-ALF90702.1|cds-ALF90703.1|cdiA2_3	1	7.2111	0.9350	0.0650	11.9000
regX3_1|ycdT|trg	1	4.9914	0.7500	0.2500	5.3400

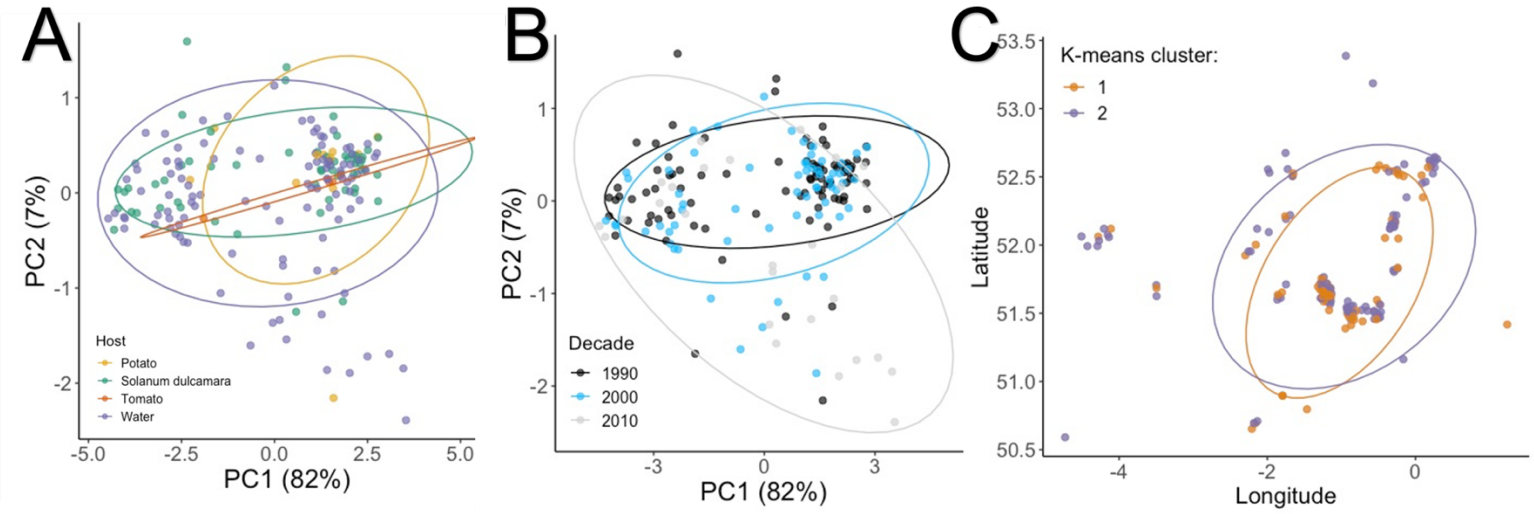
Appendix D: Chapter 4



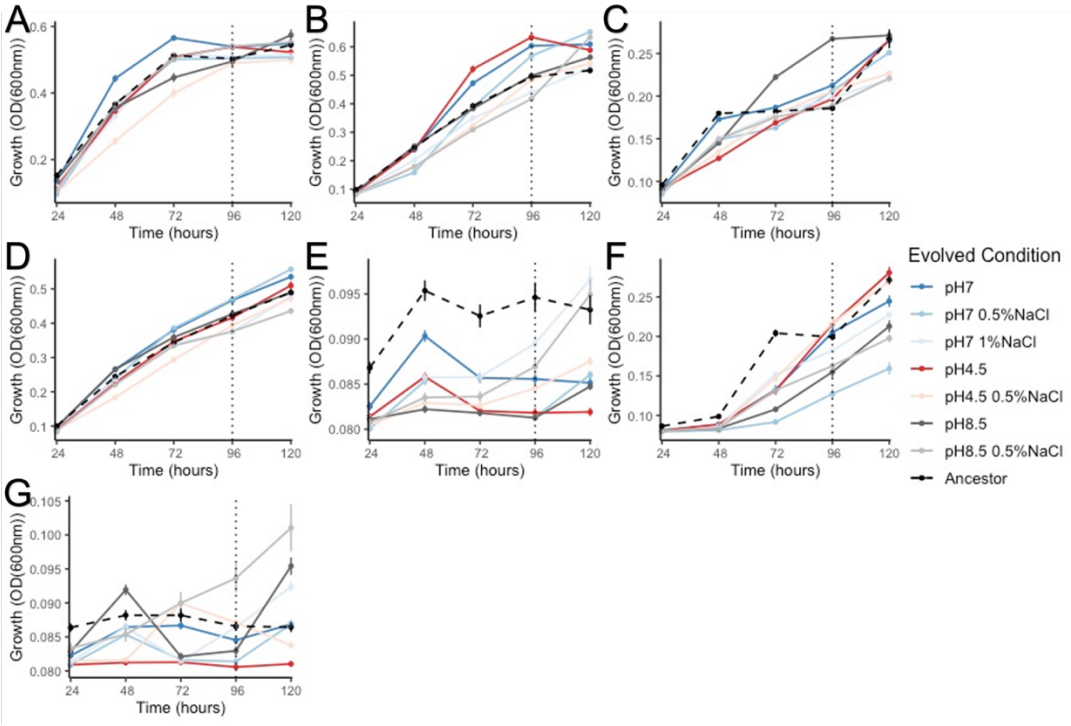
Appendix Figure D.1: (A) loadings of PCA plot in figure 1.2B. (B) Calinski method determined optimal clusters as 2.



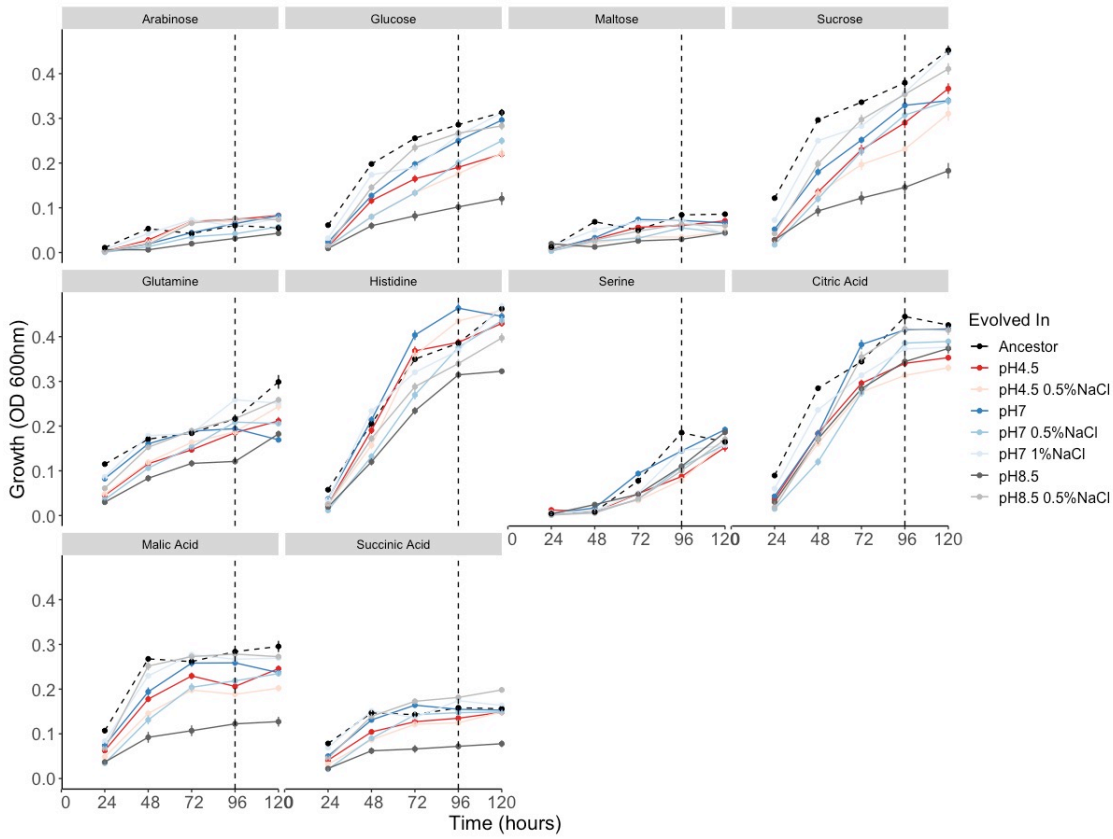
Appendix Figure D.2: Preliminary results indicate that growth above pH 8.5 is too lethal and growth still occurs at pH 4.5 justifying the pH range used for the selection experiment in figure 4.4. A single YO336 colony was grown in 6 different pH conditions (pH 4.5, 5, 7, 8, 8.5 and 9) and three different salinity concentrations (0, 0.5 and 1%). Three replicates were conducted per condition. Optical density (OD) at 600nm was taken every 20 minutes as a measure of growth.



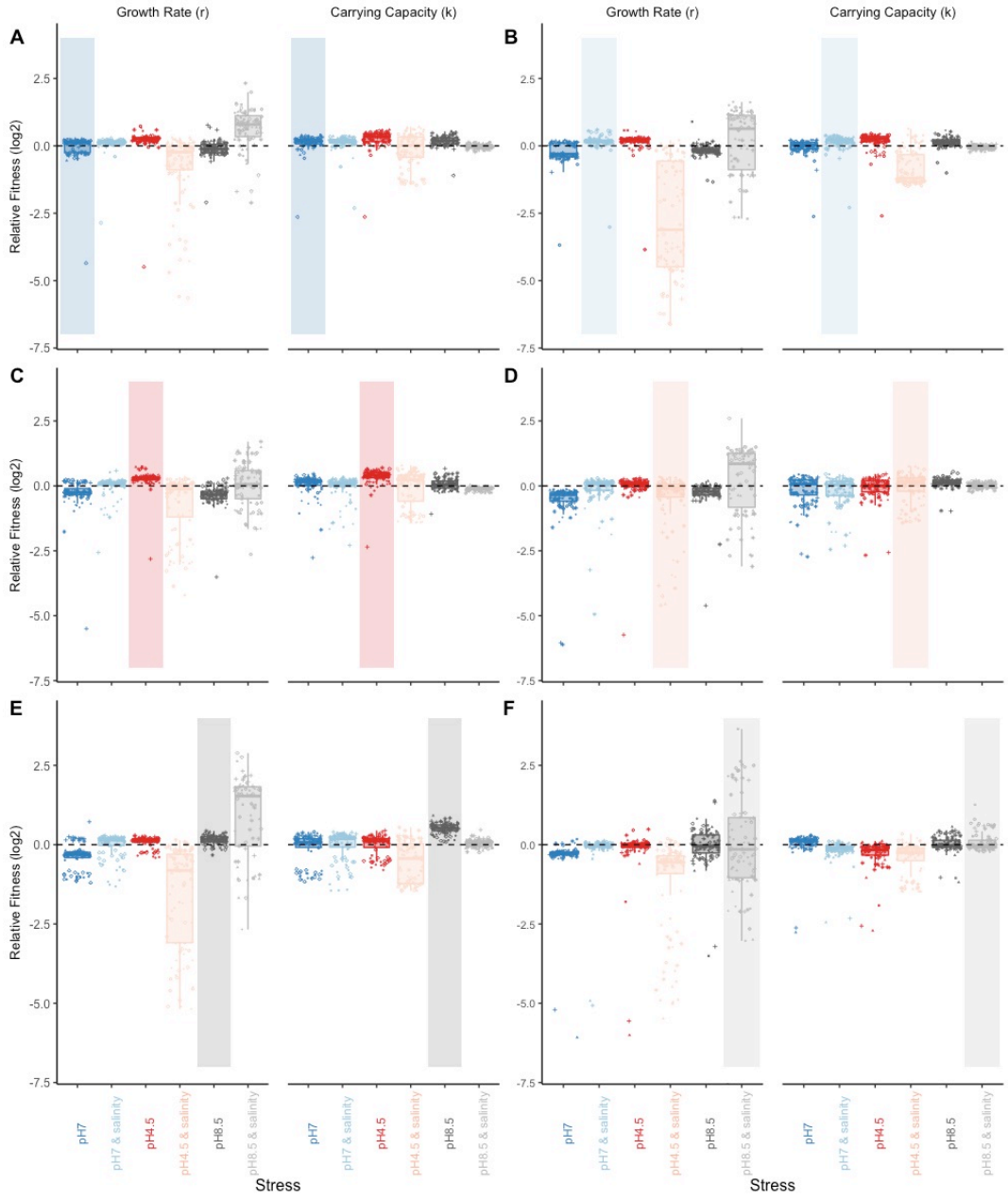
Appendix Figure D.3: Host, decade and location all cannot explain abiotic stress tolerance diversity among UK isolates. PCA plot seen in figure 1B with each point representing one UK isolate, coloured by (A) host (N=182) and (B) decade (N=180) isolated from. (C) Each point represents one UK isolate (N=174) plotted by coordinates of location isolated from and coloured by assigned cluster through k-means clustering (k=2). Eclipses shows the 95% confidence intervals around the centroids for each group.



Appendix Figure D.4: Growth curves showing *R. solanacearum* adaptation to different stressors. Growth curves used to calculate area under the curve, growth rate, and carrying capacity from. Graphs show the mean average growth of evolved clones (as measured by optical density reads at 600nm) every 24 hours over the course of 5 days. Conditions in which they are growing in differs between panels; pH 7 (A), pH 4.5 (B), pH 8.5 (C), pH 7 0.5% NaCl (D), pH 7 1% NaCl (E), pH 4.5 0.5% NaCl (F) and pH 8.5 0.5% NaCl (G). Colour indicates the condition clones were evolved under; blue shows growth of clones evolved in pH 7 in a variety of salinity concentrations (lighter the colour the higher the salinity concentration). Red shows *R. solanacearum* evolved in acidic conditions with and without salinity, and grey shows the same for alkaline evolved clones. Each point shows the mean average OD read across all clones (N = 96 for all but ancestor, where N = 16) and the bars show one standard error from the mean. The dashed vertical line indicates the cut-off point used to calculate area under the curve, growth rate and carrying capacity (96 hours).



Appendix Figure D.5: Growth curves showing evolved *R. solanacearum* clone's growth in 10 carbon resources. Growth curves used to calculate area under the curve data. Graphs show the mean average growth of evolved clones (as measured by optical density reads at 600nm) every 24 hours over the course of 5 days. Carbons in which they are growing in is labelled above each graph and were at a concentration of 10mM. Colour indicates the condition clones were evolved under; blue shows growth of clones evolved in pH 7 in a variety of salinity concentrations (lighter the colour the higher the salinity concentration). Red shows *R. solanacearum* evolved in acidic conditions with and without salinity, and grey shows the same for alkaline evolved clones. Each point shows the mean average OD read across all clones (N = 96 for all but ancestor, where N = 16) and the bars show one standard error from the mean. The dashed vertical line indicates the cut-off point used to calculate area under the curve, growth rate and carrying capacity (96 hours).



Appendix Figure D.6: Relative growth from growth rate and carrying capacity data showing *R. solanacearum* adaptation to different stressors. Relative growth (log2) in growth rate (left facet) and carrying capacity (right facet) was calculated by dividing clone's growth rate or carrying capacity by the mean ancestor's value (N = 16) and transforming by log2. Dashed line shows the ancestor value (log2(1) = 0), anything above the line has increased growth compared to the ancestor and anything below has reduced growth. Boxplot lines show the median relative fitness (log2) of all evolved clones in each stressor with the box indicating the interquartile range and whiskers showing the 95% quantile (N = 96). Blue boxplots show relative growth of isolates in pH 7 in the presence (light blue) and absence (darker blue) of salinity. Red boxplots show R. solanacearum relative growth in acidic conditions with and without salinity, and grey show the same for alkaline evolved clones. Each point shows an individual clone with shape referring to which population (N=8) that clone evolved in (12 clones per population). Each panel refers to a different selection condition in the selection experiment, (A and B) show clones evolved in neutral pH conditions (pH 7), in the absence of salinity (A) and presence of 0.5% NaCl (B). (C and D) are isolates evolved to acidic pH conditions in the absence (C) and presence (D) of 0.5% salinity. Finally, (E and F) show clones exposed to alkaline pH conditions throughout the selection experiment, in the absence (E) and presence (F) of 0.5% salinity.

Appendix Table D.1: Table of statistics for area under the curve (AUC) data. One-way ANOVAs conducted per condition where growth (AUC) between populations from different evolutionary histories is compared, along with the ancestor.

pH7						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	4.2	0.002			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	5.7801	2.1025	57	2.7492	0.1047
4.5	7	-4.1470	2.1025	57	-1.9724	0.4430
4.5	7_0.5	0.9405	2.1025	57	0.4473	0.9993
4.5	8.5	1.2551	2.1025	57	0.5970	0.9967
4.5	8.5_0.5	-0.1761	2.1025	57	-0.0838	1.0000
4.5	Ancestor	-1.0638	1.8208	57	-0.5842	0.9970
4.5_0.5	7	-9.9271	2.1025	57	-4.7216	0.0003
4.5_0.5	7_0.5	-4.8396	2.1025	57	-2.3019	0.2611
4.5_0.5	8.5	-4.5250	2.1025	57	-2.1522	0.3374
4.5_0.5	8.5_0.5	-5.9563	2.1025	57	-2.8330	0.0863
4.5_0.5	Ancestor	-6.8439	1.8208	57	-3.7587	0.0070
7	7_0.5	5.0875	2.1025	57	2.4198	0.2095
7	8.5	5.4021	2.1025	57	2.5694	0.1550
7	8.5_0.5	3.9709	2.1025	57	1.8887	0.4960
7	Ancestor	3.0832	1.8208	57	1.6933	0.6232
7_0.5	8.5	0.3146	2.1025	57	0.1496	1.0000
7_0.5	8.5_0.5	-1.1166	2.1025	57	-0.5311	0.9983
7_0.5	Ancestor	-2.0043	1.8208	57	-1.1008	0.9253
8.5	8.5_0.5	-1.4313	2.1025	57	-0.6807	0.9932
8.5	Ancestor	-2.3189	1.8208	57	-1.2735	0.8610
8.5_0.5	Ancestor	-0.8876	1.8208	57	-0.4875	0.9989
pH4.5						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	12.2	<0.0001			

Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	8.1913	1.3568	57	6.0370	0.0000
4.5	7	1.4566	1.3568	57	1.0735	0.9333
4.5	7_0.5	6.0714	1.3568	57	4.4746	0.0007
4.5	8.5	4.5917	1.3568	57	3.3841	0.0208
4.5	8.5_0.5	9.2021	1.3568	57	6.7820	0.0000
4.5	Ancestor	4.2995	1.1751	57	3.6589	0.0094
4.5_0.5	7	-6.7346	1.3568	57	-4.9634	0.0001
4.5_0.5	7_0.5	-2.1199	1.3568	57	-1.5624	0.7062
4.5_0.5	8.5	-3.5995	1.3568	57	-2.6528	0.1297
4.5_0.5	8.5_0.5	1.0109	1.3568	57	0.7450	0.9890
4.5_0.5	Ancestor	-3.8918	1.1751	57	-3.3119	0.0254
7	7_0.5	4.6148	1.3568	57	3.4011	0.0199
7	8.5	3.1351	1.3568	57	2.3106	0.2570
7	8.5_0.5	7.7455	1.3568	57	5.7084	0.0000
7	Ancestor	2.8429	1.1751	57	2.4193	0.2097
7_0.5	8.5	-1.4796	1.3568	57	-1.0905	0.9284
7_0.5	8.5_0.5	3.1307	1.3568	57	2.3074	0.2585
7_0.5	Ancestor	-1.7719	1.1751	57	-1.5079	0.7390
8.5	8.5_0.5	4.6104	1.3568	57	3.3979	0.0200
8.5	Ancestor	-0.2923	1.1751	57	-0.2487	1.0000
8.5_0.5	Ancestor	-4.9026	1.1751	57	-4.1722	0.0019
pH8.5						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	12.7	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	-0.5138	0.3826	57	-1.3426	0.8288
4.5	7	-1.6902	0.3826	57	-4.4173	0.0009
4.5	7_0.5	-0.2880	0.3826	57	-0.7527	0.9884
4.5	8.5	-2.5184	0.3826	57	-6.5816	0.0000

4.5	8.5_0.5	-0.5297	0.3826	57	-1.3845	0.8077
4.5	Ancestor	-1.6769	0.3314	57	-5.0603	0.0001
4.5_0.5	7	-1.1765	0.3826	57	-3.0747	0.0477
4.5_0.5	7_0.5	0.2258	0.3826	57	0.5900	0.9969
4.5_0.5	8.5	-2.0046	0.3826	57	-5.2389	0.0000
4.5_0.5	8.5_0.5	-0.0160	0.3826	57	-0.0418	1.0000
4.5_0.5	Ancestor	-1.1631	0.3314	57	-3.5100	0.0146
7	7_0.5	1.4022	0.3826	57	3.6647	0.0093
7	8.5	-0.8281	0.3826	57	-2.1642	0.3309
7	8.5_0.5	1.1605	0.3826	57	3.0329	0.0530
7	Ancestor	0.0134	0.3314	57	0.0404	1.0000
7_0.5	8.5	-2.2304	0.3826	57	-5.8289	0.0000
7_0.5	8.5_0.5	-0.2417	0.3826	57	-0.6318	0.9954
7_0.5	Ancestor	-1.3889	0.3314	57	-4.1912	0.0018
8.5	8.5_0.5	1.9886	0.3826	57	5.1971	0.0001
8.5	Ancestor	0.8415	0.3314	57	2.5394	0.1650
8.5_0.5	Ancestor	-1.1471	0.3314	57	-3.4617	0.0167
Salinity						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	3.1	0.01			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	2.5890	1.4366	57	1.8022	0.5522
4.5	7	-2.5361	1.4366	57	-1.7654	0.5762
4.5	7_0.5	-1.7398	1.4366	57	-1.2110	0.8870
4.5	8.5	-1.5881	1.4366	57	-1.1055	0.9239
4.5	8.5_0.5	0.9565	1.4366	57	0.6658	0.9939
4.5	Ancestor	-0.7545	1.2441	57	-0.6064	0.9964
4.5_0.5	7	-5.1251	1.4366	57	-3.5675	0.0123
4.5_0.5	7_0.5	-4.3288	1.4366	57	-3.0132	0.0557
4.5_0.5	8.5	-4.1771	1.4366	57	-2.9076	0.0722

4.5_0.5	8.5_0.5	-1.6325	1.4366	57	-1.1364	0.9140
4.5_0.5	Ancestor	-3.3435	1.2441	57	-2.6874	0.1202
7	7_0.5	0.7964	1.4366	57	0.5543	0.9978
7	8.5	0.9480	1.4366	57	0.6599	0.9942
7	8.5_0.5	3.4926	1.4366	57	2.4312	0.2049
7	Ancestor	1.7816	1.2441	57	1.4320	0.7823
7_0.5	8.5	0.1516	1.4366	57	0.1055	1.0000
7_0.5	8.5_0.5	2.6963	1.4366	57	1.8768	0.5036
7_0.5	Ancestor	0.9853	1.2441	57	0.7919	0.9849
8.5	8.5_0.5	2.5446	1.4366	57	1.7713	0.5723
8.5	Ancestor	0.8336	1.2441	57	0.6700	0.9937
8.5_0.5	Ancestor	-1.7110	1.2441	57	-1.3752	0.8125
pH4.5 + Salinity						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	16.9	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	-0.2873	0.5508	57	-0.5215	0.9984
4.5	7	0.0963	0.5508	57	0.1747	1.0000
4.5	7_0.5	2.2243	0.5508	57	4.0379	0.0029
4.5	8.5	1.4288	0.5508	57	2.5937	0.1473
4.5	8.5_0.5	0.7470	0.5508	57	1.3561	0.8222
4.5	Ancestor	-1.9700	0.4770	57	-4.1296	0.0022
4.5_0.5	7	0.3835	0.5508	57	0.6962	0.9923
4.5_0.5	7_0.5	2.5115	0.5508	57	4.5594	0.0005
4.5_0.5	8.5	1.7160	0.5508	57	3.1152	0.0430
4.5_0.5	8.5_0.5	1.0343	0.5508	57	1.8776	0.5032
4.5_0.5	Ancestor	-1.6828	0.4770	57	-3.5274	0.0139
7	7_0.5	2.1280	0.5508	57	3.8632	0.0051
7	8.5	1.3325	0.5508	57	2.4190	0.2098
7	8.5_0.5	0.6507	0.5508	57	1.1814	0.8983
7	Ancestor	-2.0663	0.4770	57	-4.3314	0.0011
7_0.5	8.5	-0.7955	0.5508	57	-1.4441	0.7756
7_0.5	8.5_0.5	-1.4773	0.5508	57	-2.6818	0.1217
7_0.5	Ancestor	-4.1943	0.4770	57	-8.7922	0.0000

8.5	8.5_0.5	-0.6818	0.5508	57	-1.2376	0.8763
8.5	Ancestor	-3.3988	0.4770	57	-7.1246	0.0000
8.5_0.5	Ancestor	-2.7170	0.4770	57	-5.6955	0.0000
pH8.5 + Salinity						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	10.9	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	-0.3426	0.1028	57	-3.3314	0.0241
4.5	7	-0.3574	0.1028	57	-3.4749	0.0161
4.5	7_0.5	-0.1440	0.1028	57	-1.4002	0.7995
4.5	8.5	-0.3926	0.1028	57	-3.8176	0.0058
4.5	8.5_0.5	-0.5680	0.1028	57	-5.5228	0.0000
4.5	Ancestor	-0.6133	0.0891	57	-6.8853	0.0000
4.5_0.5	7	-0.0147	0.1028	57	-0.1434	1.0000
4.5_0.5	7_0.5	0.1986	0.1028	57	1.9313	0.4688
4.5_0.5	8.5	-0.0500	0.1028	57	-0.4862	0.9989
4.5_0.5	8.5_0.5	-0.2254	0.1028	57	-2.1914	0.3163
4.5_0.5	Ancestor	-0.2706	0.0891	57	-3.0384	0.0523
7	7_0.5	0.2134	0.1028	57	2.0747	0.3813
7	8.5	-0.0353	0.1028	57	-0.3427	0.9999
7	8.5_0.5	-0.2106	0.1028	57	-2.0480	0.3971
7	Ancestor	-0.2559	0.0891	57	-2.8728	0.0785
7_0.5	8.5	-0.2486	0.1028	57	-2.4175	0.2104
7_0.5	8.5_0.5	-0.4240	0.1028	57	-4.1227	0.0022
7_0.5	Ancestor	-0.4692	0.0891	57	-5.2685	0.0000
8.5	8.5_0.5	-0.1754	0.1028	57	-1.7052	0.6155
8.5	Ancestor	-0.2206	0.0891	57	-2.4771	0.1872
8.5_0.5	Ancestor	-0.0452	0.0891	57	-0.5080	0.9986

Arabinose						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	6.4	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0844	0.6110	57	0.1381	1.0000
4.5	7	0.9920	0.6110	57	1.6236	0.6679
4.5	7_0.5	1.6745	0.6110	57	2.7406	0.1068
4.5	8.5	2.1728	0.6110	57	3.5561	0.0128
4.5	8.5_0.5	0.3426	0.6110	57	0.5608	0.9976
4.5	Ancestor	-0.5763	0.5291	57	-1.0891	0.9288
4.5_0.5	7	0.9076	0.6110	57	1.4855	0.7521
4.5_0.5	7_0.5	1.5901	0.6110	57	2.6026	0.1445
4.5_0.5	8.5	2.0884	0.6110	57	3.4180	0.0189
4.5_0.5	8.5_0.5	0.2582	0.6110	57	0.4227	0.9995
4.5_0.5	Ancestor	-0.6606	0.5291	57	-1.2485	0.8718
7	7_0.5	0.6825	0.6110	57	1.1170	0.9203
7	8.5	1.1808	0.6110	57	1.9325	0.4680
7	8.5_0.5	-0.6494	0.6110	57	-1.0628	0.9363
7	Ancestor	-1.5683	0.5291	57	-2.9638	0.0630
7_0.5	8.5	0.4982	0.6110	57	0.8155	0.9824
7_0.5	8.5_0.5	-1.3319	0.6110	57	-2.1799	0.3224
7_0.5	Ancestor	-2.2508	0.5291	57	-4.2537	0.0015
8.5	8.5_0.5	-1.8301	0.6110	57	-2.9954	0.0582
8.5	Ancestor	-2.7490	0.5291	57	-5.1953	0.0001
8.5_0.5	Ancestor	-0.9189	0.5291	57	-1.7366	0.5950
Citric Acid						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	8	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	1.4014	1.8491	57	0.7579	0.9880
4.5	7	-3.1892	1.8491	57	-1.7247	0.6028

4.5	7_0.5	1.8666	1.8491	57	1.0095	0.9498
4.5	8.5	0.5837	1.8491	57	0.3157	0.9999
4.5	8.5_0.5	-1.7505	1.8491	57	-0.9467	0.9630
4.5	Ancestor	-6.6809	1.6014	57	-4.1719	0.0019
4.5_0.5	7	-4.5906	1.8491	57	-2.4826	0.1851
4.5_0.5	7_0.5	0.4653	1.8491	57	0.2516	1.0000
4.5_0.5	8.5	-0.8176	1.8491	57	-0.4422	0.9994
4.5_0.5	8.5_0.5	-3.1519	1.8491	57	-1.7045	0.6159
4.5_0.5	Ancestor	-8.0823	1.6014	57	-5.0470	0.0001
7	7_0.5	5.0559	1.8491	57	2.7342	0.1083
7	8.5	3.7730	1.8491	57	2.0404	0.4016
7	8.5_0.5	1.4387	1.8491	57	0.7781	0.9862
7	Ancestor	-3.4916	1.6014	57	-2.1804	0.3222
7_0.5	8.5	-1.2829	1.8491	57	-0.6938	0.9925
7_0.5	8.5_0.5	-3.6171	1.8491	57	-1.9561	0.4532
7_0.5	Ancestor	-8.5475	1.6014	57	-5.3375	0.0000
8.5	8.5_0.5	-2.3343	1.8491	57	-1.2624	0.8659
8.5	Ancestor	-7.2646	1.6014	57	-4.5364	0.0006
8.5_0.5	Ancestor	-4.9304	1.6014	57	-3.0788	0.0472
Glucose						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	14.7	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	1.7020	1.6468	57	1.0335	0.9439
4.5	7	-2.0649	1.6468	57	-1.2539	0.8695
4.5	7_0.5	1.5136	1.6468	57	0.9191	0.9680
4.5	8.5	4.4869	1.6468	57	2.7246	0.1107
4.5	8.5_0.5	-3.5611	1.6468	57	-2.1625	0.3318
4.5	Ancestor	-6.8748	1.4262	57	-4.8205	0.0002
4.5_0.5	7	-3.7669	1.6468	57	-2.2874	0.2679

4.5_0.5	7_0.5	-0.1884	1.6468	57	-0.1144	1.0000
4.5_0.5	8.5	2.7849	1.6468	57	1.6911	0.6246
4.5_0.5	8.5_0.5	-5.2631	1.6468	57	-3.1960	0.0348
4.5_0.5	Ancestor	-8.5768	1.4262	57	-6.0139	0.0000
7	7_0.5	3.5785	1.6468	57	2.1730	0.3261
7	8.5	6.5518	1.6468	57	3.9785	0.0035
7	8.5_0.5	-1.4963	1.6468	57	-0.9086	0.9697
7	Ancestor	-4.8099	1.4262	57	-3.3726	0.0215
7_0.5	8.5	2.9733	1.6468	57	1.8055	0.5500
7_0.5	8.5_0.5	-5.0748	1.6468	57	-3.0816	0.0468
7_0.5	Ancestor	-8.3884	1.4262	57	-5.8818	0.0000
8.5	8.5_0.5	-8.0480	1.6468	57	-4.8871	0.0002
8.5	Ancestor	-11.3616	1.4262	57	-7.9666	0.0000
8.5_0.5	Ancestor	-3.3136	1.4262	57	-2.3235	0.2511
Glutamine						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	8.2	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	-0.6454	1.3745	57	-0.4695	0.9991
4.5	7	-3.1401	1.3745	57	-2.2845	0.2693
4.5	7_0.5	0.0521	1.3745	57	0.0379	1.0000
4.5	8.5	2.5838	1.3745	57	1.8797	0.5018
4.5	8.5_0.5	-2.5009	1.3745	57	-1.8194	0.5409
4.5	Ancestor	-4.7334	1.1904	57	-3.9763	0.0036
4.5_0.5	7	-2.4947	1.3745	57	-1.8150	0.5438
4.5_0.5	7_0.5	0.6975	1.3745	57	0.5074	0.9987
4.5_0.5	8.5	3.2291	1.3745	57	2.3492	0.2394
4.5_0.5	8.5_0.5	-1.8555	1.3745	57	-1.3499	0.8253
4.5_0.5	Ancestor	-4.0880	1.1904	57	-3.4342	0.0181
7	7_0.5	3.1922	1.3745	57	2.3224	0.2515

7	8.5	5.7239	1.3745	57	4.1642	0.0020
7	8.5_0.5	0.6392	1.3745	57	0.4651	0.9992
7	Ancestor	-1.5933	1.1904	57	-1.3384	0.8309
7_0.5	8.5	2.5316	1.3745	57	1.8418	0.5263
7_0.5	8.5_0.5	-2.5530	1.3745	57	-1.8574	0.5162
7_0.5	Ancestor	-4.7855	1.1904	57	-4.0201	0.0031
8.5	8.5_0.5	-5.0846	1.3745	57	-3.6992	0.0084
8.5	Ancestor	-7.3171	1.1904	57	-6.1469	0.0000
8.5_0.5	Ancestor	-2.2325	1.1904	57	-1.8754	0.5045
Histidine						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	11	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.4908	1.4070	57	0.3488	0.9998
4.5	7	-2.7190	1.4070	57	-1.9325	0.4680
4.5	7_0.5	4.1770	1.4070	57	2.9688	0.0622
4.5	8.5	5.9799	1.4070	57	4.2502	0.0015
4.5	8.5_0.5	2.8568	1.4070	57	2.0304	0.4075
4.5	Ancestor	-1.4289	1.2185	57	-1.1727	0.9014
4.5_0.5	7	-3.2098	1.4070	57	-2.2813	0.2708
4.5_0.5	7_0.5	3.6863	1.4070	57	2.6200	0.1392
4.5_0.5	8.5	5.4891	1.4070	57	3.9014	0.0045
4.5_0.5	8.5_0.5	2.3660	1.4070	57	1.6816	0.6307
4.5_0.5	Ancestor	-1.9196	1.2185	57	-1.5754	0.6981
7	7_0.5	6.8960	1.4070	57	4.9014	0.0002
7	8.5	8.6989	1.4070	57	6.1828	0.0000
7	8.5_0.5	5.5757	1.4070	57	3.9630	0.0037
7	Ancestor	1.2901	1.2185	57	1.0588	0.9374
7_0.5	8.5	1.8029	1.4070	57	1.2814	0.8575
7_0.5	8.5_0.5	-1.3203	1.4070	57	-0.9384	0.9646
7_0.5	Ancestor	-5.6059	1.2185	57	-4.6008	0.0005
8.5	8.5_0.5	-3.1231	1.4070	57	-2.2198	0.3015

8.5	Ancestor	-7.4088	1.2185	57	-6.0804	0.0000
8.5_0.5	Ancestor	-4.2856	1.2185	57	-3.5172	0.0143
Malic Acid						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	9.7	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	2.0485	2.0914	57	0.9795	0.9564
4.5	7	-1.9711	2.0914	57	-0.9425	0.9638
4.5	7_0.5	2.3661	2.0914	57	1.1314	0.9157
4.5	8.5	6.7110	2.0914	57	3.2089	0.0336
4.5	8.5_0.5	-3.8313	2.0914	57	-1.8319	0.5327
4.5	Ancestor	-5.3298	1.8112	57	-2.9427	0.0663
4.5_0.5	7	-4.0196	2.0914	57	-1.9220	0.4747
4.5_0.5	7_0.5	0.3176	2.0914	57	0.1519	1.0000
4.5_0.5	8.5	4.6625	2.0914	57	2.2294	0.2966
4.5_0.5	8.5_0.5	-5.8798	2.0914	57	-2.8114	0.0907
4.5_0.5	Ancestor	-7.3783	1.8112	57	-4.0737	0.0026
7	7_0.5	4.3373	2.0914	57	2.0739	0.3818
7	8.5	8.6821	2.0914	57	4.1514	0.0020
7	8.5_0.5	-1.8601	2.0914	57	-0.8894	0.9728
7	Ancestor	-3.3586	1.8112	57	-1.8544	0.5182
7_0.5	8.5	4.3449	2.0914	57	2.0775	0.3797
7_0.5	8.5_0.5	-6.1974	2.0914	57	-2.9633	0.0631
7_0.5	Ancestor	-7.6959	1.8112	57	-4.2491	0.0015
8.5	8.5_0.5	-10.5423	2.0914	57	-5.0408	0.0001
8.5	Ancestor	-12.0408	1.8112	57	-6.6480	0.0000
8.5_0.5	Ancestor	-1.4985	1.8112	57	-0.8274	0.9811
Maltose						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	15.2	<0.0001			

Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.8331	0.4555	57	1.8291	0.5346
4.5	7	-0.8605	0.4555	57	-1.8892	0.4957
4.5	7_0.5	0.7298	0.4555	57	1.6022	0.6815
4.5	8.5	1.1893	0.4555	57	2.6110	0.1420
4.5	8.5_0.5	0.0570	0.4555	57	0.1251	1.0000
4.5	Ancestor	-1.7861	0.3945	57	-4.5281	0.0006
4.5_0.5	7	-1.6936	0.4555	57	-3.7184	0.0079
4.5_0.5	7_0.5	-0.1034	0.4555	57	-0.2270	1.0000
4.5_0.5	8.5	0.3561	0.4555	57	0.7819	0.9858
4.5_0.5	8.5_0.5	-0.7761	0.4555	57	-1.7040	0.6163
4.5_0.5	Ancestor	-2.6193	0.3945	57	-6.6402	0.0000
7	7_0.5	1.5903	0.4555	57	3.4914	0.0154
7	8.5	2.0498	0.4555	57	4.5002	0.0006
7	8.5_0.5	0.9175	0.4555	57	2.0144	0.4172
7	Ancestor	-0.9256	0.3945	57	-2.3466	0.2406
7_0.5	8.5	0.4595	0.4555	57	1.0088	0.9499
7_0.5	8.5_0.5	-0.6728	0.4555	57	-1.4770	0.7570
7_0.5	Ancestor	-2.5159	0.3945	57	-6.3781	0.0000
8.5	8.5_0.5	-1.1323	0.4555	57	-2.4859	0.1839
8.5	Ancestor	-2.9754	0.3945	57	-7.5430	0.0000
8.5_0.5	Ancestor	-1.8431	0.3945	57	-4.6726	0.0004
Serine						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	7.6	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.5953	0.6379	57	0.9332	0.9655
4.5	7	-1.6228	0.6379	57	-2.5440	0.1634
4.5	7_0.5	0.3270	0.6379	57	0.5126	0.9986
4.5	8.5	-0.3149	0.6379	57	-0.4936	0.9988

4.5	8.5_0.5	0.5084	0.6379	57	0.7970	0.9844
4.5	Ancestor	-2.0591	0.5524	57	-3.7275	0.0077
4.5_0.5	7	-2.2180	0.6379	57	-3.4771	0.0160
4.5_0.5	7_0.5	-0.2683	0.6379	57	-0.4205	0.9995
4.5_0.5	8.5	-0.9101	0.6379	57	-1.4268	0.7851
4.5_0.5	8.5_0.5	-0.0869	0.6379	57	-0.1362	1.0000
4.5_0.5	Ancestor	-2.6544	0.5524	57	-4.8050	0.0002
7	7_0.5	1.9498	0.6379	57	3.0566	0.0499
7	8.5	1.3079	0.6379	57	2.0503	0.3956
7	8.5_0.5	2.1311	0.6379	57	3.3410	0.0235
7	Ancestor	-0.4364	0.5524	57	-0.7899	0.9851
7_0.5	8.5	-0.6419	0.6379	57	-1.0063	0.9505
7_0.5	8.5_0.5	0.1814	0.6379	57	0.2843	1.0000
7_0.5	Ancestor	-2.3861	0.5524	57	-4.3194	0.0012
8.5	8.5_0.5	0.8233	0.6379	57	1.2906	0.8534
8.5	Ancestor	-1.7443	0.5524	57	-3.1575	0.0385
8.5_0.5	Ancestor	-2.5675	0.5524	57	-4.6477	0.0004
Succinic Acid						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	6.8	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.9531	1.4088	57	0.6766	0.9934
4.5	7	-2.0194	1.4088	57	-1.4334	0.7815
4.5	7_0.5	0.3421	1.4088	57	0.2429	1.0000
4.5	8.5	3.8239	1.4088	57	2.7143	0.1133
4.5	8.5_0.5	-2.5171	1.4088	57	-1.7867	0.5622
4.5	Ancestor	-2.9784	1.2200	57	-2.4412	0.2009
4.5_0.5	7	-2.9725	1.4088	57	-2.1100	0.3610
4.5_0.5	7_0.5	-0.6110	1.4088	57	-0.4337	0.9994
4.5_0.5	8.5	2.8708	1.4088	57	2.0378	0.4031

4.5_0.5	8.5_0.5	-3.4703	1.4088	57	-2.4633	0.1924
4.5_0.5	Ancestor	-3.9315	1.2200	57	-3.2224	0.0324
7	7_0.5	2.3615	1.4088	57	1.6763	0.6342
7	8.5	5.8432	1.4088	57	4.1477	0.0021
7	8.5_0.5	-0.4978	1.4088	57	-0.3533	0.9998
7	Ancestor	-0.9590	1.2200	57	-0.7860	0.9854
7_0.5	8.5	3.4818	1.4088	57	2.4715	0.1893
7_0.5	8.5_0.5	-2.8593	1.4088	57	-2.0296	0.4080
7_0.5	Ancestor	-3.3205	1.2200	57	-2.7216	0.1114
8.5	8.5_0.5	-6.3410	1.4088	57	-4.5011	0.0006
8.5	Ancestor	-6.8023	1.2200	57	-5.5754	0.0000
8.5_0.5	Ancestor	-0.4612	1.2200	57	-0.3781	0.9998
Sucrose						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	14.1	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	1.6321	2.2773	57	0.7167	0.9910
4.5	7	-2.6741	2.2773	57	-1.1742	0.9009
4.5	7_0.5	0.5550	2.2773	57	0.2437	1.0000
4.5	8.5	5.3334	2.2773	57	2.3420	0.2427
4.5	8.5_0.5	-4.2608	2.2773	57	-1.8710	0.5074
4.5	Ancestor	-10.3035	1.9722	57	-5.2244	0.0001
4.5_0.5	7	-4.3063	2.2773	57	-1.8909	0.4946
4.5_0.5	7_0.5	-1.0771	2.2773	57	-0.4730	0.9991
4.5_0.5	8.5	3.7013	2.2773	57	1.6253	0.6669
4.5_0.5	8.5_0.5	-5.8929	2.2773	57	-2.5877	0.1492
4.5_0.5	Ancestor	-11.9356	1.9722	57	-6.0519	0.0000
7	7_0.5	3.2291	2.2773	57	1.4180	0.7899
7	8.5	8.0075	2.2773	57	3.5162	0.0143
7	8.5_0.5	-1.5866	2.2773	57	-0.6967	0.9923

7	Ancestor	-7.6294	1.9722	57	-3.8684	0.0050
7_0.5	8.5	4.7784	2.2773	57	2.0983	0.3677
7_0.5	8.5_0.5	-4.8158	2.2773	57	-2.1147	0.3583
7_0.5	Ancestor	-10.8585	1.9722	57	-5.5058	0.0000
8.5	8.5_0.5	-9.5941	2.2773	57	-4.2129	0.0017
8.5	Ancestor	-15.6369	1.9722	57	-7.9286	0.0000
8.5_0.5	Ancestor	-6.0427	1.9722	57	-3.0640	0.0490

Appendix Table D.2: Table of statistics for growth rate (r) data. One-way ANOVAs conducted per condition where growth (r) between populations from different evolutionary histories is compared, along with the ancestor.

pH7						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	5.2	0.0002			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0033	0.0015	57	2.1670	0.3294
4.5	7	-0.0035	0.0015	57	-2.3346	0.2460
4.5	7_0.5	-0.0007	0.0015	57	-0.4810	0.9990
4.5	8.5	0.0017	0.0015	57	1.1046	0.9241
4.5	8.5_0.5	0.0010	0.0015	57	0.6848	0.9930
4.5	Ancestor	-0.0021	0.0013	57	-1.6269	0.6658
4.5_0.5	7	-0.0068	0.0015	57	-4.5015	0.0006
4.5_0.5	7_0.5	-0.0040	0.0015	57	-2.6479	0.1311
4.5_0.5	8.5	-0.0016	0.0015	57	-1.0624	0.9364
4.5_0.5	8.5_0.5	-0.0022	0.0015	57	-1.4822	0.7540
4.5_0.5	Ancestor	-0.0054	0.0013	57	-4.1291	0.0022
7	7_0.5	0.0028	0.0015	57	1.8536	0.5187
7	8.5	0.0052	0.0015	57	3.4392	0.0178
7	8.5_0.5	0.0045	0.0015	57	3.0193	0.0549
7	Ancestor	0.0014	0.0013	57	1.0688	0.9346
7_0.5	8.5	0.0024	0.0015	57	1.5856	0.6918
7_0.5	8.5_0.5	0.0018	0.0015	57	1.1658	0.9039
7_0.5	Ancestor	-0.0014	0.0013	57	-1.0715	0.9338
8.5	8.5_0.5	-0.0006	0.0015	57	-0.4198	0.9995
8.5	Ancestor	-0.0038	0.0013	57	-2.9024	0.0731
8.5_0.5	Ancestor	-0.0031	0.0013	57	-2.4177	0.2103
pH4.5						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	20	<0.0001			
Residuals	57					

Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0037	0.0006	57	6.2450	0.0000
4.5	7	0.0007	0.0006	57	1.2521	0.8703
4.5	7_0.5	0.0015	0.0006	57	2.5244	0.1701
4.5	8.5	0.0029	0.0006	57	4.8968	0.0002
4.5	8.5_0.5	0.0048	0.0006	57	8.1110	0.0000
4.5	Ancestor	0.0039	0.0005	57	7.6398	0.0000
4.5_0.5	7	-0.0029	0.0006	57	-4.9929	0.0001
4.5_0.5	7_0.5	-0.0022	0.0006	57	-3.7206	0.0078
4.5_0.5	8.5	-0.0008	0.0006	57	-1.3482	0.8261
4.5_0.5	8.5_0.5	0.0011	0.0006	57	1.8660	0.5106
4.5_0.5	Ancestor	0.0002	0.0005	57	0.4288	0.9995
7	7_0.5	0.0008	0.0006	57	1.2723	0.8616
7	8.5	0.0022	0.0006	57	3.6447	0.0098
7	8.5_0.5	0.0041	0.0006	57	6.8589	0.0000
7	Ancestor	0.0032	0.0005	57	6.1940	0.0000
7_0.5	8.5	0.0014	0.0006	57	2.3724	0.2293
7_0.5	8.5_0.5	0.0033	0.0006	57	5.5866	0.0000
7_0.5	Ancestor	0.0024	0.0005	57	4.7249	0.0003
8.5	8.5_0.5	0.0019	0.0006	57	3.2142	0.0331
8.5	Ancestor	0.0010	0.0005	57	1.9855	0.4349
8.5_0.5	Ancestor	-0.0009	0.0005	57	-1.7259	0.6020
pH8.5						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	2.6	0.025			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	-0.0006	0.0013	57	-0.4489	0.9993
4.5	7	-0.0016	0.0013	57	-1.2314	0.8789
4.5	7_0.5	-0.0012	0.0013	57	-0.8693	0.9757
4.5	8.5	-0.0035	0.0013	57	-2.6129	0.1414
4.5	8.5_0.5	-0.0025	0.0013	57	-1.8852	0.4982
4.5	Ancestor	-0.0035	0.0011	57	-3.0782	0.0473

4.5_0.5	7	-0.0010	0.0013	57	-0.7825	0.9858
4.5_0.5	7_0.5	-0.0006	0.0013	57	-0.4205	0.9995
4.5_0.5	8.5	-0.0029	0.0013	57	-2.1640	0.3310
4.5_0.5	8.5_0.5	-0.0019	0.0013	57	-1.4363	0.7799
4.5_0.5	Ancestor	-0.0029	0.0011	57	-2.5599	0.1581
7	7_0.5	0.0005	0.0013	57	0.3620	0.9998
7	8.5	-0.0018	0.0013	57	-1.3816	0.8092
7	8.5_0.5	-0.0009	0.0013	57	-0.6539	0.9945
7	Ancestor	-0.0019	0.0011	57	-1.6564	0.6470
7_0.5	8.5	-0.0023	0.0013	57	-1.7436	0.5905
7_0.5	8.5_0.5	-0.0013	0.0013	57	-1.0159	0.9483
7_0.5	Ancestor	-0.0024	0.0011	57	-2.0744	0.3815
8.5	8.5_0.5	0.0010	0.0013	57	0.7277	0.9903
8.5	Ancestor	-0.0001	0.0011	57	-0.0611	1.0000
8.5_0.5	Ancestor	-0.0010	0.0011	57	-0.9013	0.9709
Salinity						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	3.4	0.006			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0008	0.0008	57	0.8953	0.9718
4.5	7	-0.0013	0.0008	57	-1.5075	0.7392
4.5	7_0.5	-0.0021	0.0008	57	-2.4698	0.1899
4.5	8.5	0.0000	0.0008	57	0.0250	1.0000
4.5	8.5_0.5	0.0009	0.0008	57	1.1292	0.9164
4.5	Ancestor	0.0001	0.0007	57	0.1495	1.0000
4.5_0.5	7	-0.0020	0.0008	57	-2.4028	0.2164
4.5_0.5	7_0.5	-0.0028	0.0008	57	-3.3651	0.0220
4.5_0.5	8.5	-0.0007	0.0008	57	-0.8703	0.9755
4.5_0.5	8.5_0.5	0.0002	0.0008	57	0.2339	1.0000

4.5_0.5	Ancestor	-0.0006	0.0007	57	-0.8843	0.9735
7	7_0.5	-0.0008	0.0008	57	-0.9623	0.9600
7	8.5	0.0013	0.0008	57	1.5326	0.7243
7	8.5_0.5	0.0022	0.0008	57	2.6367	0.1343
7	Ancestor	0.0014	0.0007	57	1.8903	0.4950
7_0.5	8.5	0.0021	0.0008	57	2.4949	0.1806
7_0.5	8.5_0.5	0.0030	0.0008	57	3.5990	0.0113
7_0.5	Ancestor	0.0022	0.0007	57	3.0014	0.0574
8.5	8.5_0.5	0.0009	0.0008	57	1.1042	0.9243
8.5	Ancestor	0.0001	0.0007	57	0.1206	1.0000
8.5_0.5	Ancestor	-0.0008	0.0007	57	-1.1544	0.9079
pH4.5 + Salinity						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	11.8	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	-0.0007	0.0012	57	-0.5991	0.9966
4.5	7	0.0003	0.0012	57	0.2707	1.0000
4.5	7_0.5	0.0053	0.0012	57	4.3432	0.0011
4.5	8.5	0.0034	0.0012	57	2.7220	0.1113
4.5	8.5_0.5	0.0020	0.0012	57	1.6322	0.6625
4.5	Ancestor	-0.0025	0.0011	57	-2.3503	0.2390
4.5_0.5	7	0.0011	0.0012	57	0.8698	0.9756
4.5_0.5	7_0.5	0.0061	0.0012	57	4.9423	0.0001
4.5_0.5	8.5	0.0041	0.0012	57	3.3211	0.0248
4.5_0.5	8.5_0.5	0.0027	0.0012	57	2.2313	0.2956
4.5_0.5	Ancestor	-0.0018	0.0011	57	-1.6585	0.6456
7	7_0.5	0.0050	0.0012	57	4.0725	0.0026
7	8.5	0.0030	0.0012	57	2.4513	0.1970
7	8.5_0.5	0.0017	0.0012	57	1.3615	0.8195
7	Ancestor	-0.0028	0.0011	57	-2.6629	0.1269
7_0.5	8.5	-0.0020	0.0012	57	-1.6212	0.6694
7_0.5	8.5_0.5	-0.0033	0.0012	57	-2.7111	0.1141
7_0.5	Ancestor	-0.0079	0.0011	57	-7.3654	0.0000

8.5	8.5_0.5	-0.0013	0.0012	57	-1.0898	0.9286
8.5	Ancestor	-0.0059	0.0011	57	-5.4934	0.0000
8.5_0.5	Ancestor	-0.0045	0.0011	57	-4.2350	0.0016
pH4.5 + Salinity						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	1.4	0.22			
Residuals	57					
Posthoc: TUKEY						
N/A						
Arabinose						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	5.1	0.0003			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0006	0.0013	57	0.4328	0.9995
4.5	7	0.0027	0.0013	57	2.0941	0.3700
4.5	7_0.5	0.0033	0.0013	57	2.5320	0.1675
4.5	8.5	0.0056	0.0013	57	4.2998	0.0013
4.5	8.5_0.5	0.0001	0.0013	57	0.0443	1.0000
4.5	Ancestor	0.0012	0.0011	57	1.0571	0.9378
4.5_0.5	7	0.0022	0.0013	57	1.6613	0.6438
4.5_0.5	7_0.5	0.0027	0.0013	57	2.0992	0.3672
4.5_0.5	8.5	0.0050	0.0013	57	3.8669	0.0050
4.5_0.5	8.5_0.5	-0.0005	0.0013	57	-0.3885	0.9997
4.5_0.5	Ancestor	0.0006	0.0011	57	0.5573	0.9977
7	7_0.5	0.0006	0.0013	57	0.4378	0.9994
7	8.5	0.0029	0.0013	57	2.2056	0.3088
7	8.5_0.5	-0.0027	0.0013	57	-2.0498	0.3960
7	Ancestor	-0.0015	0.0011	57	-1.3610	0.8197
7_0.5	8.5	0.0023	0.0013	57	1.7678	0.5746
7_0.5	8.5_0.5	-0.0032	0.0013	57	-2.4876	0.1833
7_0.5	Ancestor	-0.0021	0.0011	57	-1.8666	0.5103
8.5	8.5_0.5	-0.0055	0.0013	57	-4.2554	0.0015
8.5	Ancestor	-0.0044	0.0011	57	-3.9078	0.0044

8.5_0.5	Ancestor	0.0011	0.0011	57	1.0059	0.9506
Citric Acid						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	5.3	0.0002			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0007	0.0010	57	0.7208	0.9908
4.5	7	-0.0043	0.0010	57	-4.1749	0.0019
4.5	7_0.5	-0.0012	0.0010	57	-1.1497	0.9095
4.5	8.5	0.0001	0.0010	57	0.0977	1.0000
4.5	8.5_0.5	-0.0018	0.0010	57	-1.7804	0.5664
4.5	Ancestor	-0.0011	0.0009	57	-1.2708	0.8622
4.5_0.5	7	-0.0050	0.0010	57	-4.8957	0.0002
4.5_0.5	7_0.5	-0.0019	0.0010	57	-1.8705	0.5077
4.5_0.5	8.5	-0.0006	0.0010	57	-0.6231	0.9958
4.5_0.5	8.5_0.5	-0.0026	0.0010	57	-2.5011	0.1784
4.5_0.5	Ancestor	-0.0019	0.0009	57	-2.1030	0.3649
7	7_0.5	0.0031	0.0010	57	3.0252	0.0541
7	8.5	0.0044	0.0010	57	4.2726	0.0014
7	8.5_0.5	0.0025	0.0010	57	2.3945	0.2199
7	Ancestor	0.0032	0.0009	57	3.5500	0.0130
7_0.5	8.5	0.0013	0.0010	57	1.2474	0.8723
7_0.5	8.5_0.5	-0.0006	0.0010	57	-0.6306	0.9955
7_0.5	Ancestor	0.0001	0.0009	57	0.0568	1.0000
8.5	8.5_0.5	-0.0019	0.0010	57	-1.8780	0.5029
8.5	Ancestor	-0.0012	0.0009	57	-1.3835	0.8082
8.5_0.5	Ancestor	0.0007	0.0009	57	0.7850	0.9855
Glucose						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	7.7	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value

4.5	4.5_0.5	0.0014	0.0015	57	0.9645	0.9595
4.5	7	-0.0020	0.0015	57	-1.3319	0.8341
4.5	7_0.5	0.0000	0.0015	57	-0.0128	1.0000
4.5	8.5	0.0051	0.0015	57	3.4661	0.0165
4.5	8.5_0.5	-0.0022	0.0015	57	-1.5139	0.7354
4.5	Ancestor	-0.0027	0.0013	57	-2.1204	0.3551
4.5_0.5	7	-0.0034	0.0015	57	-2.2964	0.2637
4.5_0.5	7_0.5	-0.0014	0.0015	57	-0.9772	0.9569
4.5_0.5	8.5	0.0037	0.0015	57	2.5016	0.1782
4.5_0.5	8.5_0.5	-0.0037	0.0015	57	-2.4784	0.1867
4.5_0.5	Ancestor	-0.0041	0.0013	57	-3.2341	0.0314
7	7_0.5	0.0020	0.0015	57	1.3191	0.8402
7	8.5	0.0071	0.0015	57	4.7980	0.0002
7	8.5_0.5	-0.0003	0.0015	57	-0.1821	1.0000
7	Ancestor	-0.0007	0.0013	57	-0.5825	0.9971
7_0.5	8.5	0.0052	0.0015	57	3.4788	0.0159
7_0.5	8.5_0.5	-0.0022	0.0015	57	-1.5012	0.7429
7_0.5	Ancestor	-0.0027	0.0013	57	-2.1057	0.3634
8.5	8.5_0.5	-0.0074	0.0015	57	-4.9800	0.0001
8.5	Ancestor	-0.0079	0.0013	57	-6.1227	0.0000
8.5_0.5	Ancestor	-0.0005	0.0013	57	-0.3723	0.9998
Glutamine						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	3	0.01			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	-0.0002	0.0010	57	-0.1887	1.0000
4.5	7	-0.0025	0.0010	57	-2.4641	0.1921
4.5	7_0.5	-0.0010	0.0010	57	-0.9242	0.9671
4.5	8.5	0.0015	0.0010	57	1.5007	0.7432

4.5	8.5_0.5	-0.0014	0.0010	57	-1.3338	0.8331
4.5	Ancestor	-0.0006	0.0009	57	-0.6242	0.9957
4.5_0.5	7	-0.0023	0.0010	57	-2.2754	0.2737
4.5_0.5	7_0.5	-0.0008	0.0010	57	-0.7355	0.9897
4.5_0.5	8.5	0.0017	0.0010	57	1.6894	0.6257
4.5_0.5	8.5_0.5	-0.0012	0.0010	57	-1.1451	0.9111
4.5_0.5	Ancestor	-0.0004	0.0009	57	-0.4063	0.9996
7	7_0.5	0.0016	0.0010	57	1.5399	0.7199
7	8.5	0.0041	0.0010	57	3.9648	0.0037
7	8.5_0.5	0.0012	0.0010	57	1.1302	0.9160
7	Ancestor	0.0020	0.0009	57	2.2211	0.3008
7_0.5	8.5	0.0025	0.0010	57	2.4249	0.2074
7_0.5	8.5_0.5	-0.0004	0.0010	57	-0.4097	0.9996
7_0.5	Ancestor	0.0004	0.0009	57	0.4430	0.9994
8.5	8.5_0.5	-0.0029	0.0010	57	-2.8345	0.0860
8.5	Ancestor	-0.0021	0.0009	57	-2.3570	0.2360
8.5_0.5	Ancestor	0.0008	0.0009	57	0.9160	0.9685
Histidine						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	10.9	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0013	0.0008	57	1.7686	0.5741
4.5	7	0.0005	0.0008	57	0.6374	0.9952
4.5	7_0.5	0.0030	0.0008	57	4.0311	0.0030
4.5	8.5	0.0043	0.0008	57	5.6454	0.0000
4.5	8.5_0.5	0.0040	0.0008	57	5.3438	0.0000
4.5	Ancestor	0.0033	0.0007	57	5.0297	0.0001
4.5_0.5	7	-0.0009	0.0008	57	-1.1312	0.9157
4.5_0.5	7_0.5	0.0017	0.0008	57	2.2625	0.2800
4.5_0.5	8.5	0.0029	0.0008	57	3.8768	0.0049
4.5_0.5	8.5_0.5	0.0027	0.0008	57	3.5752	0.0121
4.5_0.5	Ancestor	0.0020	0.0007	57	2.9875	0.0594
7	7_0.5	0.0026	0.0008	57	3.3936	0.0203
7	8.5	0.0038	0.0008	57	5.0080	0.0001

7	8.5_0.5	0.0036	0.0008	57	4.7064	0.0003
7	Ancestor	0.0028	0.0007	57	4.2937	0.0013
7_0.5	8.5	0.0012	0.0008	57	1.6144	0.6738
7_0.5	8.5_0.5	0.0010	0.0008	57	1.3127	0.8432
7_0.5	Ancestor	0.0002	0.0007	57	0.3750	0.9998
8.5	8.5_0.5	-0.0002	0.0008	57	-0.3016	0.9999
8.5	Ancestor	-0.0010	0.0007	57	-1.4891	0.7500
8.5_0.5	Ancestor	-0.0007	0.0007	57	-1.1408	0.9125
Malic Acid						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	4.3	0.001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0019	0.0024	57	0.8027	0.9838
4.5	7	-0.0005	0.0024	57	-0.2331	1.0000
4.5	7_0.5	0.0015	0.0024	57	0.6395	0.9951
4.5	8.5	0.0088	0.0024	57	3.7364	0.0075
4.5	8.5_0.5	-0.0011	0.0024	57	-0.4814	0.9990
4.5	Ancestor	-0.0001	0.0020	57	-0.0577	1.0000
4.5_0.5	7	-0.0024	0.0024	57	-1.0358	0.9434
4.5_0.5	7_0.5	-0.0004	0.0024	57	-0.1632	1.0000
4.5_0.5	8.5	0.0069	0.0024	57	2.9337	0.0678
4.5_0.5	8.5_0.5	-0.0030	0.0024	57	-1.2841	0.8563
4.5_0.5	Ancestor	-0.0020	0.0020	57	-0.9846	0.9554
7	7_0.5	0.0021	0.0024	57	0.8727	0.9752
7	8.5	0.0093	0.0024	57	3.9696	0.0036
7	8.5_0.5	-0.0006	0.0024	57	-0.2483	1.0000
7	Ancestor	0.0004	0.0020	57	0.2115	1.0000
7_0.5	8.5	0.0073	0.0024	57	3.0969	0.0450
7_0.5	8.5_0.5	-0.0026	0.0024	57	-1.1209	0.9190
7_0.5	Ancestor	-0.0016	0.0020	57	-0.7961	0.9844
8.5	8.5_0.5	-0.0099	0.0024	57	-4.2179	0.0016

8.5	Ancestor	-0.0089	0.0020	57	-4.3722	0.0010
8.5_0.5	Ancestor	0.0010	0.0020	57	0.4982	0.9988
Maltose						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	5.4	0.0002			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0019	0.0008	57	2.5540	0.1600
4.5	7	-0.0017	0.0008	57	-2.1855	0.3194
4.5	7_0.5	0.0010	0.0008	57	1.3583	0.8211
4.5	8.5	0.0001	0.0008	57	0.1641	1.0000
4.5	8.5_0.5	0.0009	0.0008	57	1.1685	0.9029
4.5	Ancestor	-0.0007	0.0007	57	-1.1184	0.9198
4.5_0.5	7	-0.0036	0.0008	57	-4.7396	0.0003
4.5_0.5	7_0.5	-0.0009	0.0008	57	-1.1957	0.8929
4.5_0.5	8.5	-0.0018	0.0008	57	-2.3899	0.2218
4.5_0.5	8.5_0.5	-0.0010	0.0008	57	-1.3855	0.8072
4.5_0.5	Ancestor	-0.0027	0.0007	57	-4.0675	0.0027
7	7_0.5	0.0027	0.0008	57	3.5438	0.0132
7	8.5	0.0018	0.0008	57	2.3497	0.2392
7	8.5_0.5	0.0025	0.0008	57	3.3541	0.0226
7	Ancestor	0.0009	0.0007	57	1.4052	0.7968
7_0.5	8.5	-0.0009	0.0008	57	-1.1942	0.8935
7_0.5	8.5_0.5	-0.0001	0.0008	57	-0.1898	1.0000
7_0.5	Ancestor	-0.0018	0.0007	57	-2.6868	0.1204
8.5	8.5_0.5	0.0008	0.0008	57	1.0044	0.9509
8.5	Ancestor	-0.0009	0.0007	57	-1.3079	0.8455
8.5_0.5	Ancestor	-0.0016	0.0007	57	-2.4677	0.1907
Serine						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	8.3	<0.0001			
Residuals	57					

Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0009	0.0011	57	0.7992	0.9841
4.5	7	-0.0032	0.0011	57	-2.8732	0.0785
4.5	7_0.5	-0.0009	0.0011	57	-0.7847	0.9856
4.5	8.5	-0.0009	0.0011	57	-0.7827	0.9858
4.5	8.5_0.5	-0.0007	0.0011	57	-0.6407	0.9951
4.5	Ancestor	-0.0046	0.0010	57	-4.7629	0.0003
4.5_0.5	7	-0.0041	0.0011	57	-3.6724	0.0091
4.5_0.5	7_0.5	-0.0018	0.0011	57	-1.5839	0.6929
4.5_0.5	8.5	-0.0018	0.0011	57	-1.5820	0.6941
4.5_0.5	8.5_0.5	-0.0016	0.0011	57	-1.4400	0.7779
4.5_0.5	Ancestor	-0.0055	0.0010	57	-5.6858	0.0000
7	7_0.5	0.0023	0.0011	57	2.0885	0.3733
7	8.5	0.0023	0.0011	57	2.0905	0.3722
7	8.5_0.5	0.0025	0.0011	57	2.2324	0.2950
7	Ancestor	-0.0014	0.0010	57	-1.4453	0.7749
7_0.5	8.5	0.0000	0.0011	57	0.0020	1.0000
7_0.5	8.5_0.5	0.0002	0.0011	57	0.1439	1.0000
7_0.5	Ancestor	-0.0038	0.0010	57	-3.8569	0.0052
8.5	8.5_0.5	0.0002	0.0011	57	0.1420	1.0000
8.5	Ancestor	-0.0038	0.0010	57	-3.8591	0.0051
8.5_0.5	Ancestor	-0.0039	0.0010	57	-4.0231	0.0031
Succinic Acid						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	4.2	0.001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0000	0.0017	57	0.0169	1.0000
4.5	7	-0.0034	0.0017	57	-2.0763	0.3804
4.5	7_0.5	-0.0028	0.0017	57	-1.7020	0.6176
4.5	8.5	0.0035	0.0017	57	2.1340	0.3475

4.5	8.5_0.5	-0.0025	0.0017	57	-1.5351	0.7227
4.5	Ancestor	-0.0002	0.0014	57	-0.1429	1.0000
4.5_0.5	7	-0.0035	0.0017	57	-2.0931	0.3706
4.5_0.5	7_0.5	-0.0028	0.0017	57	-1.7189	0.6066
4.5_0.5	8.5	0.0035	0.0017	57	2.1171	0.3569
4.5_0.5	8.5_0.5	-0.0026	0.0017	57	-1.5520	0.7125
4.5_0.5	Ancestor	-0.0002	0.0014	57	-0.1624	1.0000
7	7_0.5	0.0006	0.0017	57	0.3743	0.9998
7	8.5	0.0070	0.0017	57	4.2103	0.0017
7	8.5_0.5	0.0009	0.0017	57	0.5411	0.9981
7	Ancestor	0.0032	0.0014	57	2.2545	0.2839
7_0.5	8.5	0.0064	0.0017	57	3.8360	0.0055
7_0.5	8.5_0.5	0.0003	0.0017	57	0.1668	1.0000
7_0.5	Ancestor	0.0026	0.0014	57	1.8223	0.5390
8.5	8.5_0.5	-0.0061	0.0017	57	-3.6692	0.0092
8.5	Ancestor	-0.0037	0.0014	57	-2.6071	0.1431
8.5_0.5	Ancestor	0.0023	0.0014	57	1.6297	0.6641
Sucrose						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	6.5	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0001	0.0015	57	0.0949	1.0000
4.5	7	-0.0013	0.0015	57	-0.8378	0.9798
4.5	7_0.5	-0.0011	0.0015	57	-0.7173	0.9910
4.5	8.5	0.0055	0.0015	57	3.5922	0.0115
4.5	8.5_0.5	-0.0017	0.0015	57	-1.1031	0.9246
4.5	Ancestor	-0.0024	0.0013	57	-1.7872	0.5619
4.5_0.5	7	-0.0014	0.0015	57	-0.9327	0.9656
4.5_0.5	7_0.5	-0.0012	0.0015	57	-0.8122	0.9828
4.5_0.5	8.5	0.0054	0.0015	57	3.4973	0.0151
4.5_0.5	8.5_0.5	-0.0018	0.0015	57	-1.1980	0.8921

4.5_0.5	Ancestor	-0.0025	0.0013	57	-1.8968	0.4908
7	7_0.5	0.0002	0.0015	57	0.1205	1.0000
7	8.5	0.0068	0.0015	57	4.4300	0.0008
7	8.5_0.5	-0.0004	0.0015	57	-0.2653	1.0000
7	Ancestor	-0.0011	0.0013	57	-0.8199	0.9819
7_0.5	8.5	0.0066	0.0015	57	4.3095	0.0012
7_0.5	8.5_0.5	-0.0006	0.0015	57	-0.3858	0.9997
7_0.5	Ancestor	-0.0013	0.0013	57	-0.9590	0.9606
8.5	8.5_0.5	-0.0072	0.0015	57	-4.6953	0.0003
8.5	Ancestor	-0.0079	0.0013	57	-5.9352	0.0000
8.5_0.5	Ancestor	-0.0007	0.0013	57	-0.5135	0.9986

Appendix Table D.3: Table of statistics for carrying capacity (k) data. One-way ANOVAs conducted per condition where growth (k) between populations from different evolutionary histories is compared, along with the ancestor.

pH7						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	1.5	0.19			
Residuals	57					
Posthoc: TUKEY						
N/A						
pH4.5						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	18.7	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.1505	0.0254	57	5.9212	0.0000
4.5	7	0.0304	0.0254	57	1.1959	0.8929
4.5	7_0.5	0.0667	0.0254	57	2.6253	0.1377
4.5	8.5	0.1367	0.0254	57	5.3769	0.0000
4.5	8.5_0.5	0.2175	0.0254	57	8.5559	0.0000
4.5	Ancestor	0.1415	0.0220	57	6.4264	0.0000
4.5_0.5	7	-0.1201	0.0254	57	-4.7253	0.0003
4.5_0.5	7_0.5	-0.0838	0.0254	57	-3.2958	0.0266
4.5_0.5	8.5	-0.0138	0.0254	57	-0.5442	0.9980
4.5_0.5	8.5_0.5	0.0670	0.0254	57	2.6348	0.1349
4.5_0.5	Ancestor	-0.0090	0.0220	57	-0.4108	0.9996
7	7_0.5	0.0363	0.0254	57	1.4295	0.7837
7	8.5	0.1063	0.0254	57	4.1810	0.0019
7	8.5_0.5	0.1871	0.0254	57	7.3601	0.0000
7	Ancestor	0.1111	0.0220	57	5.0455	0.0001
7_0.5	8.5	0.0699	0.0254	57	2.7516	0.1041
7_0.5	8.5_0.5	0.1507	0.0254	57	5.9306	0.0000
7_0.5	Ancestor	0.0747	0.0220	57	3.3949	0.0202
8.5	8.5_0.5	0.0808	0.0254	57	3.1790	0.0364
8.5	Ancestor	0.0048	0.0220	57	0.2177	1.0000
8.5_0.5	Ancestor	-0.0760	0.0220	57	-3.4531	0.0172
pH8.5						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	25.9	<0.0001			

Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	-0.0094	0.0078	57	-1.2142	0.8858
4.5	7	-0.0171	0.0078	57	-2.2008	0.3114
4.5	7_0.5	-0.0098	0.0078	57	-1.2543	0.8693
4.5	8.5	-0.0705	0.0078	57	-9.0641	0.0000
4.5	8.5_0.5	0.0053	0.0078	57	0.6854	0.9929
4.5	Ancestor	0.0082	0.0067	57	1.2165	0.8849
4.5_0.5	7	-0.0077	0.0078	57	-0.9866	0.9549
4.5_0.5	7_0.5	-0.0003	0.0078	57	-0.0402	1.0000
4.5_0.5	8.5	-0.0611	0.0078	57	-7.8499	0.0000
4.5_0.5	8.5_0.5	0.0148	0.0078	57	1.8996	0.4890
4.5_0.5	Ancestor	0.0176	0.0067	57	2.6185	0.1397
7	7_0.5	0.0074	0.0078	57	0.9464	0.9631
7	8.5	-0.0534	0.0078	57	-6.8633	0.0000
7	8.5_0.5	0.0225	0.0078	57	2.8862	0.0761
7	Ancestor	0.0253	0.0067	57	3.7577	0.0070
7_0.5	8.5	-0.0608	0.0078	57	-7.8098	0.0000
7_0.5	8.5_0.5	0.0151	0.0078	57	1.9397	0.4635
7_0.5	Ancestor	0.0180	0.0067	57	2.6649	0.1263
8.5	8.5_0.5	0.0759	0.0078	57	9.7495	0.0000
8.5	Ancestor	0.0787	0.0067	57	11.6828	0.0000
8.5_0.5	Ancestor	0.0029	0.0067	57	0.4251	0.9995
Salinity						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	3.2	0.009			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0226	0.0274	57	0.8240	0.9814
4.5	7	-0.0501	0.0274	57	-1.8242	0.5378
4.5	7_0.5	-0.0531	0.0274	57	-1.9335	0.4674
4.5	8.5	-0.0105	0.0274	57	-0.3837	0.9997
4.5	8.5_0.5	0.0408	0.0274	57	1.4875	0.7509
4.5	Ancestor	-0.0056	0.0238	57	-0.2336	1.0000
4.5_0.5	7	-0.0727	0.0274	57	-2.6482	0.1310
4.5_0.5	7_0.5	-0.0757	0.0274	57	-2.7575	0.1027
4.5_0.5	8.5	-0.0331	0.0274	57	-1.2078	0.8883
4.5_0.5	8.5_0.5	0.0182	0.0274	57	0.6635	0.9941
4.5_0.5	Ancestor	-0.0282	0.0238	57	-1.1851	0.8969

7	7_0.5	-0.0030	0.0274	57	-0.1093	1.0000
7	8.5	0.0395	0.0274	57	1.4404	0.7776
7	8.5_0.5	0.0909	0.0274	57	3.3117	0.0254
7	Ancestor	0.0445	0.0238	57	1.8728	0.5062
7_0.5	8.5	0.0425	0.0274	57	1.5498	0.7139
7_0.5	8.5_0.5	0.0939	0.0274	57	3.4210	0.0188
7_0.5	Ancestor	0.0475	0.0238	57	1.9990	0.4266
8.5	8.5_0.5	0.0514	0.0274	57	1.8713	0.5072
8.5	Ancestor	0.0050	0.0238	57	0.2095	1.0000
8.5_0.5	Ancestor	-0.0464	0.0238	57	-1.9512	0.4563
pH4.5 + Salinity						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	7.9	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	-0.0006	0.0186	57	-0.0297	1.0000
4.5	7	0.0109	0.0186	57	0.5877	0.9969
4.5	7_0.5	0.0886	0.0186	57	4.7727	0.0003
4.5	8.5	0.0605	0.0186	57	3.2584	0.0294
4.5	8.5_0.5	0.0523	0.0186	57	2.8176	0.0894
4.5	Ancestor	0.0061	0.0161	57	0.3769	0.9998
4.5_0.5	7	0.0115	0.0186	57	0.6175	0.9960
4.5_0.5	7_0.5	0.0892	0.0186	57	4.8024	0.0002
4.5_0.5	8.5	0.0611	0.0186	57	3.2881	0.0271
4.5_0.5	8.5_0.5	0.0529	0.0186	57	2.8473	0.0834
4.5_0.5	Ancestor	0.0066	0.0161	57	0.4112	0.9996
7	7_0.5	0.0777	0.0186	57	4.1849	0.0018
7	8.5	0.0496	0.0186	57	2.6707	0.1248
7	8.5_0.5	0.0414	0.0186	57	2.2299	0.2963
7	Ancestor	-0.0049	0.0161	57	-0.3018	0.9999
7_0.5	8.5	-0.0281	0.0186	57	-1.5142	0.7352
7_0.5	8.5_0.5	-0.0363	0.0186	57	-1.9551	0.4539
7_0.5	Ancestor	-0.0826	0.0161	57	-5.1341	0.0001
8.5	8.5_0.5	-0.0082	0.0186	57	-0.4408	0.9994
8.5	Ancestor	-0.0545	0.0161	57	-3.3856	0.0207
8.5_0.5	Ancestor	-0.0463	0.0161	57	-2.8766	0.0778
pH8.5 + Salinity						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	6	<0.0001			

Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	-0.0087	0.0025	57	-3.4302	0.0183
4.5	7	-0.0058	0.0025	57	-2.2841	0.2695
4.5	7_0.5	-0.0039	0.0025	57	-1.5296	0.7261
4.5	8.5	-0.0104	0.0025	57	-4.1105	0.0023
4.5	8.5_0.5	-0.0133	0.0025	57	-5.2607	0.0000
4.5	Ancestor	-0.0071	0.0022	57	-3.2610	0.0292
4.5_0.5	7	0.0029	0.0025	57	1.1461	0.9107
4.5_0.5	7_0.5	0.0048	0.0025	57	1.9006	0.4883
4.5_0.5	8.5	-0.0017	0.0025	57	-0.6803	0.9932
4.5_0.5	8.5_0.5	-0.0046	0.0025	57	-1.8305	0.5336
4.5_0.5	Ancestor	0.0015	0.0022	57	0.6998	0.9921
7	7_0.5	0.0019	0.0025	57	0.7545	0.9882
7	8.5	-0.0046	0.0025	57	-1.8264	0.5363
7	8.5_0.5	-0.0075	0.0025	57	-2.9767	0.0610
7	Ancestor	-0.0014	0.0022	57	-0.6236	0.9958
7_0.5	8.5	-0.0065	0.0025	57	-2.5809	0.1513
7_0.5	8.5_0.5	-0.0094	0.0025	57	-3.7312	0.0076
7_0.5	Ancestor	-0.0033	0.0022	57	-1.4948	0.7466
8.5	8.5_0.5	-0.0029	0.0025	57	-1.1503	0.9093
8.5	Ancestor	0.0033	0.0022	57	1.4853	0.7522
8.5_0.5	Ancestor	0.0062	0.0022	57	2.8135	0.0903
Arabinose						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	4	0.002			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0006	0.0135	57	0.0423	1.0000
4.5	7	0.0150	0.0135	57	1.1094	0.9227
4.5	7_0.5	0.0340	0.0135	57	2.5099	0.1752
4.5	8.5	0.0457	0.0135	57	3.3791	0.0211
4.5	8.5_0.5	-0.0048	0.0135	57	-0.3549	0.9998
4.5	Ancestor	0.0096	0.0117	57	0.8232	0.9815
4.5_0.5	7	0.0144	0.0135	57	1.0671	0.9351
4.5_0.5	7_0.5	0.0334	0.0135	57	2.4675	0.1908
4.5_0.5	8.5	0.0451	0.0135	57	3.3367	0.0238
4.5_0.5	8.5_0.5	-0.0054	0.0135	57	-0.3973	0.9997
4.5_0.5	Ancestor	0.0091	0.0117	57	0.7743	0.9865

7	7_0.5	0.0189	0.0135	57	1.4004	0.7993
7	8.5	0.0307	0.0135	57	2.2697	0.2765
7	8.5_0.5	-0.0198	0.0135	57	-1.4643	0.7642
7	Ancestor	-0.0054	0.0117	57	-0.4578	0.9992
7_0.5	8.5	0.0118	0.0135	57	0.8692	0.9757
7_0.5	8.5_0.5	-0.0388	0.0135	57	-2.8648	0.0800
7_0.5	Ancestor	-0.0243	0.0117	57	-2.0749	0.3812
8.5	8.5_0.5	-0.0505	0.0135	57	-3.7340	0.0075
8.5	Ancestor	-0.0361	0.0117	57	-3.0786	0.0472
8.5_0.5	Ancestor	0.0144	0.0117	57	1.2330	0.8782
Citric Acid						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	8.6	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0258	0.0277	57	0.9322	0.9657
4.5	7	-0.0895	0.0277	57	-3.2333	0.0315
4.5	7_0.5	-0.0460	0.0277	57	-1.6631	0.6427
4.5	8.5	-0.0053	0.0277	57	-0.1908	1.0000
4.5	8.5_0.5	-0.0781	0.0277	57	-2.8227	0.0884
4.5	Ancestor	-0.1095	0.0240	57	-4.5679	0.0005
4.5_0.5	7	-0.1153	0.0277	57	-4.1655	0.0020
4.5_0.5	7_0.5	-0.0718	0.0277	57	-2.5954	0.1468
4.5_0.5	8.5	-0.0311	0.0277	57	-1.1231	0.9184
4.5_0.5	8.5_0.5	-0.1039	0.0277	57	-3.7549	0.0071
4.5_0.5	Ancestor	-0.1353	0.0240	57	-5.6444	0.0000
7	7_0.5	0.0435	0.0277	57	1.5702	0.7014
7	8.5	0.0842	0.0277	57	3.0425	0.0517
7	8.5_0.5	0.0114	0.0277	57	0.4106	0.9996
7	Ancestor	-0.0200	0.0240	57	-0.8344	0.9802
7_0.5	8.5	0.0408	0.0277	57	1.4723	0.7597
7_0.5	8.5_0.5	-0.0321	0.0277	57	-1.1596	0.9061
7_0.5	Ancestor	-0.0635	0.0240	57	-2.6475	0.1312
8.5	8.5_0.5	-0.0728	0.0277	57	-2.6319	0.1357
8.5	Ancestor	-0.1042	0.0240	57	-4.3476	0.0011
8.5_0.5	Ancestor	-0.0314	0.0240	57	-1.3085	0.8452
Glucose						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	11.5	<0.0001			

Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0149	0.0294	57	0.5078	0.9986
4.5	7	-0.0593	0.0294	57	-2.0169	0.4157
4.5	7_0.5	-0.0092	0.0294	57	-0.3112	0.9999
4.5	8.5	0.0873	0.0294	57	2.9684	0.0623
4.5	8.5_0.5	-0.0792	0.0294	57	-2.6932	0.1187
4.5	Ancestor	-0.0985	0.0255	57	-3.8679	0.0050
4.5_0.5	7	-0.0743	0.0294	57	-2.5247	0.1700
4.5_0.5	7_0.5	-0.0241	0.0294	57	-0.8190	0.9820
4.5_0.5	8.5	0.0724	0.0294	57	2.4606	0.1934
4.5_0.5	8.5_0.5	-0.0942	0.0294	57	-3.2010	0.0343
4.5_0.5	Ancestor	-0.1135	0.0255	57	-4.4543	0.0008
7	7_0.5	0.0502	0.0294	57	1.7057	0.6152
7	8.5	0.1467	0.0294	57	4.9853	0.0001
7	8.5_0.5	-0.0199	0.0294	57	-0.6763	0.9934
7	Ancestor	-0.0392	0.0255	57	-1.5390	0.7204
7_0.5	8.5	0.0965	0.0294	57	3.2796	0.0278
7_0.5	8.5_0.5	-0.0701	0.0294	57	-2.3820	0.2252
7_0.5	Ancestor	-0.0894	0.0255	57	-3.5085	0.0146
8.5	8.5_0.5	-0.1666	0.0294	57	-5.6616	0.0000
8.5	Ancestor	-0.1859	0.0255	57	-7.2955	0.0000
8.5_0.5	Ancestor	-0.0193	0.0255	57	-0.7580	0.9879
Glutamine						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	5.4	0.0002			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	-0.0046	0.0204	57	-0.2252	1.0000
4.5	7	-0.0146	0.0204	57	-0.7171	0.9910
4.5	7_0.5	-0.0231	0.0204	57	-1.1373	0.9137
4.5	8.5	0.0587	0.0204	57	2.8847	0.0763
4.5	8.5_0.5	-0.0299	0.0204	57	-1.4700	0.7610
4.5	Ancestor	-0.0355	0.0176	57	-2.0124	0.4184
4.5_0.5	7	-0.0100	0.0204	57	-0.4919	0.9989
4.5_0.5	7_0.5	-0.0186	0.0204	57	-0.9121	0.9691
4.5_0.5	8.5	0.0633	0.0204	57	3.1099	0.0436
4.5_0.5	8.5_0.5	-0.0253	0.0204	57	-1.2448	0.8734
4.5_0.5	Ancestor	-0.0309	0.0176	57	-1.7524	0.5847

7	7_0.5	-0.0086	0.0204	57	-0.4202	0.9995
7	8.5	0.0733	0.0204	57	3.6018	0.0112
7	8.5_0.5	-0.0153	0.0204	57	-0.7529	0.9884
7	Ancestor	-0.0209	0.0176	57	-1.1844	0.8972
7_0.5	8.5	0.0819	0.0204	57	4.0220	0.0031
7_0.5	8.5_0.5	-0.0068	0.0204	57	-0.3327	0.9999
7_0.5	Ancestor	-0.0123	0.0176	57	-0.6992	0.9921
8.5	8.5_0.5	-0.0886	0.0204	57	-4.3547	0.0010
8.5	Ancestor	-0.0942	0.0176	57	-5.3433	0.0000
8.5_0.5	Ancestor	-0.0056	0.0176	57	-0.3150	0.9999
Histidine						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	10.8	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	-0.0333	0.0229	57	-1.4586	0.7675
4.5	7	-0.0663	0.0229	57	-2.8998	0.0736
4.5	7_0.5	0.0264	0.0229	57	1.1545	0.9079
4.5	8.5	0.0887	0.0229	57	3.8830	0.0048
4.5	8.5_0.5	0.0625	0.0229	57	2.7339	0.1084
4.5	Ancestor	0.0099	0.0198	57	0.4995	0.9988
4.5_0.5	7	-0.0329	0.0229	57	-1.4412	0.7772
4.5_0.5	7_0.5	0.0597	0.0229	57	2.6131	0.1413
4.5_0.5	8.5	0.1221	0.0229	57	5.3416	0.0000
4.5_0.5	8.5_0.5	0.0958	0.0229	57	4.1925	0.0018
4.5_0.5	Ancestor	0.0432	0.0198	57	2.1837	0.3204
7	7_0.5	0.0927	0.0229	57	4.0544	0.0028
7	8.5	0.1550	0.0229	57	6.7828	0.0000
7	8.5_0.5	0.1288	0.0229	57	5.6337	0.0000
7	Ancestor	0.0762	0.0198	57	3.8479	0.0053
7_0.5	8.5	0.0624	0.0229	57	2.7284	0.1097
7_0.5	8.5_0.5	0.0361	0.0229	57	1.5794	0.6957
7_0.5	Ancestor	-0.0165	0.0198	57	-0.8337	0.9803
8.5	8.5_0.5	-0.0263	0.0229	57	-1.1491	0.9097
8.5	Ancestor	-0.0789	0.0198	57	-3.9842	0.0035
8.5_0.5	Ancestor	-0.0526	0.0198	57	-2.6574	0.1284
Malic Acid						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	10.9	<0.0001			

Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0262	0.0264	57	0.9922	0.9537
4.5	7	-0.0382	0.0264	57	-1.4439	0.7757
4.5	7_0.5	0.0019	0.0264	57	0.0725	1.0000
4.5	8.5	0.1100	0.0264	57	4.1611	0.0020
4.5	8.5_0.5	-0.0519	0.0264	57	-1.9651	0.4476
4.5	Ancestor	-0.0564	0.0229	57	-2.4657	0.1915
4.5_0.5	7	-0.0644	0.0264	57	-2.4361	0.2029
4.5_0.5	7_0.5	-0.0243	0.0264	57	-0.9197	0.9679
4.5_0.5	8.5	0.0837	0.0264	57	3.1689	0.0373
4.5_0.5	8.5_0.5	-0.0781	0.0264	57	-2.9572	0.0640
4.5_0.5	Ancestor	-0.0826	0.0229	57	-3.6114	0.0109
7	7_0.5	0.0401	0.0264	57	1.5165	0.7339
7	8.5	0.1481	0.0264	57	5.6050	0.0000
7	8.5_0.5	-0.0138	0.0264	57	-0.5211	0.9984
7	Ancestor	-0.0183	0.0229	57	-0.7984	0.9842
7_0.5	8.5	0.1080	0.0264	57	4.0886	0.0025
7_0.5	8.5_0.5	-0.0538	0.0264	57	-2.0376	0.4033
7_0.5	Ancestor	-0.0583	0.0229	57	-2.5494	0.1616
8.5	8.5_0.5	-0.1619	0.0264	57	-6.1262	0.0000
8.5	Ancestor	-0.1664	0.0229	57	-7.2705	0.0000
8.5_0.5	Ancestor	-0.0045	0.0229	57	-0.1966	1.0000
Maltose						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	11.3	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0212	0.0093	57	2.2763	0.2733
4.5	7	-0.0162	0.0093	57	-1.7357	0.5956
4.5	7_0.5	0.0077	0.0093	57	0.8220	0.9817
4.5	8.5	0.0276	0.0093	57	2.9576	0.0640
4.5	8.5_0.5	-0.0007	0.0093	57	-0.0737	1.0000
4.5	Ancestor	-0.0268	0.0081	57	-3.3145	0.0253
4.5_0.5	7	-0.0374	0.0093	57	-4.0119	0.0032
4.5_0.5	7_0.5	-0.0136	0.0093	57	-1.4542	0.7699
4.5_0.5	8.5	0.0064	0.0093	57	0.6813	0.9932
4.5_0.5	8.5_0.5	-0.0219	0.0093	57	-2.3500	0.2391
4.5_0.5	Ancestor	-0.0480	0.0081	57	-5.9429	0.0000

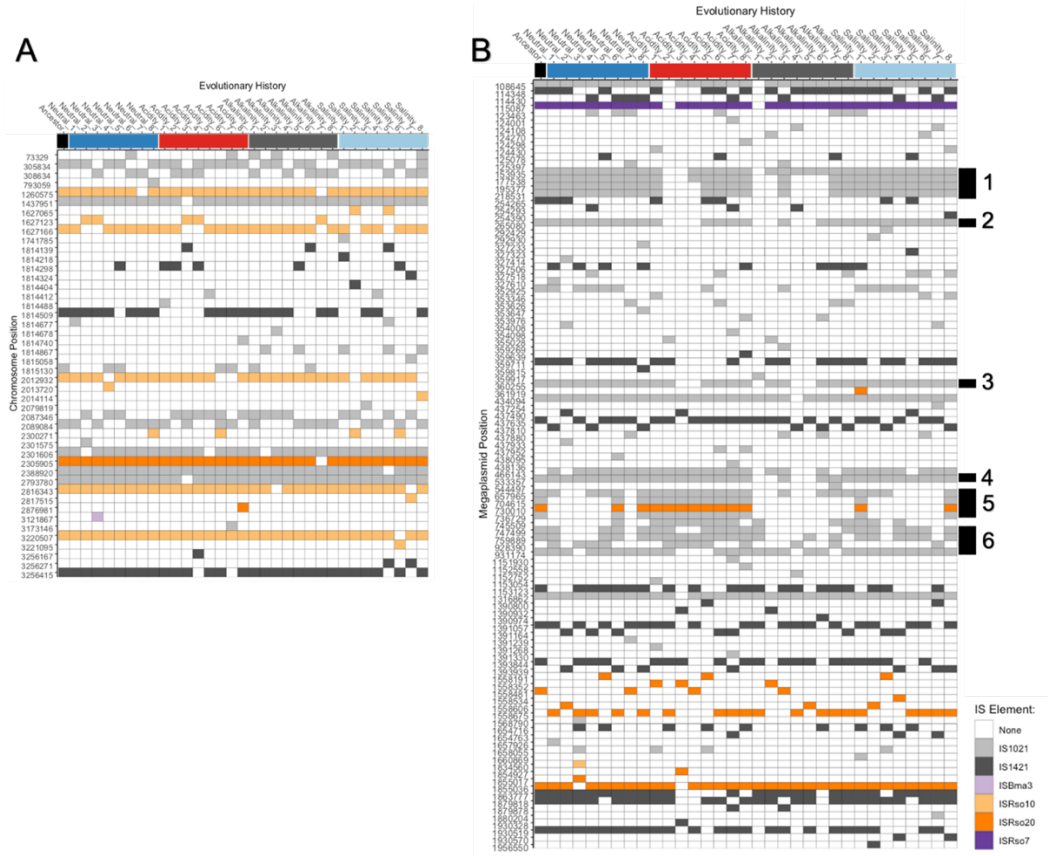
7	7_0.5	0.0239	0.0093	57	2.5577	0.1588
7	8.5	0.0438	0.0093	57	4.6933	0.0003
7	8.5_0.5	0.0155	0.0093	57	1.6620	0.6434
7	Ancestor	-0.0106	0.0081	57	-1.3103	0.8443
7_0.5	8.5	0.0199	0.0093	57	2.1355	0.3466
7_0.5	8.5_0.5	-0.0084	0.0093	57	-0.8958	0.9718
7_0.5	Ancestor	-0.0344	0.0081	57	-4.2637	0.0014
8.5	8.5_0.5	-0.0283	0.0093	57	-3.0313	0.0532
8.5	Ancestor	-0.0544	0.0081	57	-6.7296	0.0000
8.5_0.5	Ancestor	-0.0261	0.0081	57	-3.2294	0.0318
Serine						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	13	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0099	0.0182	57	0.5417	0.9981
4.5	7	-0.0542	0.0182	57	-2.9770	0.0610
4.5	7_0.5	-0.0186	0.0182	57	-1.0198	0.9473
4.5	8.5	-0.0196	0.0182	57	-1.0753	0.9328
4.5	8.5_0.5	-0.0116	0.0182	57	-0.6367	0.9953
4.5	Ancestor	-0.0994	0.0158	57	-6.3097	0.0000
4.5_0.5	7	-0.0640	0.0182	57	-3.5186	0.0142
4.5_0.5	7_0.5	-0.0284	0.0182	57	-1.5615	0.7067
4.5_0.5	8.5	-0.0294	0.0182	57	-1.6170	0.6721
4.5_0.5	8.5_0.5	-0.0214	0.0182	57	-1.1784	0.8994
4.5_0.5	Ancestor	-0.1093	0.0158	57	-6.9352	0.0000
7	7_0.5	0.0356	0.0182	57	1.9572	0.4525
7	8.5	0.0346	0.0182	57	1.9016	0.4877
7	8.5_0.5	0.0426	0.0182	57	2.3402	0.2435
7	Ancestor	-0.0453	0.0158	57	-2.8722	0.0786
7_0.5	8.5	-0.0010	0.0182	57	-0.0555	1.0000
7_0.5	8.5_0.5	0.0070	0.0182	57	0.3831	0.9997
7_0.5	Ancestor	-0.0809	0.0158	57	-5.1321	0.0001
8.5	8.5_0.5	0.0080	0.0182	57	0.4386	0.9994
8.5	Ancestor	-0.0798	0.0158	57	-5.0680	0.0001
8.5_0.5	Ancestor	-0.0878	0.0158	57	-5.5744	0.0000
Succinic Acid						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	7	0.001			

Residuals	57					
Posthoc: TUKEY						
4.5	4.5_0.5	0.0073	0.0192	57	0.3826	0.9997
4.5	7	-0.0288	0.0192	57	-1.5020	0.7425
4.5	7_0.5	-0.0203	0.0192	57	-1.0549	0.9384
4.5	8.5	0.0618	0.0192	57	3.2194	0.0327
4.5	8.5_0.5	-0.0463	0.0192	57	-2.4120	0.2127
4.5	Ancestor	-0.0244	0.0166	57	-1.4662	0.7632
4.5_0.5	7	-0.0362	0.0192	57	-1.8845	0.4987
4.5_0.5	7_0.5	-0.0276	0.0192	57	-1.4374	0.7793
4.5_0.5	8.5	0.0545	0.0192	57	2.8369	0.0855
4.5_0.5	8.5_0.5	-0.0536	0.0192	57	-2.7945	0.0944
4.5_0.5	Ancestor	-0.0317	0.0166	57	-1.9079	0.4837
7	7_0.5	0.0086	0.0192	57	0.4471	0.9993
7	8.5	0.0906	0.0192	57	4.7214	0.0003
7	8.5_0.5	-0.0175	0.0192	57	-0.9100	0.9695
7	Ancestor	0.0045	0.0166	57	0.2682	1.0000
7_0.5	8.5	0.0821	0.0192	57	4.2743	0.0014
7_0.5	8.5_0.5	-0.0261	0.0192	57	-1.3571	0.8217
7_0.5	Ancestor	-0.0041	0.0166	57	-0.2481	1.0000
8.5	8.5_0.5	-0.1081	0.0192	57	-5.6314	0.0000
8.5	Ancestor	-0.0862	0.0166	57	-5.1836	0.0001
8.5_0.5	Ancestor	0.0219	0.0166	57	1.3189	0.8403
8.5_0.5	Ancestor					
Sucrose						
Overall ANOVA						
	DF	F value	p			
EvolvedIn	6	10.5	<0.0001			
Residuals	57					
Posthoc: TUKEY						
EvolvedIn1	EvolvedIn2	estimate	SE	df	t.ratio	p.value
4.5	4.5_0.5	0.0513	0.0380	57	1.3514	0.8245
4.5	7	-0.0393	0.0380	57	-1.0345	0.9437
4.5	7_0.5	-0.0169	0.0380	57	-0.4445	0.9994
4.5	8.5	0.1429	0.0380	57	3.7656	0.0068
4.5	8.5_0.5	-0.0639	0.0380	57	-1.6843	0.6290
4.5	Ancestor	-0.0955	0.0329	57	-2.9036	0.0729
4.5_0.5	7	-0.0906	0.0380	57	-2.3859	0.2235
4.5_0.5	7_0.5	-0.0682	0.0380	57	-1.7959	0.5562
4.5_0.5	8.5	0.0916	0.0380	57	2.4142	0.2118
4.5_0.5	8.5_0.5	-0.1152	0.0380	57	-3.0357	0.0526
4.5_0.5	Ancestor	-0.1468	0.0329	57	-4.4641	0.0007

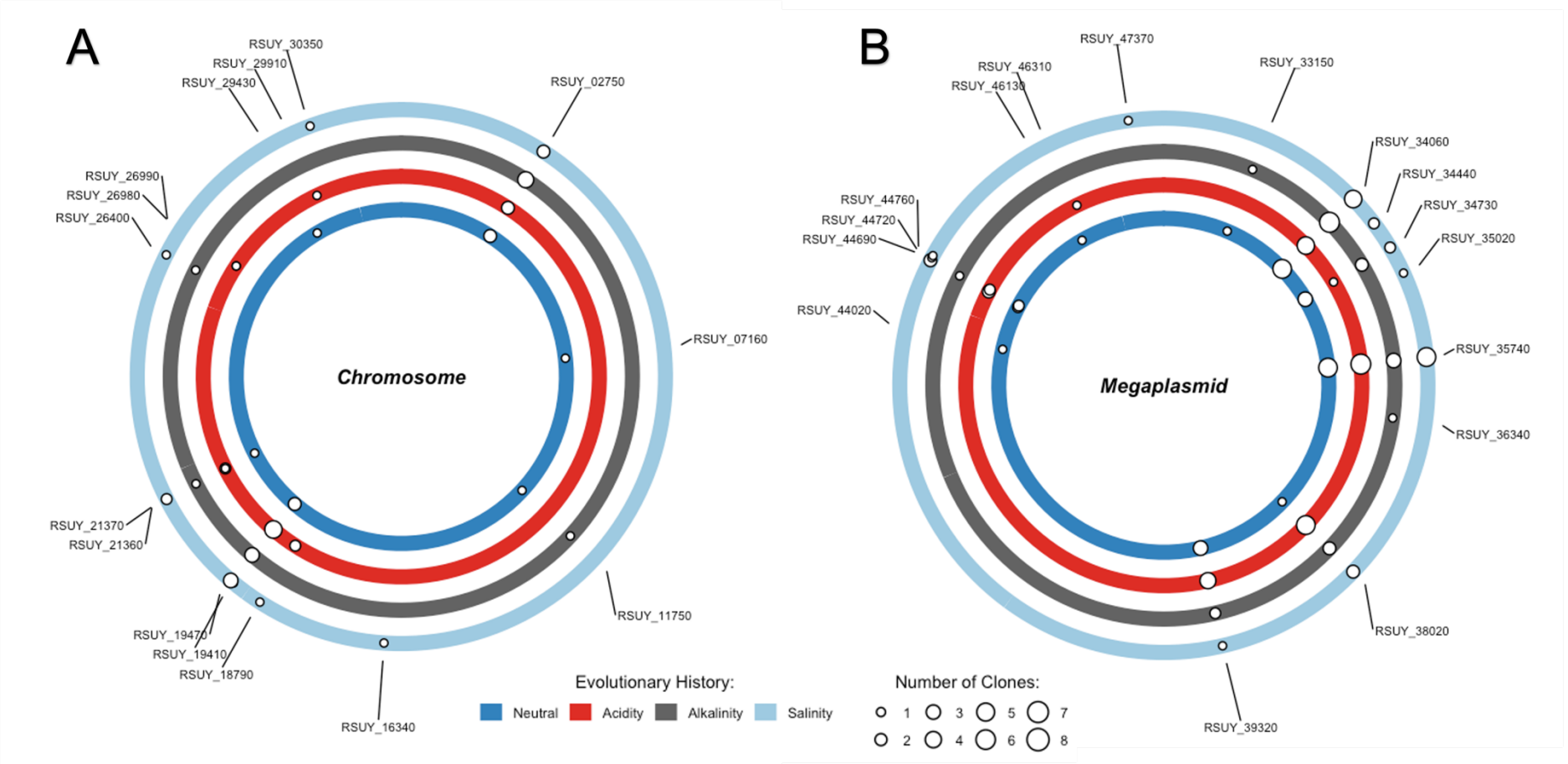
7	7_0.5	0.0224	0.0380	57	0.5900	0.9969
7	8.5	0.1822	0.0380	57	4.8001	0.0002
7	8.5_0.5	-0.0247	0.0380	57	-0.6498	0.9947
7	Ancestor	-0.0562	0.0329	57	-1.7091	0.6129
7_0.5	8.5	0.1598	0.0380	57	4.2101	0.0017
7_0.5	8.5_0.5	-0.0471	0.0380	57	-1.2397	0.8755
7_0.5	Ancestor	-0.0786	0.0329	57	-2.3903	0.2217
8.5	8.5_0.5	-0.2069	0.0380	57	-5.4498	0.0000
8.5	Ancestor	-0.2384	0.0329	57	-7.2517	0.0000
8.5_0.5	Ancestor	-0.0315	0.0329	57	-0.9588	0.9607

Appendix Table D.4: Table of filtered genetic variants present within evolved clones. Where CP012687 is the chromosome and CP012688 is the megaplasmid and position refers to place within their respective chromosome. Reference allele refers to the allele present in the ancestor (YO336) and alternative allele the genetic variant present in the clones highlighted. Predicted gene function was taken from the annotated reference file. Where each clone has a highlighted section, this indicates that the alternative allele is present within the genome, whereas no colour indicates that the reference allele is present in that clone. Colour indicates evolved condition, where blue shows neutral evolved clones (pH 7 0% NaCl), red acidic (pH 4.5 0% NaCl), grey alkaline (pH 8.5 0% NaCl) and light blue salinity (pH 7 0.5% NaCl) evolved clones.

Pipeline Discovered In	SNIPPY	Both	Mine	Both	Both	Both	Both	SNIPPY	SNIPPY	Both	Mine	Mine	Mine	Mine	Mine	Mine	Mine	Both	Both	Both	Mine	Mine
Looked at manually?	YES		YES			YES	YES	YES	YES		YES	YES										
Comments:	Manual check confirmed variant		Manual check confirmed variant			Manual check confirmed variant	Manual check confirmed variant	Manual check confirmed variant	Manual check confirmed variant		In all clones but not ancestor. Manually changed.	Manual check confirmed variant						Mine found more clones with this variant compared to SNIPPY	Mine found more clones with this variant compared to SNIPPY	Mine found more clones with this variant compared to SNIPPY		
Chromosome	CP012687	CP012687	CP012687	CP012687	CP012687	CP012687	CP012687	CP012687	CP012687	CP012687	CP012688	CP012688	CP012688	CP012688	CP012688	CP012688	CP012688	CP012688	CP012688	CP012688	CP012688	CP012688
Locus Tag	RSUY_RS04350	RSUY_RS14065	RSUY_RS01195	RSUY_RS03495	RSUY_RS03550	RSUY_RS03815	RSUY_RS03815	RSUY_RS03815	RSUY_RS03815	RSUY_RS04095	RSUY_RS21215	RSUY_RS21740	RSUY_RS17095	RSUY_RS17095	RSUY_RS17095	RSUY_RS17095	RSUY_RS17095	RSUY_RS17095	RSUY_RS17095	RSUY_RS17095	RSUY_RS17095	RSUY_RS17095
Position	978476	3061170	658326	793548	805796	863455	863590	863797	863952	926782	1500065	1649968	356411	356488	356576	356722	356737	357148	357163	357235	357255	357400
Reference Allele	TGGCA	CCCCGGCCGCGCGGACGGCCCAGGCGGCTCCGCCCGACGGCAGCAGCCGCGCGGCGGTCCAGCTCGAAAGGTC	TCTAGAGGCTAGAGGCTAGAGGCTAGAGGCTAGAGGCTAGAGGCTAGAGGC	C	G	C	C	G	C	T	C	T	A	GCCTA	GGC	G	T	C	C	CTAC	TCTT	GAAC
Alternative Allele	A	CC	TCTAGAGGCTAGAGGCTAGAGGCTAGAGGCTAGAGGCTAGAGGCTAGAGGC	T	A	G	T	T	G	G	T	C	G	ACCTA	GGT	A	C	T	T	CTAT	TCTC	CAAT
Predicted Gene Function	Hpt domain-containing protein	tetratricopeptide repeat protein	tRNAscan-SE	LysR family transcriptional regulator	hybrid sensor histidine kinase/response regulator	type IV pilus twitching motility protein	type IV pilus twitching motility protein	type IV pilus twitching motility protein	type IV pilus twitching motility protein	AI-2E family transporter	Ig-like domain-containing protein	IS5 family transposase	DUF2778 domain-containing protein	DUF2778 domain-containing protein	DUF2778 domain-containing protein	DUF2778 domain-containing protein	DUF2778 domain-containing protein	DUF2778 domain-containing protein	DUF2778 domain-containing protein	DUF2778 domain-containing protein	DUF2778 domain-containing protein	DUF2778 domain-containing protein
Predicted Mutation	frameshift_variant	frameshift_variant	intragenic_variant	missense_variant	missense_variant	missense_variant	missense_variant	missense_variant	missense_variant	missense_variant	synonymous_variant	missense_variant	missense_variant	synonymous_variant	synonymous_variant	synonymous_variant	synonymous_variant	synonymous_variant	synonymous_variant	synonymous_variant	missense_variant	synonymous_variant
Predicted Impact		HIGH	MODIFIER	MODERATE	MODERATE	MODERATE	MODERATE			MODERATE	LOW	MODERATE	MODERATE	LOW	LOW	LOW	LOW	LOW	LOW	LOW	MODERATE	LOW
neutral_1																						
neutral_2																						
neutral_3																						
neutral_4																						
neutral_5																						
neutral_6																						
neutral_7																						
neutral_8																						
acid_1																						
acid_2																						
acid_3																						
acid_4																						
acid_5																						
acid_6																						
acid_7																						
acid_8																						
alkaline_1																						
alkaline_2																						
alkaline_3																						
alkaline_4																						
alkaline_5																						
alkaline_6																						
alkaline_7																						
alkaline_8																						
salinity_1																						
salinity_2																						
salinity_3																						
salinity_4																						
salinity_5																						
salinity_6																						
salinity_7																						
salintiy_8																						

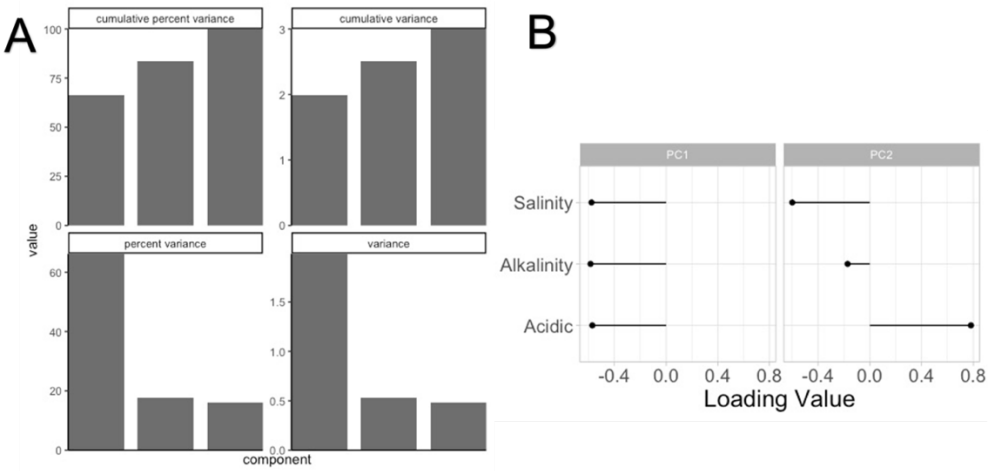


Appendix Figure D.7: Insertion sequence (IS) movement contributes to stress adaptation. Presence (coloured) and absence (white) of IS across genomic positions across the (A) chromosome and (B) megaplasmid of *R. solanacearum*. Specific colours indicate the type of IS present. Evolutionary history indicates the stress conditions in which the clone was exposed to during the selection experiment (n=8 per stress condition) plus the ancestor clone. Black numbered blocks along the right-hand side indicate the areas of interest which differ between different evolutionary history groups. Block 1 includes positions near a PhoPQ-activated pathogenicity-related protein, an endoglucanase precursor, a thioesterase superfamily protein, a TRP repeat-containing protein (YrrB), a leukotokin, and hypothetical proteins. Block 2 involves a cyclic di-GMP phosphodiesterase Gmr and a H-NS histone family protein. Block 3 a hypothetical protein and integrase core domain protein. Block 4 a tRNA-Pro and a filamentous hemagglutinin. 5 an aminopeptidase N, a type II secretion system protein F, a type II secretion system protein G precursor and hypothetical proteins. Finally, block 6 includes a putative deoxyribonuclease (RhsB), tRNA3(Ser)-specific nuclease WapA precursor, minor extracellular protease Epr precursor and hypothetical protein



Appendix Figure D.8: Insertion sequence (IS) movement within adapted clones compared to the ancestor. IS found in the chromosome (left, ~3.5Mbp) and megaplasmid (right, ~2Mbp) within different positions compared to the ancestor. Colour indicates the evolutionary history of the clones (N=8 per evolved condition), either from neutral (pH 7 with 0% NaCl), acidic (pH 4.5 with 0% NaCl), alkalinity (pH 8.5 with 0% NaCl) and salinity (pH 7 with 0.5% NaCl) stress conditions. Each circle represents a gene in which

an insertion sequence has been called in the *R. solanacearum* genome, size of circles indicates the number of clones per evolutionary history that has that IS present within that position in the genome.



Appendix Figure D.9: Variance (A) and loading (B) values used to calculate the PCA plot figure 4.8.

References

Adhikari, T. B., and R. C. Basnyat. 1998. Effect of crop rotation and cultivar resistance on bacterial wilt of tomato in Nepal. Canadian Journal of Plant Pathology 20:283–287.

Ailloud, F., T. Lowe, G. Cellier, D. Roche, C. Allen, and P. Prior. 2011. Comparative genomic analysis of *Ralstonia solanacearum* reveals candidate genes for host specificity. BMC Genomics 16:270.

Alderley, C. L., S. T. E. Greenrod, and V. Friman. 2022. Plant pathogenic bacterium can rapidly evolve tolerance to an antimicrobial plant allelochemical. Evolutionary Applications 15:735–750.

Allen, H. K., J. Donato, H. H. Wang, K. A. Cloud-Hansen, J. Davies, and J. Handelsman. 2010. Call of the wild: Antibiotic resistance genes in natural environments. Nature Reviews Microbiology 8:251–259.

Álvarez, B., E. G. Biosca, and M. M. López. 2010. On the life of *Ralstonia solanacearum*, a destructive bacterial plant pathogen. In: Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology. Méndez-Vilas, A. (Ed.). Formatex Research Center, Badajoz, Spain. 267-279.

Álvarez, B., M. M. López, and E. G. Biosca. 2008. Survival strategies and pathogenicity of *Ralstonia solanacearum* phylotype II subjected to prolonged starvation in environmental water microcosms. Microbiology. 154:3590-3598.

Arora, N. K., and M. Verma. 2017. Modified microplate method for rapid and efficient estimation of siderophore produced by bacteria. 3 Biotech 7:1–9.

Bahrun, A., C. R. Jensen, F. Asch, and V. O. Mogensen. 2002. Drought-induced changes in xylem pH, ionic composition, and ABA concentration act as early signals in field-grown maize (*Zea mays* L.). Journal of Experimental Botany 53:251–263.

Bajic, D., and A. Sanchez. 2020. The ecology and evolution of microbial metabolic strategies. Current Opinion in Biotechnology 62:123–128.

Bankevich, P. A., S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, and P. A. Pevzner. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology 19:455–477.

Baroukh, C., M. Zemouri, and S. Genin. 2021. Trophic preferences of the pathogen *Ralstonia solanacearum* and consequences on its growth in xylem sap. MicrobiologyOpen 11:1–14.

Bhanwar, R. 2022. Effect of temperature and pH level on growth of bacterial wilt causing *R. solanacearum* under in vitro condition. The Pharma Innovation Journal 11:1656–1659.

Bhatt, G., and T. P. Denny. 2004. *Ralstonia solanacearum* iron scavenging by the siderophore staphyloferrin B is controlled by PhcA, the global virulence regulator. Journal of Bacteriology 186:7896–7904.

Bhunchoth, A., N. Phironrit, C. Leksomboon, O. Chatchawankanphanich, S. Kotera, E. Narulita, T. Kawasaki, M. Fujie, and T. Yamada. 2015. Isolation of *Ralstonia solanacearum*-infecting bacteriophages from tomato fields in Chiang Mai, Thailand, and their experimental use as biocontrol agents. Journal of Applied Microbiology 118(4):1023-1033.

Blackwell, P. A., P. Kay, and A. B. A. Boxall. 2007. The dissipation and transport of veterinary antibiotics in a sandy loam soil. Chemosphere 67:292–299.

Bleuven, C., and C. R. Landry. 2016. Molecular and cellular bases of adaptation to a changing environment in microoragnisms. Proceedings of the Royal Society B 283:20161458.

Blumenstein, K., D. Macaya-Sanz, J. A. Martín, B. R. Albrectsen, and J. Witzell. 2015. Phenotype microarrays as a complementary tool to next generation sequencing for characterization of tree endophytes. Frontiers in Microbiology 6:1033.

Bochner, B. R. 2009. Global phenotypic characterization of bacteria. FEMS Microbiology Reviews. 33:191-205.

Bocsanczy, A. M., A. S. Espindola, and D. J. Norman. 2019. Whole-Genome sequences of *Ralstonia solanacearum*sStrains P816, P822, and P824, emerging pathogens of blueberry in Florida . Microbiology Resource Announcements 8.

Bocsanczy, A. M., J. C. Huguet-Tapia, and D. J. Norman. 2017. Comparative genomics of *Ralstonia solanacearum* identifies candidate genes associated with cool virulence. Frontiers in Plant Science 8.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

Bragard, C., K. Dehnen-Schmutz, F. Di Serio, P. Gonthier, J. A. Jaques Miret, A. F. Justesen, A. MacLeod, C. S. Magnusson, P. Milonas, J. A. Navas-Cortes, S. Parnell, R. Potting, P. L. Reignault, H. Thulke, W. Van der Werf, A. Vicent Civera, J. Yuen, L. Zappalà, J. Van der Wolf, T. Kaluski, M. Pautasso, and M. Jacques. 2019. Pest categorisation of the *Ralstonia solanacearum* species complex. EFSA Journal 17.

Brennan, G., and S. Collins. 2015. Growth responses of a green alga to multiple environmental drivers. Nature Climate Change 5:892–897.

Buchfink, B., K. Reuter, and H.-G. Drost. 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nature Methods 18:366–368.

Bull, J. J. 1987. Evolution of phenotypic variance. Evolution 41:303–315.

Burton, E., N. Yakandawala, K. LoVetri, and M. S. Madhyastha. 2007. A microplate spectrofluorometric assay for bacterial biofilms. Journal of Industrial Microbiology and Biotechnology 34:1–4.

Campos, H., and O. Ortiz, editors. 2020. The potato crop: Its agricultural, nutritional and social contribution to humankind. Springer International Publishing, Cham, Switzerland.

Cantalapiedra, C. P., A. Hernández-Plaza, I. Letunic, P. Bork, and J. Huerta-Cepas. 2021. eggNOG-mapper v2: Functional annotation, orthology assignments,

and domain prediction at the metagenomic scale. Molecular Biology and Evolution 38:5825–5829.

Caruso, P., E. G. Biosca, E. Bertolini, E. Marco-Noales, M. T. Gorris, C. Licciardello, and M. M. López. 2017. Genetic diversity reflects geographical origin of *Ralstonia solanacearum* strains isolated from plant and water sources in Spain. International Microbiology 20:155–164.

Castillo, J. A., and J. T. Greenberg. 2007. Evolutionary dynamics of *Ralstonia solanacearum*. Applied and Environmental Microbiology 73:1225–1238.

Cellier, G., and P. Prior. 2010. Deciphering phenotypic diversity of *Ralstonia solanacearum* strains pathogenic to potato. Phytopathology 11.

Champoiseau, P. G., J. B. Jones, and C. Allen. 2009. *Ralstonia solanacearum* Race 3 Biovar 2 causes tropicall losses and temperate anxieties. Plant Health Progress 10:35.

Chen, M., W. Zhang, L. Han, X. Ru, Y. Cao, Y. Hikichi, K. Ohnishi, G. Pan, and Y. Zhang. 2022. A CysB regulator positively regulates cysteine synthesis, expression of type III secretion system genes, and pathogenicity in *Ralstonia solanacearum*. Molecular Plant Pathology 23:679–692.

Chen, P. E., and B. J. Shapiro. 2015. The advent of genome-wide association studies for bacteria. Current Opinion in Microbiology 25:17–24.

Chen, Y. H., C. W. Lu, Y. T. Shyu, and S. S. Lin. 2017. Revealing the saline adaptation strategies of the halophilic bacterium *Halomonas beimenensis* through high-

throughput omics and transposon mutagenesis approaches. Scientific Reports 7:1–15.

Cho, H., E. S. Song, Y. K. Lee, S. Lee, S. W. Lee, A. Jo, B. M. Lee, J. G. Kim, and I. Hwang. 2018. Analysis of genetic and pathogenic diversity of *Ralstonia solanacearum* causing potato bacterial wilt in Korea. Plant Pathology Journal 34:23–34.

Cho, H., E.-S. Song, S. Heu, J. Baek, Y. K. Lee, S. Lee, S.-W. Lee, D. S. Park, T.-H. Lee, J.-G. Kim, and I. Hwang. 2019. Prediction of host-specific genes by pan-genome analyses of the Korean *Ralstonia solanacearum* species complex. Frontiers in Microbiology 10.

Choi, S. Y., N. G. Kim, S.-M. Kim, and B. C. Lee. 2022. First report of bacterial wilt by *Ralstonia pseudosolanacearum* on peanut in Korea. Research in Plant Disease 28:54–56.

Cianciotto, N. P. 2005. Type II secretion: a protein secretion system for all seasons. Trends in Microbiology 13:581–588.

Cianciotto, N. P., and R. C. White. 2017. Expanding role of type II secretion in bacterial pathogenesis and beyond. Infection and Immunity 85:e00014-17.

Cingolani, P., A. Platts, L. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, and D. M. Ruden. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w ^{1118} ; iso-2; iso-3. Fly 6:80–92.

Coenye, T., and P. Vandamme. 2003. Simple sequence repeats and compositional bias in the bipartite *Ralstonia solanacearum* GMI1000 genome. BMC Genomics 4:10.

Collins, C., and X. Didelot. 2018. A phylogenetic method to perform genome-wide association studies in microbes that accounts for population structure and recombination. PLoS Computational Biology.

Cordero, O. X., and M. F. Polz. 2014. Explaining microbial genomic diversity in light of evolutionary ecology. Nature Reviews Microbiology 12:263–273.

Croll, D., and B. A. McDonald. 2017. The genetic basis of local adaptation for pathogenic fungi in agricultural ecosystems. Molecular Ecology 26:2027–2040.

Cruz, L., M. Eloy, F. Quirino, H. Oliveira, and R. Tenreiro. 2012. Molecular epidemiology of *Ralstonia solanacearum* strains from plants and environmental sources in Portugal. European Journal of Plant Pathology 133:687–706.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, and 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158.

Deslandes, L., and S. Rivas. 2012. Catch me if you can: Bacterial effectors and plant targets. Trends in Plant Science 17:644–655.

Dimitriu, T. 2022. Evolution of horizontal transmission in antimicrobial resistance plasmids. Microbiology 168.

Doan, T. K. T., C. T. Lu, V. L. Pham, and T. T. N. Nguyen. 2022. Efficacy of bacteriophages in controlling bacterial vascular wilt caused by *Ralstonia solanacearum* Smith on eggplants. Can Tho University Journal of Science 14:81–85.

Drenkard, E., and F. M. Ausubel. 2002. Competing interests statement Pseudomonas biofilm formation and antibiotic resistance are linked to phenotypic variation.

Ebbert, D. 2019. chisq.posthoc.test: A post hoc analysis for Pearson's chi-squared test for count data.

Ekstrøm, C. T. 2019. MESS: Miscellaneous Esoteric Statistical Scripts.

Elabed, H., E. González-Tortuero, C. Ibacache-Quiroga, A. Bakhrouf, P. Johnston, K. Gaddour, J. Blázquez, and A. Rodríguez-Rojas. 2019. Seawater salt-trapped *Pseudomonas aeruginosa* survives for years and gets primed for salinity tolerance. BMC Microbiology 19:1–13.

Elena, S. F., and R. E. Lenski. 2003. Evolution experiments with microorganisms: The dynamics and genetic bases of adaptation. Nature Reviews Genetics 4:457–469.

Elphinstone, J. G., J. Hennessy, J. K. Wilson, and D. E. Stead. 1996. Sensitivity of different methods for the detection of *Ralstonia solanacearum* in potato tuber extracts. Page Bulletin OEPP/EPPO Bulletin.

Elphinstone, J., and S. Matthews-Berry. 2017. Monitoring and control of the potato brown rot bacterium in irrigation water. DEFRA.

Elphinstone, J., H. M. Stanford, and D. E. Stead. 1997. Detection of *Ralstonia solanacearum* in potato tubers, *Solanum dulcamara* and associated irrigation water. Pages 133–139 *in* P. Prior, C. Allen, and J. Elphinstone, editors. Bacterial Wilt Disease. Springer Verlag.

EPPO. 2018. PM 7/21 (2) *Ralstonia solanacearum*, *R. pseudosolanacearum* and *R. syzygii* (*Ralstonia solanacearum* species complex). Bulletin OEPP/EPPO Bulletin.

EPPO. 2022. EPPO global database. https://gd.eppo.int/.

Erill, I., M. Puigvert, L. Legrand, R. Guarischi-Sousa, C. Vandecasteele, J. C. Setubal, S. Genin, A. Guidot, and M. Valls. 2017. Comparative analysis of *Ralstonia solanacearum* methylomes. Frontiers in Plant Science 8:1–16.

Fegan, M., and P. Prior. 2005. How complex is the "*Ralstonia solanacearum* Species Complex." In: Bacterial wilt Disease and the *Ralstonia solanacearum* Species Complex. Allen, C., Priot, P., Hayward, A.C., (Ed.). APS Press, St. Paul, MN, USA. 449-461.

Fegan, M., M. Taghavi, L. I. Sly, and A. C. Hayward. 1998. Phylogeny, diversity and molecular diagnostics of *Ralstonia solanacearum*. Bacterial Wilt Disease. In: Bacterial Wilt Disease. Prior, P., Allen, C., Elphinstone, J. (Eds.). Springer, Berlin, Heidelberg.

Ferenci, T. 2016. Trade-off Mechanisms Shaping the Diversity of Bacteria. Trends in Microbiology 24:209–223.

Flores-Cruz, Z., and C. Allen. 2009. *Ralstonia solanacearum* encounters an oxidative environment during tomato infection. Molecular Plant-Microbe Interactions 22:773–782.

Foster, K. R., G. Shaulsky, J. E. Strassmann, D. C. Queller, and C. R. L. Thompson. 2004. Pleiotropy as a mechanism to stabilize cooperation. Nature 431:693–696.

de la Fuente-Núñez, C., F. Reffuveille, L. Fernández, and R. E. Hancock. 2013. Bacterial biofilm development as a multicellular adaptation: antibiotic resistance and new therapeutic strategies. Current Opinion in Microbiology 16:580–589.

Fukui, R. 2003. Suppression of soilborne plant pathogens through community evolution of soil microorganisms. Microbes and Environments 18:1–9.

van der Gaag, D. J., M. Camilleri, M. Diakaki, M. Schenk, and S. Vos. 2019. Pest survey card on potato brown rot, *Ralstonia solanacearum*. EFSA Supporting Publications 16.

Gao, C.-H. 2021. ggVennDiagram: A “ggplot2” Implement of Venn Diagram.

Garćia, R. O., J. P. Kerns, and L. Thiessen. 2019. *Ralstonia solanacearum* species complex: A quick diagnostic guide. Plant Health Progress 20:7–13.

Garrison, E., and G. Marth. 2012. Haplotype-based variant detection from short-read sequencing.

Geng, R., L. Cheng, C. Cao, Z. Liu, D. Liu, Z. Xiao, X. Wu, Z. Huang, Q. Feng, C. Luo, Z. Chen, Z. Zhang, C. Jiang, M. Ren, and A. Yang. 2022. Comprehensive analysis reveals the genetic and pathogenic diversity of *Ralstonia solanacearum* species complex and benefits Its taxonomic classification. Frontiers in Microbiology. 13.

Genin, S. 2010. Molecular traits controlling host range and adaptation to plants in Ralstonia solanacearum. New Phytologist 187:920-928.

Genin, S., and C. Boucher. 2002. *Ralstonia solanacearum*: Secrets of a major pathogen unveiled by analysis of its genome. Molecular Plant Pathology 3(3):111-118.

Genin, S., and C. Boucher. 2004. Lessons learned from the genome analysis of *Ralstonia solanacearum*. Annual Review of Phytopathology 42:107–134.

Genin, S., and T. P. Denny. 2012. Pathogenomics of the *Ralstonia solanacearum* Species Complex. Annual Review of Phytopathology 50:67-89.

Gibbons, S. M., and J. A. Gilbert. 2015. Microbial diversity — exploration of natural ecosystems and microbiomes. Current Opinion in Genetics & Development 35:66–72.

Gonçalves, O. S., K. F. Campos, J. C. S. de Assis, A. S. Fernandes, T. S. Souza, L. G. D. C. Rodrigues, M. V. de Queiroz, and M. F. Santana. 2020. Transposable elements contribute to the genome plasticity of *Ralstonia solanacearum* species complex. Microbial Genomics 6:1–12.

Gonçalves, O. S., F. D. O. Souza, F. P. Bruckner, M. F. Santana, and P. Alfenas-zerbini. 2021. Widespread distribution of prophages signaling the potential for adaptability and pathogenicity evolution of *Ralstonia solanacearum* species complex. Genomics 113:992–1000.

Greenrod, S. T., M. Stoycheva, J. Elphinstone, and V.-P. Friman. 2022. Influence of insertion sequences on population structure of phytopathogenic bacteria in the *Ralstonia solanacearum* species complex. bioRxiv.

Gregory, G. J., and E. F. Boyd. 2021. Stressed out: Bacterial response to high salinity using compatible solute biosynthesis and uptake systems, lessons from *Vibrionaceae*. Computational and Structural Biotechnology Journal 19:1014–1027.

Gu, S., Z. Wei, Z. Shao, V. P. Friman, K. Cao, T. Yang, J. Kramer, X. Wang, M. Li, X. Mei, Y. Xu, Q. Shen, R. Kümmerli, and A. Jousset. 2020. Competition for iron drives phytopathogen control by natural rhizosphere microbiomes. Nature Microbiology 5:1002-1010.

Gu, Z., R. Eils, and M. Schlesner. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32:2847–2849.

Guidot, A., B. Coupat, S. Fall, P. Prior, and F. Bertolla. 2009. Horizontal gene transfer between *Ralstonia solanacearum* strains detected by comparative genomic hybridization on microarrays. ISME Journal 3:549–562.

Hadfield, J., N. J. Croucher, R. J. Goater, K. Abudahab, D. M. Aanensen, and S. R. Harris. 2018. Phandango: an interactive viewer for bacterial population genomics. Bioinformatics 34:292–293.

Hanson, C. A., J. A. Fuhrman, M. C. Horner-Devine, and J. B. H. Martiny. 2012. Beyond biogeographic patterns: Processes shaping the microbial landscape. Nature Reviews Microbiology 10:497–506.

Harrell Jr, F. E. 2022. Hmisc: Harrell Miscellaneous. R.

Hawkey, J., M. Hamidian, R. R. Wick, D. J. Edwards, H. Billman-Jacobe, R. M. Hall, and K. E. Holt. 2015. ISMapper: identifying transposase insertion sites in bacterial genomes from short read sequence data. BMC Genomics 16:667.

Hayward, A. C. 1991. Biology and Epidemiology of Bacterial Wilt Caused by *Pseudomonas Solanacearum*. Annual Review of Phytopathology 29:65–87.

Hervé, M. 2022. RVAideMemoire: Testing and plotting procedures for biostatistics.

Hider, R. C., and X. Kong. 2010. Chemistry and biology of siderophores. Natural Product Reports 27:637–657.

Hijmans, R. J. 2003. The effect of climate change on global potato production. American Journal of Potato Research 80:271-279.

Høiby, N., T. Bjarnsholt, M. Givskov, S. Molin, and O. Ciofu. 2010. Antibiotic resistance of bacterial biofilms. International Journal of Antimicrobial Agents 35:322–332.

Horner-Devine, M. C., K. M. Carney, and B. J. M. Bohannan. 2004. An ecological perspective on bacterial biodiversity. Proceedings of the Royal Society B: Biological Sciences 271:113–122.

Hosseinidoust, Z., N. Tufenkji, and T. G. M. van de Ven. 2013. Formation of biofilms under phage predation: considerations concerning a biofilm increase. Biofouling 29:457–468.

Houle, D., D. R. Govindaraju, and S. Omholt. 2010, December. Phenomics: The next challenge. Nature Reviews Genetics 11(12):855-866

Huerta-Cepas, J., D. Szklarczyk, D. Heller, A. Hernández-Plaza, S. K. Forslund, H. Cook, D. R. Mende, I. Letunic, T. Rattei, L. J. Jensen, C. von Mering, and P. Bork. 2019. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Research 47:309–314.

Jacobs, J. M., L. Babujee, F. Meng, A. Milling, and C. Allen. 2012. The in planta transcriptome of *Ralstonia solanacearum*: Conserved physiological and virulence strategies during bacterial wilt of tomato. mBio 3.

Jaillard, M., L. Lima, M. Tournoud, P. Mahé, A. van Belkum, V. Lacroix, and L. Jacob. 2018. A fast and agnostic method for bacterial genome-wide association studies: Bridging the gap between k-mers and genetic events. PLoS genetics 14(11):e1007758.

Janse, J. D. 1996. Potato brown rot in western Europe – history, present occurrence and some remarks on possible origin, epidemiology and control strategies. EPPO Bulletin 26:679–695.

Jeong, E. L., and J. N. Timmis. 2000. Novel insertion sequence elements associated with genetic heterogeneity and phenotype conversion in *Ralstonia solanacearum*. Journal of Bacteriology 182:4673–4676.

Jezbera, J., J. Jezberová, U. Brandt, and M. W. Hahn. 2011. Ubiquity of *Polynucleobacter necessarius* subspecies asymbioticus results from ecological diversification. Environmental Microbiology 13:922–931.

Jezbera, J., J. Jezberová, V. Kasalický, K. Šimek, and M. W. Hahn. 2013. Patterns of *Limnohabitans* microdiversity across a large set of freshwater habitats as revealed by reverse line blot hybridization. PLoS ONE 8(3):e58527.

Jiang, G., Z. Wei, J. Xu, H. Chen, Y. Zhang, X. She, A. P. Macho, W. Ding, and B. Liao. 2017. Bacterial wilt in China: History, current status, and future perspectives. Frontiers in Plant Science 8:1–10.

Jiang, G., N. Wang, Y. Zhang, Z. Wang, Y. Zhang, J. Yu, Y. Zhang, Z. Wei, Y. Xu, S. Geisen, V-P, Friman, and Q. Shen. 2021. The relative importance of soil moisture in predicting bacterial wilt disease occurrence. Soil Ecology Letters. 3:356-366.

Jiang, S., X. Wu, S. Du, Q. Wang, and D. Han. 2022. Are UK rivers getting saltier and more alkaline? Water 14(18):2813.

Jousset, A., E. Lara, L. G. Wall, and C. Valverde. 2006. Secondary metabolites help biocontrol strain *Pseudomonas fluorescen*s CHA0 to escape protozoan grazing. Applied and Environmental Microbiology 72:7083–7090.

Kang, Y., H. Liu, S. Genin, M. A. Schell, and T. P. Denny. 2002. *Ralstonia solanacearum* requires type 4 pili to adhere to multiple surfaces and for natural transformation and virulence: *R. solanacearum* type 4 pili. Molecular Microbiology 46:427–437.

Kelman, A. 1954. The relationship of pathogenicity in *Pseudomonas solanacearum* to colony appearance on a tetrazolium medium. Phytopathology 44:693–695.

Kent, A. G., C. L. Dupont, S. Yooseph, and A. C. Martiny. 2016. Global biogeography of *Prochlorococcus* genome diversity in the surface ocean. The ISME Journal 10:1856–1865.

Khokhani, D., T. M. Lowe-Power, T. M. Tran, and C. Allen. 2017. A single regulator mediates strategic switching between attachment/spread and growth/virulence in the plant pathogen *Ralstonia solanacearum*. mBio 8(5).

Kleter, G. A., M. J. Groot, M. Poelman, E. J. Kok, and H. J. P. Marvin. 2009, May. Timely awareness and prevention of emerging chemical and biochemical risks in foods: Proposal for a strategy based on experience with recent cases. Food and Chemical Toxicology 47(5):992-1008.

Knaus, B. J., and N. J. Grünwald. 2017. VCFR : a package to manipulate and visualize variant call format data in R. Molecular Ecology Resources 17:44–53.

Krassowski, M., M. Arts, C. Lagger, and Max. 2020. krassowski/complex-upset: v1.3.5. R, Zenodo.

Kraupner, N., S. Ebmeyer, J. Bengtsson-Palme, J. Fick, E. Kristiansson, C.-F. Flach, and D. G. J. Larsson. 2018. Selective concentration for ciprofloxacin resistance in *Escherichia coli* grown in complex aquatic bacterial biofilms. Environment International 116:255–268.

Kuhn, M., and H. Wickham. 2020. Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles. R.

Kumar, D., M. Saleem Dar, K. Khan, D. Kumar Choudhary, R. Srinagar, I. Mohammad Saleem Dar, S. Shalimar Srinagar, I. Kamran Ahmad Khan, S. Foa, W. Sopore, S. U. Nabi, and K. Ahmad Khan. 2018. *Ralstonia solanacearum*: A wide spread and global bacterial plant wilt pathogen. Journal of Pharmacognosy and Phytochemistry 7.

Kunwar, S., F. Iriarte, Q. Fan, E. Evaristo da Silva, L. Ritchie, N. S. Nguyen, J. H. Freeman, R. E. Stall, J. B. Jones, G. V. Minsavage, J. Colee, J. W. Scott, G. E. Vallad, C. Zipfel, D. Horvath, J. Westwood, S. F. Hutton, and M. L. Paret. 2018. Transgenic expression of *EFR* and *Bs2* genes for field management of bacterial wilt and bacterial spot of tomato. Phytopathology 108:1402–1411.

Landry, D., M. González-Fuente, L. Deslandes, and N. Peeters. 2020. The large, diverse, and robust arsenal of *Ralstonia solanacearum* type III effectors and their in planta functions. Molecular Plant Pathology 21:1377–1388.

Lannou, C. 2012. Variation and selection of quantitative traits in plant pathogens. Annual Review of Phytopathology 50:319–338.

Larkin, A. A., and A. C. Martiny. 2017. Microdiversity shapes the traits, niche space, and biogeography of microbial taxa. Environmental Microbiology Reports 9:55–70.

Laxminarayan, R., T. Van Boeckel, I. Frost, S. Kariuki, E. A. Khan, D. Limmathurotsakul, D. G. J. Larsson, G. Levy-Hara, M. Mendelson, K. Outterson, S. J. Peacock, and Y.-G. Zhu. 2020. The *Lancet Infectious Diseases* Commission on antimicrobial resistance: 6 years later. The Lancet Infectious Diseases 20:51–60.

Lebeau, A., M. C. Daunay, A. Frary, A. Palloix, J. F. Wang, J. Dintinger, F. Chiroleu, E. Wicker, and P. Prior. 2011. Bacterial wilt resistance in tomato, pepper, and eggplant: Genetic resources respond to diverse strains in the *Ralstonia solanacearum* species complex. Phytopathology 101:154–165.

Lees, J. A., M. Galardini, S. D. Bentley, J. N. Weiser, and J. Corander. 2018. pyseer: A comprehensive tool for microbial pangenome-wide association studies. Bioinformatics 34:4310–4312.

Lenth, R. V. 2021. emmeans: Estimated marginal means, aka least-squares means.

Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25(14):1754-1760.

Li, M., Z. Wei, J. Wang, A. Jousset, V. P. Friman, Y. Xu, Q. Shen, and T. Pommier. 2019. Facilitation promotes invasions in plant-associated microbial communities. Ecology Letters 22:149–158.

Li, P., M. Liu, G. Li, K. Liu, T. Liu, M. Wu, M. Saleem, and Z. Li. 2021. Phosphorus availability increases pathobiome abundance and invasion of rhizosphere microbial networks by *Ralstonia*. Environmental Microbiology 23:5992–6003.

Li, S., Y. Liu, J. Wang, L. Yang, S. Zhang, C. Xu, and W. Ding. 2017. Soil acidification aggravates the occurrence of bacterial wilt in south China. Frontiers in Microbiology 8:1–12.

Liao, J., X. Guo, D. L. Weller, S. Pollak, D. H. Buckley, M. Wiedmann, and O. X. Cordero. 2021. Nationwide genomic atlas of soil-dwelling *Listeria* reveals effects of selection and population ecology on pangenome evolution. Nature Microbiology 6:1021–1030.

Liu, Y., X. Tan, Y. Pan, J. Yu, Y. Du, X. Liu, and W. Ding. 2022. Mutation in phcA enhanced the adaptation of *Ralstonia solanacearum* to long-term acid stress. Frontiers in Microbiology 13.

Lowe-Power, T. M., D. Khokhani, and C. Allen. 2018. How *Ralstonia solanacearum* exploits and thrives in the flowing plant xylem environment. Trends in Microbiology 26(11)929-942.

Lundström, S. V., M. Östman, J. Bengtsson-Palme, C. Rutgersson, M. Thoudal, T. Sircar, H. Blanck, K. M. Eriksson, M. Tysklind, C.-F. Flach, and D. G. J. Larsson. 2016. Minimal selective concentrations of tetracycline in complex aquatic bacterial biofilms. Science of The Total Environment 553:587–595.

Mansfield, J., S. Genin, S. Magori, V. Citovsky, M. Sriariyanum, P. Ronald, M. Dow, V. Verdier, S. V Beer, M. A. Machado, I. Toth, G. Salmond, and G. D. Foster.

2012. Top 10 plant pathogenic bacteria in molecular plant pathology. Molecular Plant Pathology 13(6):614-629.

Manzoni, S., J. P. Schimel, and A. Porporato. 2012. Responses of soil microbial communities to water stress: Results from a meta-analysis. Ecology 93:930–938.

McArthur, J. V., D. A. Kovacic, and M. H. Smith. 1988. Genetic diversity in natural populations of a soil bacterium across a landscape gradient. Proceedings of the National Academy of Sciences 85:9621–9624.

McGee, L. W., A. M. Sackman, A. J. Morrison, J. Pierce, J. Anisman, and D. R. Rokyta. 2016. Synergistic pleiotropy overrides the costs of complexity in viral adaptation. Genetics 202:285–295.

Mehan, V. K., and D. McDonald. 1995. Techniques for diagnosis of *Pseudomonas solanacearum*, and for resistance screening against ground- nut bacterial wilt. Page Technical Manual no. 1. (In En. Abstracts in En, Ch, Fr.) Patancheru 502 324, Andhra Pradesh, India: International Crops Research Institute for the Semi-Arid Tropics.

Meng, F. 2013. The virulence factors of the bacterial wilt pathogen *Ralstonia solanacearum*. Journal of Plant Pathology & Microbiology 4(3):1000168.

Meng, F., L. Babujee, J. M. Jacobs, and C. Allen. 2015. Comparative transcriptome analysis reveals cool virulence factors of *Ralstonia solanacearum* race 3 biovar 2. PLoS ONE 10(10):e0139090.

Met Office. 2022. Oxford (Oxfordshire) UK climate averages. https://www.metoffice.gov.uk/research/climate/maps-and-data/uk-climate-averages/gcpn7mp10.

Milling, A., F. Meng, T. P. Denny, and C. Allen. 2009. Interactions with hosts at cool temperatures, not cold tolerance, explain the unique epidemiology of R*alstonia solanacearum* Race 3 Biovar 2. Phytopathology 99(10):1127-1134.

National Center for Biotechnology Information (NCBI). 2022. https://www.ncbi.nlm.nih.gov/genome/490.

Nguyen, L.-T., H. A. Schmidt, A. von Haeseler, and B. Q. Minh. 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molecular Biology and Evolution 32:268–274.

Oerke, E. 2006. Crop losses to pests. The Journal of Agricultural Science 144:31–43.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. Henry, E. S. H. Stevens, and H. Wagner. 2019. vegan: Community Ecology Package. R package version 2.5-6.

Olson, E. R. 1993. Influence of pH on bacterial gene expression. Molecular Microbiology 8:5–14.

Ostrowski, E. A., D. E. Rozen, and R. E. Lenski. 2005. Pleiotopic effects of beneficial mutations in *Escherichia coli*. Evolution 59:2343–2352.

O'Toole, G. A., and R. Kolter. 1998. Initiation of biofilm formation in *Pseudomonas fluorescens* WCS365 proceeds via multiple, convergent signalling pathways: A genetic analysis. Molecular Microbiology 28:449–461.

Parkinson, N., R. Bryant, J. Bew, C. Conyers, R. Stones, M. Alcock, and J. Elphinstone. 2013. Application of variable-number tandem-repeat typing to discriminate *Ralstonia solanacearum* strains associated with English watercourses and disease outbreaks. Applied and Environmental Microbiology 79(19):6016-6022.

Paudel, S., S. Dobhal, A. M. Alvarez, and M. Arif. 2020. Taxonomy and phylogenetic research on *Ralstonia solanacearum*: a complex pathogen with extraordinary economic consequences. Pathogens 9(11):886.

Peeters, N., A. Guidot, F. Vailleau, and M. Valls. 2013. *Ralstonia solanacearum*, a widespread bacterial plant pathogen in the post-genomic era. Molecular Plant Pathology 14:651–662.

Peng, S., J. Huang, J. E. Sheehy, R. C. Laza, R. M. Visperas, X. Zhong, G. S. Centeno, G. S. Khush, and K. G. Cassman. 2004. Rice yields decline with higher night temperature from global warming. Proceedings of the National Academy of Sciences 101:9971–9975.

Perrier, A., X. Barlet, D. Rengel, P. Prior, S. Poussier, S. Genin, and A. Guidot. 2019. Spontaneous mutations in a regulatory gene induce phenotypic heterogeneity and adaptation of *Ralstonia solanacearum* to changing environments. Environmental Microbiology 21:3140–3152.

Perrier, A., R. Peyraud, D. Rengel, X. Barlet, E. Lucasson, J. Gouzy, N. Peeters, S. Genin, and A. Guidot. 2016. Enhanced *in planta* fitness through adaptive mutations in EfpR, a dual regulator of virulence and metabolic functions in the plant pathogen *Ralstonia solanacearum*. PLoS Pathogens 12(2):e1006044.

Peyraud, R., L. Cottret, L. Marmiesse, J. Gouzy, and S. Genin. 2016. A resource allocation trade-off between virulence and proliferation drives metabolic versatility in the plant pathogen *Ralstonia solanacearum*. PLoS Pathogens 12(10):e1005939.

Phillips, M. A., and M. K. Burke. 2021. Can laboratory evolution experiments teach us about natural populations? Molecular Ecology 30:877–879.

Power, R. A., J. Parkhill, and T. De Oliveira. 2016. Microbial genome-wide association studies: lessons from human GWAS. Nature Reviews Genetics 18:41-50.

Prior, P., F. Ailloud, B. L. Dalsing, B. Remenant, B. Sanchez, and C. Allen. 2016. Genomic and proteomic evidence supporting the division of the plant pathogen *Ralstonia solanacearum* into three species. BMC Genomics 17:90.

Quivey, R. G., W. L. Kuhnert, and K. Hahn. 2000. Adaptation of oral *streptococci* to low pH. Advances in Microbial Physiology 42:239–274.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

R Core Team. 2022. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rainey, P. B., H. J. E. Beaumont, G. C. Ferguson, J. Gallie, C. Kost, E. Libby, and X. X. Zhang. 2011. The evolutionary emergence of stochastic phenotype switching in bacteria. Microbial Cell Factories 10.

Rainey, P. B., and M. Travisano. 1998. Adaptive radiation in a heterogeneous environment. Nature 394:69–72.

Read, T. D., and R. C. Massey. 2014. Characterizing the genetic basis of bacterial phenotypes using genome-wide association studies: A new direction for bacteriology. Genome Medicine 6:109.

Remenant, B., J. C. de Cambiaire, G. Cellier, J. M. Jacobs, S. Mangenot, V. Barbe, A. Lajus, D. Vallenet, C. Medigue, M. Fegan, C. Allen, and P. Prior. 2011. *Ralstonia syzygii*, the blood disease bacterium and some Asian *R. solanacearum* strains form a single genomic species despite divergent lifestyles. PLoS ONE 6(9)e24356.

Remenant, B., B. Coupat-Goutaland, A. Guidot, G. Cellier, E. Wicker, C. Allen, M. Fegan, O. Pruvost, M. Elbaz, A. Calteau, G. Salvignol, D. Mornico, S. Mangenot, V. Barbe, C. Médigue, and P. Prior. 2010. Genomes of three tomato pathogens within the *Ralstonia solanacearum* species complex reveal significant evolutionary divergence. BMC Genomics 11:379.

Retana-Cordero, M., P. R. Fisher, and C. Gómez. 2021. Modeling the effect of temperature on ginger and turmeric rhizome sprouting. Agronomy 11(10):1931.

Robinson, J. T., H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, and J. P. Mesirov. 2011. Integrative genomics viewer. Nature Biotechnology 29:24–26.

Rowbury, R. J. 1997. Regulatory components, including integration host factor, CysB and H-NS, that influence pH responses in *Escherichia coli*. Letters in Applied Microbiology 24:319–328.

Safni, I., I. Cleenwerck, P. De Vos, M. Fegan, L. Sly, and U. Kappler. 2014. Polyphasic taxonomic revision of the *Ralstonia solanacearum* species complex: proposal to emend the descriptions of *R. solanacearum* and *R. syzygii* and reclassify current *R. syzygii* strains as *Ralstonia syzygii* subsp. *syzygii*, *R. solanacearum* phylotype IV s. International Journal of Systematic and Evolutionary Microbiology 64:3087–3103.

Safni, I., S. Subandiyah, and M. Fegan. 2018, March 1. Ecology, epidemiology and disease management of *Ralstonia syzygii* in Indonesia. Frontiers Media S.A.

Saito, H., and H. Kobayashi. 2003. Bacterial responses to alkaline stress. Science Progress 86:271–282.

Salanoubat, M., S. Genin, F. Artiguenave, J. Gouzy, S. Mangenot, M. Arlat, A. Billault, P. Brottiert, J. C. Camus, L. Cattolico, M. Chandler, N. Choisne, C. Claudel-Renard, S. Cunnac, N. Demange, C. Gaspin, M. Lavie, A. Moisan, C. Robert, W. Saurin, T. Schiex, P. Siguler, P. Thébault, M. Whalen, P. Wincker, M. Levy, J. Weissenbach, and C. A. Boucher. 2002. Genome sequence of the plant pathogen *Ralstonia solanacearum*. Nature 415:497-502.

Saltz, J. B., F. C. Hessel, and M. W. Kelly. 2017. Trait correlations in the genomics era. Trends in Ecology and Evolution 32:279–290.

Sardinha, M., T. Müller, H. Schmeisky, and R. G. Joergensen. 2003. Microbial performance in soils along a salinity gradient under acidic conditions. Applied Soil Ecology 23:237–244.

Sarmah, A. K., M. T. Meyer, and A. B. A. Boxall. 2006. A global perspective on the use, sales, exposure pathways, occurrence, fate and effects of veterinary antibiotics (VAs) in the environment. Chemosphere 65:725–759.

Schell, M. A. 2000. Control of virulence and pathogenicity genes of *Ralstonia solanacearum* by an elaborate sensory network. Annual Review of Phytopathology 38:263–292.

Schenk, M. F., S. Witte, M. L. M. Salverda, B. Koopmanschap, J. Krug, and J. A. G. M. de Visser. 2015. Role of pleiotropy during adaptation of TEM-1 β-lactamase to two novel antibiotics. Evolutionary Applications 8:248–260.

Schloerke, B., D. Cook, J. Larmarange, F. Briatte, M. Marbach, E. Thoen, A. Elberg, and J. Crowley. 2021. GGally: Extension to “ggplot2.”

Schwyn, B., and J. B. Neilands. 1987. Universal chemical assay for the detection and determination of siderophores. Analytical Biochemistry 160:47–56.

Secchi, F., and M. A. Zwieniecki. 2016. Accumulation of sugars in the xylem apoplast observed under water stress conditions is controlled by xylem pH. Plant Cell and Environment 39:2350–2360.

Seemann, T. 2014. Prokka: rapid prokaryotic genome annotation. 30:2068–2069.

Seemann, T. 2015. snippy: fast bacterial variant calling from NGS reads.

Sgrò, C. M., and A. A. Hoffmann. 2004. Genetic correlations, tradeoffs and environmental variation. Heredity 93:241–248.

Sharma, P., M. Johnson, R. Mazloom, C. Allen, L. Heath, T. Lowe-Power, and B. Vinatzer. 2022. Meta Analysis of the *Ralstonia solanacearum* species complex (RSSC) based on comparative evolutionary genomics and reverse ecology. bioRxiv.

Sheppard, S. K., X. Didelot, G. Meric, A. Torralbo, K. A. Jolley, D. J. Kelly, S. D. Bentley, M. C. J. Maiden, J. Parkhill, and D. Falush. 2013. Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in *Campylobacter*. Proceedings of the National Academy of Sciences 110:11923–11927.

Shrivastava, P., and R. Kumar. 2015. Soil salinity: A serious environmental issue and plant growth promoting bacteria as one of the tools for its alleviation. Saudi Journal of Biological Sciences 22:123–131.

Sibinelli-Sousa, S., J. T. Hespanhol, G. G. Nicastro, B. Y. Matsuyama, S. Mesnage, A. Patel, R. F. de Souza, C. R. Guzzo, E. Bayer-Santos. 2020. A family of T6SS antibacterial effectors related to l,d-transpeptidases targets the peptidoglycan. Cell Reports 31(12):107813.

Siliakus, M. F., J. van der Oost, and S. W. M. Kengen. 2017. Adaptations of archaeal and bacterial membranes to variations in temperature, pH and pressure. Extremophiles 21:651–670.

Silver, L. L. 2016. Appropriate targets for antibacterial drugs. Cold Spring Harbor Perspectives in Medicine 6:1–7.

Singer, A. C., J. D. Järhult, R. Grabic, G. A. Khan, R. H. Lindberg, G. Fedorova, J. Fick, M. J. Bowes, B. Olsen, and H. Söderström. 2014. Intra- and Inter-pandemic variations of antiviral, antibiotics and decongestants in wastewater treatment plants and receiving rivers. PLoS ONE 9:e108621.

Singh, D., and D. Kumar Yadav. 2016. Potential of *Bacillus amyloliquefaciens* for biocontrol of bacterial wilt of tomato incited by *Ralstonia solanacearum*. Journal of Plant Pathology & Microbiology 7(1):1000327.

Smith, E. F. 1896. A bacterial disease of the tomato, eggplant, and irish potato. Bulletin of the U.S. Department of Agriculture:1–27.

Smits, W. K., O. P. Kuipers, and J. W. Veening. 2006. Phenotypic variation in bacteria: The role of feedback regulation. Nature Reviews Microbiology 4:259–271.

Sniegowski, P., P. Gerrish, and R. E. Lenski. 1997. Evolution of high mutation rates in experimental populations of E. coli. Nature 387:703–705.

Söderberg, M. A., O. Rossier, and N. P. Cianciotto. 2004. The type II protein secretion system of *Legionella pneumophila* promotes growth at low temperatures. Journal of Bacteriology 186:3712–3720.

Stanton, I. C., A. K. Murray, L. Zhang, J. Snape, and W. H. Gaze. 2020. Evolution of antibiotic resistance at low antibiotic concentrations including selection below the minimal selective concentration. Communications Biology 3:467.

Stearns, S. C. 1989. Trade-offs in life-history evolution. Function Ecology 3:259–268.

Stevens, L. H., P. S. van der Zouwen, C. A. M. van Tongeren, P. Kastelein, and J. M. van der Wolf. 2018. Survival of *Ralstonia solanacearum* and *Ralstonia pseudosolanacearum* in drain water. EPPO Bulletin.

Stevens, P., and J. D. Van Elsas. 2010. Genetic and phenotypic diversity of *Ralstonia solanacearum* biovar 2 strains obtained from Dutch waterways. Antonie van Leeuwenhoek 97(2):171-188.

Taghavi, M., C. Hayward, L. I. Sly, and M. Fegan. 1996. Analysis of the phylogenetic relationships of strains of *Burkholderia solanacearum*, *Pseudomonas syzygii*, and the blood disease bacterium of banana based on 16S rRNA gene sequences. International Journal of Systematic Bacteriology 46:10–15.

Timilsina, S., N. Potnis, E. A. Newberry, P. Liyanapathiranage, F. Iruegas-Bocardo, F. F. White, E. M. Goss, and J. B. Jones. 2020. *Xanthomonas* diversity, virulence and plant–pathogen interactions. Nature Reviews Microbiology 18:415–427.

Timms-Wilson, T. M., K. Bryant, and M. J. Bailey. 2001. Strain characterization and 16S-23S probe development for differentiating geographically dispersed isolates of the phytopathogen *Ralstonia solanacearum*. Environmental Microbiology 3(2):785-797.

Tonkin-Hill, G., N. MacAlasdair, C. Ruis, A. Weimann, G. Horesh, J. A. Lees, R. A. Gladstone, S. Lo, C. Beaudoin, R. A. Floto, S. D. W. Frost, J. Corander, S. D. Bentley, and J. Parkhill. 2020. Producing polished prokaryotic pangenomes with the Panaroo pipeline. Genome Biology 21:180.

Topp, E., S. C. Monteiro, A. Beck, B. B. Coelho, A. B. A. Boxall, P. W. Duenk, S. Kleywegt, D. R. Lapen, M. Payne, L. Sabourin, H. Li, and C. D. Metcalfe. 2008. Runoff of pharmaceuticals and personal care products following application of biosolids to an agricultural field. Science of The Total Environment 396:52–59.

United Nations. 2015. World population prospects the 2017 revision. Department of Economic and Social Affairs, Population Division.

USDA. 2020. Animal and Plant Health Inspection Service. https://www.aphis.usda.gov/aphis/ourfocus/planthealth/plant-pest-and-disease-programs/pests-and-diseases/plant-disease/sa_ralstonia/ct_ralstonia.

Venkatesh, B., L. Babujee, H. Liu, P. Hedley, T. Fujikawa, P. Birch, I. Toth, and S. Tsuyumu. 2006. The *Erwinia chrysanthemi* 3937 PhoQ sensor kinase regulates several virulence determinants. Journal of Bacteriology 188:3088–3098.

Wang, H. C., H. Guo, L. Cai, L. T. Cai, Y. S. Guo, and W. Ding. 2020. Effect of temperature on phenotype characterization of *Ralstonia solanacearum* from tobacco. Canadian Journal of Plant Pathology 42:164–181.

Wang, K., D. Zheng, G. Yuan, W. Lin, and Q. Li. 2021. A yceI Gene Involves in the adaptation of *Ralstonia solanacearum* to methyl gallate and other stresses. Microorganisms 9(9):1982.

Wang, X., Z. Wei, M. Li, X. Wang, A. Shan, X. Mei, A. Jousset, Q. Shen, Y. Xu, and V. P. Friman. 2017. Parasites and competitors suppress bacterial pathogen synergistically due to evolutionary trade-offs. Evolution 71:733–746.

Wang, X., Z. Wei, K. Yang, J. Wang, A. Jousset, Y. Xu, Q. Shen, and V. P. Friman. 2019. Phage combination therapies for bacterial wilt disease in tomato. Nature Biotechnology 37:1513–1520.

Wei, Z., Y. Gu, V.-P. Friman, G. A. Kowalchuk, Y. Xu, Q. Shen, and A. Jousset. 2019. Initial soil microbiome composition and functioning predetermine future plant health. Science Advances 5(9).

Wei, Z., T. Yang, V. P. Friman, Y. Xu, Q. Shen, and A. Jousset. 2015. Trophic network architecture of root-associated bacterial communities determines pathogen invasion and plant health. Nature Communications 6:8413.

Weimann, A., K. Mooren, J. Frank, P. B. Pope, A. Bremges, and A. C. McHardy. 2016. From genomes to phenotypes: Traitar, the microbial trait analyzer. mSystems 1(6).

Weller, S. A., J. G. Elphinstone, N. C. Smith, N. Boonham, and D. E. Stead. 2000. Detection of *Ralstonia solanacearum* strains with a quantitative, multiplex, real-time, fluorogenic PCR (TaqMan) assay. Applied and Environmental Microbiology 66:2853–2858.

Wellington, E. M., A. B. Boxall, P. Cross, E. J. Feil, W. H. Gaze, P. M. Hawkey, A. S. Johnson-Rollings, D. L. Jones, N. M. Lee, W. Otten, C. M. Thomas, and A. P. Williams. 2013. The role of the natural environment in the emergence of

antibiotic resistance in Gram-negative bacteria. The Lancet Infectious Diseases 13:155–165.

Wenneker, M., M. Verdel, R. Groeneveld, C. Kempenaar, A. Van Beuningen, and J. Janse. 1999. *Ralstonia* (*Pseudomonas*) *solanacearum* race 3 (biovar 2) in surface water and natural weed hosts: First report on stinging nettle (*Urtica dioica*). European Journal of Plant Pathology 105:307–315.

Wicker, E., L. Grassart, R. Coranson-Beaudu, D. Mian, C. Guilbaud, M. Fegan, and P. Prior. 2007. *Ralstonia solanacearum* strains from Martinique (French West Indies) exhibiting a new pathogenic potential. Applied and Environmental Microbiology 73(21):6790-6801.

Wicker, E., P. Lefeuvre, J. C. De Cambiaire, C. Lemaire, S. Poussier, and P. Prior. 2012. Contrasting recombination patterns and demographic histories of the plant pathogen *Ralstonia solanacearum* inferred from MLSA. ISME Journal 6(5):961-974.

Wickham, H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.

Wickham, H., M. Averick, J. Bryan, W. Chang, L. McGowan, R. Francois, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. Pedersen, E. Miller, S. Bache, K. Muller, J. Ooms, D. Robinson, D. Seidel, V. Spinu, K. Takahshi, D. Vaughan, C. Wilke, K. Woo, and H. Yutani. 2019. Welcome to the tidyverse. Journal of Open Source Software 4:1686.

Williamson, L., K. Nakaho, B. Hudelson, and C. Allen. 2002. *Ralstonia solanacearum* race 3, biovar 2 strains isolated from geranium are pathogenic on potato. Plant Disease 86:987–991.

van der Wolf, J. M., P. J. M. Bonants, J. J. Smith, M. Hagenaar, E. Nijhuis, R. C. M. van Beckhoven, G. S. Saddler, A. Trigalet, and R. Feuillade. 1997. Genetic diversity of *Ralstonia solanacearum* Race 3 in Western Europe determined by AFLP, RC-PFGE and Rep-PCR. In: Bacterial Wilt Disease. Pages 44–49. P. H. Prior, C. Allen, and J. Elphinstone. Springer Verlag.

Wu, K., L. Su, Z. Fang, S. Yuan, L. Wang, B. Shen, and Q. Shen. 2017. Competitive use of root exudates by *Bacillus amyloliquefaciens* with *Ralstonia solanacearum* decreases the pathogenic population density and effectively controls tomato bacterial wilt. Scientia Horticulturae 218:132–138.

Wu, K., S. Yuan, G. Xun, W. Shi, B. Pan, H. Guan, B. Shen, and Q. Shen. 2015. Root exudates from two tobacco cultivars affect colonization of *Ralstonia solanacearum* and the disease index. European Journal of Plant Pathology 141:667–677.

Wu, T., E. Hu, S. Xu, M. Chen, P. Guo, Z. Dai, T. Feng, L. Zhou, W. Tang, L. Zhan, X. Fu, S. Liu, X. Bo, and G. Yu. 2021. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. The Innovation 2.

Xian, L., G. Yu, Y. Wei, Y. Li, H. Xue, R. Morcillo, H. Rufian, and A. Macho. 2019. No A Bacterial effector protein hijacks plant metabolism to support bacterial nutrition. Cell Host and Microbe 28(4):548-557.

Xiong, W., R. Li, S. Guo, I. Karlsson, Z. Jiao, W. Xun, G. A. Kowalchuk, Q. Shen, and S. Geisen. 2019. Microbial amendments alter protist communities within the soil microbiome. Soil Biology and Biochemistry 135:379–382.

Yabuuchi, E., Y. Kosako, Y. Ikuya, H. Hisako, and Y. Nishiuchi. 1995. Transfer of Two *Burkholderia* and an *Alcaligenes* species to *Ralstonia* gen. Nov .: Proposal of R*alstonia pickettii* (Ralston, Palleroni and Doudoroff 1973) Comb. Nov., *Ralstonia solanacearum* (Smith 1896) Comb. Nov. and *Ralstonia eutropha* (Davis 1969) . Nov. Microbiol. Immunol 39:897–904.

Yamada, T., T. Kawasaki, S. Nagata, A. Fujiwara, S. Usami, and M. Fujie. 2007. New bacteriophages that infect the phytopathogen *Ralstonia solanacearum*. Microbiology 8:2630-2639.

Yan, J., P. Li, X. Wang, M. Zhu, H. Shi, G. Yu, X. Chen, H. Wang, X. Zhou, L. Liao, and L. Zhang. 2022. RasI/R quorum sensing system controls the virulence of *Ralstonia solanacearum* strain EP1. Applied and Environmental Microbiology 88(15).

Yang, C., Y. Dong, V. P. Friman, A. Jousset, Z. Wei, Y. Xu, and Q. Shen. 2019. Carbon resource richness shapes bacterial competitive interactions by alleviating growth-antibiosis trade-off. Functional Ecology 33:868–875.

Yang, T., G. Han, Q. Yang, V.-P. Friman, S. Gu, Z. Wei, G. A. Kowalchuk, Y. Xu, Q. Shen, and A. Jousset. 2018. Resource stoichiometry shapes community invasion resistance via productivity-mediated species identity effects. Proceedings of the Royal Society B: Biological Sciences 285:20182035.

Yao, J., and C. Allen. 2006. Chemotaxis is required for virulence and competitive fitness of the bacterial wilt pathogen *Ralstonia solanacearum*. Journal of bacteriology 188:3697–708.

Yao, J., and C. Allen. 2007. The plant pathogen *Ralstonia solanacearum* needs aerotaxis for normal biofilm formation and interactions with its tomato host. Journal of Bacteriology 189:6415–6424.

Yu, G. 2020. Using ggtree to visualize data on tree-like structures 69.

Yu, G. 2022. Data integration, manipulation and visualization of phylogenetic trees. 1st edition. Chapman and Hall/CRC.

Yu, G., T. T.-Y. Lam, H. Zhu, and Y. Guan. 2018. Two methods for mapping and visualizing associated data on phylogeny using *Ggtree*. Molecular Biology and Evolution 35:3041–3043.

Yu, G., D. K. Smith, H. Zhu, Y. Guan, and T. T. Lam. 2017. GGTREE : an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution 8:28–36.

Yuliar, Y. Asi Nion, and K. Toyota. 2015. Recent trends in control methods for bacterial wilt diseases caused by *Ralstonia solanacearum*. Microbes and Environments 30:1–11.

Zarattini, M., M. Farjad, A. Launay, D. Cannella, M. C. SouliÃ, G. Bernacchia, and M. Fagard. 2021. Every cloud has a silver lining: How abiotic stresses affect gene expression in plant-pathogen interactions. Journal of Experimental Botany 72:1020–1033.

Zeglin, L. H. 2015. Stream microbial diversity in response to environmental changes: review and synthesis of existing research. Frontiers in Microbiology 6:454.

Zhang, J., Q. Xiao, T. Guo, and P. Wang. 2021. Effect of sodium chloride on the expression of genes involved in the salt tolerance of *Bacillus* sp. strain "SX4" isolated from salinized greenhouse soil. Open Chemistry 19:9–22.

Zhang, N., D. Wang, Y. Liu, S. Li, Q. Shen, and R. Zhang. 2014. Effects of different plant root exudates and their organic acid components on chemotaxis, biofilm formation and colonization by beneficial rhizosphere-associated bacterial strains. Plant and Soil 374:689–700.

Zhou, T., D. Chen, C. Li, Q. Sun, L. Li, F. Liu, Q. Shen, and B. Shen. 2012. Isolation and characterization of *Pseudomonas brassicacearum* J12 as an antagonist against *Ralstonia solanacearum* and identification of its antimicrobial components. Microbiological Research 167:388–394.

Zuluaga, A. P., M. Puigvert, and M. Valls. 2013. Novel plant inputs influencing *Ralstonia solanacearum* during infection. Frontiers in Microbiology 4:349.