Comparative Genomic analysis of *Ralstonia solanacearum* species complex

Martina Metodieva Stoycheva

Doctor of Philosophy

University of York

Biology

December 2022

ABSTRACT

Ralstonia solanacearum species complex (RSSC) consist of a group of phytopathogenic bacteria that can infect many economically important crops, including tomatoes, potatoes, and bananas. RSSC are very diverse, capable of surviving up to 200 plant host species and various environmental reservoirs such as soils, river water and secondary wild plant hosts. The diversity of the RSSC is often credited to its large bipartite genome (5-6Mbp), that encodes multiple genes linked to virulence and survival across different niches. In this thesis, I investigated the genetic diversity of RSSC at worldwide (55 countries), country (the UK) and crop field (four tomato fields in China) levels. Worldwide, we found that the open pangenome of RSSC contained 18,080 genes. I estimate that the recombination across the phylogeny occurred five times frequently than mutation. Moreover, I show that insertion sequences linked to virulence and metal resistance genes played an important role in the accessory genome diversification of RSSC. Within the UK, I show that the diversification of the clonal phylotype IIB-1 strain is due to initial loss of accessory genes and movement of IS elements with estimated origin of the population dated between 1958 and 1988. At the field level, we show that two to three clonal lineages co-occur within all sampled fields. Interestingly, co-occurring lineages differ in their virulence traits and gene content, which could be due to due to adaptation to different niches within each field. The work presented here lays the groundwork for a systematic understanding of the ecological and evolutionary genomics of the RSSC species complex, at the local and global scale. It also demonstrates that instead of mutations, recombination and highly mobile transposases are important drivers of RSSC genetic diversity.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

TABLE OF SUPPLEMENTARY FIGURES

TABLE OF SUPPLEMENTARY TABLES

Declaration by the 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.

Acknowledgements

All science is a joint effort, and this work is no exception. I cannot express my gratitude to all the people that have helped me complete this work in a few lines, but I will mention the most important figures in my life and science. I want to thank my supervisor Ville Friman for his patience, never dying enthusiasm and empathy without which, this work would not have been possible. I want to thank Evie Farnham for being a top-notch lab buddy and for the cup-cakes on my doorstep she made for my pandemic birthday. I want to thank my co-supervisors Dan Jeffares and John Elphinstone for providing useful insight in genomics and plant pathology. I want to thank the technology facility team at the University of York and more specifically John Davey, Sarah Forester, and Sally James for providing invaluable support in my early struggle with genomics software and DNA extractions. Also, I want to thank Andrea Harper for providing useful comments on my annual reports and always checking up if not only the science is looked after but myself as well. In addition, despite being hundreds of miles away my friends and family in Bulgaria provided loads of love and support for which I am forever in depth and hope to repay them with loads of love and devotion in the years to come. Finally, I want to thank Lauri Mikonranta for being next to me through the pandemic and the most stressful parts of the thesis writeup. Overall this work would not have been possible if it was not for a load of kindness and love provided by the many scientists and friends that were around me so thank you all!

Work Contributions

Martina Stoycheva wrote the work presented in this thesis. Data collection for chapters 2 and 3 were done at Fera Ltd., where both Martina Stoycheva and Evie Farnham at the University of York curated a live bacterial culture collection and created a DNA sequence database from the bacterial cells. Following Martina Stoycheva collected the sequence data and created genome assemblies and annotations. Data analysis for chapter 2 was fully performed by

Martina Stoycheva. Data analysis for chapter 3 was performed mainly by Martina Stoycheva, with insertion sequence analysis performed by Samuel Greenrod at the University of York. In chapter 4, all data collection, experimental work, and initial phenotypic analysis were performed by Gaofei Jiang, Xiaofang Wang, Yuling Zhang, Zhong Wei, Yangchun Xu and Qirong Shen at The Agricultural University of Nanjing. Remi Peynard and Stephan Genin provided intellectual feedback on the manuscript. The final data analysis of phenotypic data, analysis of genome data and manuscript write-up was performed by Martina Stoycheva. In all chapters genome assemblies and sequence data quality checking was performed by Martina Stoycheva. Daniel Jeffares, John Elphinstone, and Ville Friman contributed to conceptualising each chapter. Ville Friman proofread and approved the thesis.

Project Funding

This project was funded by a NERC iCASE studentship and partially funded by Fera Ltd. and The UK Department for Environment, Food & Rural Affairs (Defra).

1. General Introduction

1.1. Overview of bacterial plant infections and evolution of agricultural pathogens

1.1.1. The importance of control and monitoring of phytopathogenic bacteria

Agriculture is a vast industry contributing to 4% of the gross domestic product globally. Around a third of the world's habitable land is used for crop and animal feed production (FAO, 2021). However, almost 770 million people were undernourished in 2020 (FAO, 2021). Therefore, producing and distributing food is an issue we must address to guarantee everyone's access to nutritious food in sufficient quantity, a concept known as food security. Moreover, we are at a crucial point in our planet's climate history, as we must slow down human-induced climate change. Therefore, we need to address the nutritional issues of humanity sustainably without further increasing the carbon pressure on our planet and the destruction of valuable habitats. It has been estimated that worldwide food production has to increase by 50% by 2050 using the same amount of land to match the threats that climate change poses to agriculture and avoid the increase of its effect on climate change (Chakraborty and Newton, 2011). Climate change will lead to the expansion of climatic zones that favour plant pathogens and lead to an increase in pathogen-associated crop losses, especially in temperate zones (Chaloner et al., 2021). Thus, increasing the efficiency of food production from agricultural land by reducing disease-driven crop losses is a crucial step in achieving food security.

One of the biggest problems in agriculture is the loss of crops to phytopathogens. Global estimates show that the reduction in staple food crop yields caused by plant diseases can be up to 10% (Savary et al., 2019a). Although fungal plant diseases are far more prevalent than bacterial ones, multiple soil bacteria can also cause devastating diseases in plants. Therefore, they hugely impact horticulture and agriculture, affecting various crops, including economically essential food staples such as potatoes, rice, and wheat (Mansfield et al., 2012).

Plant-pathogenic bacteria are found worldwide but have a greater impact in tropical and sub-tropical climates where warmth and humidity provide ideal conditions for bacterial growth (Jiang et al., 2021a; Kannan et al., 2015). Pesticide control is one of the major and most common methods to tackle plant diseases; thus, world pesticide use increased by 36% between 2000 and 2019 (FAO, 2021)**.** However, unlike the wide range of fungicides available for suppressing fungal diseases, multiple options are not available for controlling bacterial infections. One of the most popular methods is the use of copper-based pesticides, however, results are at the level of suppression at best and bacterial resistance to copper is not uncommon (Colombi et al., 2017). Alternatively, just like with human bacterial pathogens, antibiotics can be used to control plant pathogenic bacteria. For instance, in the USA, streptomycin and oxytetracycline are two commonly used compounds against *Erwinia amylovora*, *Pseudomonas spp*., and *Xanthomonas campestris* (McManus et al., 2002). However, the evolution of antibiotic resistance is a major concern in both agriculture and human healthcare and similar antibiotics are used in both fields (George W Sundin and Wang, 2018). Antibiotic and pesticide resistance can lead to large outbreaks of pathogenic strains that we do not have the ability to control. Thus, to prevent the spread of resistant strains regulation of antibiotics and pesticides in animal and plant agriculture is getting stricter (FAO and VMD, 2022). The regulation requires more controlled application of chemical control methods in a complex approach along with research-backed crop management practices such as the use of tolerant or resistant plant varieties, crop rotation, clean and verified seed materials, and monitoring and sanitising of known outbreak areas. Developing such combined approaches requires substantial financial investment and a deep understanding of the biology of bacterial plant pathogens.

1.1.2. The importance of pathogen evolution for sustainable agriculture

Agriculture is already threatened by climate change which is predicted to cause severe weather conditions in most of the world and lead to the emergence and establishment of new diseases (Chaloner et al., 2021). Warming up of the planet means expansion of the climatic zones with favourable conditions for plant pathogenic bacteria. Pathogens will be able to expand to previously unaffected areas of the world and cause new epidemics. The research

into plant-infecting bacteria and how they evolve and adapt to new areas and conditions is thus extremely important for agriculture and food security. The evolution of a pathogen to its environment is tightly coupled with the evolution of its host (Pfennig, 2001). Therefore, to avoid the emergence and spread of pathogens, and to control known contaminated areas without increasing the toll on the environment, we need to understand the pathogens' capacity to evolve and adapt to their hosts across different environments and in the context of control methods and climate change. We can only create effective resistant crop varieties and control techniques if we understand how the pathogen may respond to the pressures created by both the natural and human-made environment (Shipton, 1977). For instance, the plant pathogenic bacteria *Xylella fastidiosa* causes disease on olive trees and has spread around the world with human movement and trade (Morelli et al., 2021). However, it was shown that a 2016 outbreak in Italy was greatly facilitated by the agricultural practice of planting dense plant host monocultures and specifically planting two autochthonous susceptible olive tree cultivars (Luvisi et al., 2017). Dense plant monocultures usually used in agriculture have limited genetic diversity and are unable to evolve defences against pathogens. Thus, achieving control over bacterial plant diseases requires a deep understanding of how bacteria evolve on different scales, from the individual plant to the agricultural field and the wild plant populations surrounding them. Predicting and experimentally testing the possible evolutionary trajectories for pathogens within new locations and climatic conditions can prepare us and help us design evolution-aware control techniques. The importance of understanding the evolutionary capacity of a pathogen and its ability to jump hosts has never been as evident as it is now with the ongoing viral pandemic of the human respiratory disease COVID-19 (Machado et al., 2021). Previous research in vaccine development and related viruses allowed researchers to design a vaccine quickly. Understanding the biology, epidemiology, and bacterial adaptability to human-made agricultural environments will help us prevent outbreaks and epidemics that threaten food security and design control techniques that consider the organism's evolutionary capacity.

1.1.3. How do bacteria cause disease in plants?

An invisible arms race is continuously occurring between plants and phytopathogenic bacteria. The long coevolution of plants and bacterial pathogens has created an array of plant immune defences and bacterial virulence factors to evade them. The virulence factors need to overcome the host's defence mechanisms to establish infections and thus determine the pathogenicity of the bacterial species and can be highly specific to certain plant hosts (Ngou et al., 2022). Plant pathogenic bacteria can produce enzymes to help dissolve plant cell walls, excrete exopolysaccharides that block the vasculature of a plant, evade the immune system of the host by injecting effector molecules in and around its cells, induce tumour growth by manipulating the host DNA, and produce toxins that can lead to necrosis, chlorosis, and gummosis of the host tissue (Alfano and Collmer, 1996). In response, plants have evolved multiple defences to protect their nutrient-rich cells and kill unwanted invaders. One essential defence mechanism, termed the hypersensitive response, is when plants trigger controlled cell death (apoptosis) of infected cells and the surrounding tissue after recognising an infection (Alfano and Collmer, 1996). Plants' innate immunity depends on cell surface receptors and kinases recognising conserved pathogen-associated molecular patterns (PAMPs). PAMPs can refer to fungal or bacterial molecules, but for bacteria, they are usually proteins on the surface of the bacterial cell. For instance, the protein flagellin is the main structural component of bacterial flagella, and plants have cell surface receptors triggering a cascade of molecular defences in response to detecting flagellin. Other bacteria-associated molecules include cell wall and cell membrane proteins such as lipopolysaccharides (LPS), phospholipids and peptidoglycans, elongation factor Tu, bacterial nucleic acids, and many pathogen-host relationship-specific molecules. When the plant detects PAMPs, it triggers downstream immunity, including the activation of protein kinases, the production of reactive oxygen species (ROS), immunity-related gene expression, and the deposition of the complex branched carbohydrate molecule callose creating an additional physical defence to the cell (Jones and Dangl, 2006). This response is known as PAMP-triggered immunity and is usually sufficient to suppress the growth of most microbes (Macho and Zipfel, 2015). To avoid activating the immune response of plants, bacteria have evolved methods such as the secretion of enzymes, effector molecules, extracellular polysaccharides (EPS) and toxins

(Alfano and Collmer, 1996). The plants have a specialised effector-triggered immunity (ETI) to target the effectors produced by bacterial pathogens (Ngou et al., 2022).

Overall, the establishment and propagation of a bacterial infection on a plant can be broken up into four stages: 1. Detect and move to the host; 2. Enter the host and evade its immunity; 3. Grow and multiply inside the host; 4. Transmit to new hosts. The plants respond to each of these stages with different defence molecules and mechanisms, and the bacteria have a set of immunity suppression tricks to help them escape the defence. Understanding the molecular interplay of virulence factors and the plant's immunity during infection is essential for understanding how bacterial populations evolve.

1.1.3.1. How do bacteria find and enter new plant hosts?

To successfully invade a plant, a bacterial cell must first find it, then get to it, and finally, enter it by penetrating the external epidermal barriers of the host. The dispersal of bacteria is made easy, as they are microscopic and motile. Most bacteria have flagella or pili that allow them to move both towards and within a host organism. For instance, soil bacteria can use their flagella for chemotactic motion, with their flagella helping them to detect signal molecules such as host plant root exudates and use this signal to move along a molecular gradient towards the host (Parales and Harwood, 2002). Pilli can also play a vital role in twitching motility which is a motion that requires the attachment of pili to a surface and the movement of the cell in a grappling hook manner. This motion can be used to move across short distance such as inside the inside the plant host. For instance, the plant pathogenic bacterium *Ralstonia solanacearum* uses twitching motility to move around its plant host's vasculature and form biofilms (Corral et al., 2020). Therefore, bacterial motility can provide a means of finding a potential host, getting to it, and moving around it.

Furthermore, due to bacteria's small size, they can be easily transported by other organisms, seeds or even water droplets in the air. In addition, bacteria's involvement in cloud formation has been suggested as a potential mechanism for dispersal over large distances and continents (Joung et al., 2017). In an equally clever way, as plants use insects to help them disperse their tiny seeds, some small pathogenic bacteria use insect vectors to transport whole cells from one host to another and deliver the cells directly inside the plant. For

instance, the sap-feeding insect Asian citrus psyllid (*Diaphorina citri*) feeds on citrus plants. It can transmit bacterial plant pathogens within the *Candidatus* Liberibacter genera to their desired hosts. The insect injects the bacteria into the phloem, where they can cause Huanglongbing - the most devastating disease of citriculture globally (Bové, 2006). Most pathogenic bacteria, however, do not need a vector to enter their plant host as they can swim through openings in the plant surface, such as leaf stomata and root openings. Plants need these openings to get oxygen, water and nutrients from the environment and have evolved multiple defences against pathogens entering through them. For example, plants can close their stomata to prevent the entry of foreign organisms (Zhang et al., 2008). Stomatal closure is a mechanical defence mechanism employed by plants as a response to detecting pathogen-associated molecular patterns (PAMPs) by surface receptors that provides an effective pathogen entry control and is a good innate immunity for the plant (Bharath et al., 2021; Sawinski et al., 2013). In response to this physical defence, bacteria have evolved to exploit the plant's systems for controlling stomatal closure. One example is the plant pathogenic bacteria *Pseudomonas syringae* and the phytohormone coronatine that it can secrete. Coronatine functions as a structural and functional analogue of jasmonic acid and related signal molecules in the plant (Sakata et al., 2021). The plant hormones salicylic acid and jasmonic acid play critical roles in the plant's pathogen defence response activation and diurnal cycle control system. *P. syringae's* coronatine acting as jasmonic acid inhibits stomatal closure through guard cell–specific inhibition of NADPH oxidase-dependent reactive oxygen species (ROS) production that is triggered in response to bacterial microbe-associated molecular patterns (MAMP) or darkness (Sakata et al., 2021). Another vulnerable natural opening on the plant surface is the site of secondary root emergence. New secondary roots must emerge for roots to grow and allow the plant to get to new nutrient reserves. This rather aggressive process leaves a susceptible spot in the root cortex as the new root needs to be connected to the inner vascular tissues (Péret et al., 2009). This vulnerability provides a gateway for pathogens that infect the plant's vascular tissues. For instance, the xylem colonising bacterium *Ralstonia solanacearum* (Vasse et al., 1995) can use sites of secondary root emergence as entrance or swim through root lesions which are common on the root surface due to soil disturbances or parasitic nematodes chewing on the roots (Siddiqui et al., 2012).

Generally, motility and the presence of flagella are essential for bacterial virulence and plant invasion. Thus, it is unsurprising that plants have evolved mechanisms to recognise flagella and trigger an immune response. The non-flagellated forms of many bacterial species have been reported to be less pathogenic or avirulent (Montie et al., 1982). In many cases, the flagella are needed for expressing secondary virulence factors and attachment to the host (He and Jin, 2003). Overall, bacteria have found multiple ways to find and enter their hosts, exploiting animals, using by-products of the plant's metabolism as signals, and using vulnerabilities in their surface to invade them (Meena et al., 2019).

1.1.3.2. How do bacteria evade the plant defence responses inside the plant?

Once a pathogen has entered the plant, it must evade the host's defences inside to get the nutrients it seeks. Plant host invasion by bacteria often requires the secretion of molecules that block plant receptors to hinder the immune response. Plant pathogenic bacteria secrete enzymes such as cellulases, pectinases, ligases, and proteases (Frees et al., 2013; Maculins et al., 2016; Prade et al., 1999; Wilson, 2011). These enzymes penetrate living plant cells as they can degrade cells' building blocks, such as cellulose, pectin, and lignin. The secretion of these enzymes happens in specific pathogen–host combinations shaped by coevolution. Thus, the bacterial secreted molecules vary greatly between species and strains.

Gram-negative bacteria have seven (types I, II, III, IV, V, VI, and IX) secretion systems that aid the transport of molecules inside and outside the bacterial cell. These systems are crucial for virulence and are hugely varied depending on the pathogen-host interaction due to the long co-evolutionary time bacterial pathogens and their hosts have had together (Chang et al., 2014). One of the major systems for plant pathogens is the type III secretion system (T3SS), which delivers type III effector proteins directly inside the plant cell (Meena et al., 2019). The released proteins can be either effectors delivered into the host cell's cytosol, or translocators, which aid the effectors' transport across the eukaryotic cell membrane. These proteins are encoded by the *hrp* (hypersensitive response) genes, which are highly conserved across bacterial pathogens. They acquired their name as the T3SS was first described in phytopathogens in association with plants' hypersensitive response defence mechanism (Alfano and Collmer, 1996). Type III effectors (T3Es) are weapons against the plant's innate and specialised effector-triggered immunity (ETI). T3Es can interact with plant proteins and

DNA inside and outside the plant cell in a precise manner driven by the host and pathogen coevolution. They have various activities, such as inhibiting or eliminating the activity of host cell surface proteins, interfering with key modules of the plant's immune signalling, such as signalling kinase cascades and immune receptor complexes, activating host gene transcription for pathogenesis, etc. (Feng and Zhou, 2012).

For instance, the type III effector HopAI1 is widely conserved in bacterial pathogens of plants and animals. In the well-studied plant pathogenic bacterium *Pseudomonas syringae*, HopAI1 inhibits the model plant *Arabidopsis thaliana's* mitogen-activated protein kinases (MAPKs), usually activated by exposure to pathogen-associated molecular patterns (PAMPs) (Zhang et al., 2007). The inhibition of MAPKs by HopAI1 suppresses two independent downstream events that enable the pathogen to establish an infection: the reinforcement of cell wall defence and transcriptional activation of PAMPs response genes (Zhang et al., 2007). HopAI1 is a widely conserved T3E gene important for many pathogenic bacteria. Still, the specificity of the interplay between type III effectors and plant immunity is the limiting factor for bacterial species infecting different plant strains (Cornelis and Van Gijsegem, 2000). The diversity of effectors in pathogenic bacteria directly correlates with the number of hosts a species can infect because effectors have unique binding or enzymatic activity within the eukaryotic host cell. *Pseudomonas syringae* can infect multiple plant hosts; thus, it has a vast repertoire of type III effectors that it must maintain, possibly with some cost, to keep its ability to be a broad-host pathogen (Lindeberg et al., 2009). A study of nearly five hundred *P. syringae* strains isolated from over one hundred hosts worldwide identified 14,613 putative T3Es helping the *P. syringae* strains infect so many hosts (Dillon et al., 2019a). The study also found a strong signal of positive selection on the type III effector genes, demonstrating the importance of these proteins in bacterial infection and the continuous arms race between plant defences and T3Es. The interplay of type III effectors together is understudied but genomics is now paving the way to understanding why some pathogens have a great diversity of effectors. For instance, strains of *Pseudomonas* spp. have around 30–40 core T3Es shared across the genus but the essential genes and epistatic effect of effectors in plant pathogens is an ongoing study (Sanchez-Garrido et al., 2022; Wei et al., 2015a).

While T3Es are involved in the precise immunity suppression of plant immunity, type II secretion systems (T2SS) are often involved in the cell wall degradation enzymes and are specific to *Proteobacteria* (Cianciotto and White, 2017). For instance, *Xanthomonas campestris* pv. vesicatoria the causal agent of bacterial spot disease in tomatoes and peppers secretes proteases and xylanases via the T2SS (Solé et al., 2015). Many of the important plant pathogens use type II secretion system for secretion of enzymes including *Xanthomonas oryzae, Erwinia amylovora, Dickeya dadantii, Ralstonia solanacearum, Pectobacterium carotovorum,* and *Xylella fastidiosa* (Cianciotto and White, 2017). Type IV systems are associated with pili and attachment to eukaryotic cells and are also common in plant pathogens (Meena et al., 2019). Some pathogens use a great diversity of type IV effectors. For example, the human pathogenic *Legionella* species are reported to have over 300 type IV effectors helping them evade human immunity (Gomez-Valero et al., 2019). Overall, Type I-VI secretion systems are essential for bacterial virulence and are involved in many stages of the hide-and-seek game bacteria play with the plant cells. Despite the massive amount of literature available on secretion systems and the knowledge of their crucial role in the infection of both animal and plant cells, we are yet to uncover the function of most of these molecules. Both the large amount of redundancy in these genes within pathogen genomes and the trade-offs encountered with maintaining larger genomes with high redundancy are poorly understood.

1.1.3.3. How do bacteria control the secretion of virulence factors?

Bacteria use secreted molecules for virulence and other vital interactions with their environment. There is, of course, a cost to producing secreted molecules, and bacteria have evolved a clever method to control the secretion tightly. They can recognise molecules secreted by themselves and other genetically similar individuals around them. This way, they can sense their population size and accordingly regulate environmental responses. This density-dependent environment sensing is termed quorum sensing (Von Bodman et al., 2003). Many bacteria use self-generated molecules to activate intracellular processes associated with virulence or defence, such as forming biofilms, cell adhesion, secretion of exopolysaccharides (EPS), siderophores and antibiotics. For example, the maize pathogen *Pantoea stewartii* ssp. *stewartii* produces EPS, necessary for biofilm formation, in a cell density-dependent manner controlled by a quorum sensing system. Two genes, *esaI* and *esaR*, encode essential regulatory proteins for quorum sensing. The product of the *esaI* gene

is the molecule N-acyl homoserine lactone (AHL), which acts as a signal for the *esaR*-encoded cell receptors to ensure the bacterial population produces enough EPS according to the cell density. Mutant strains in the *esaR* gene synthesise EPS constitutively at low cell densities and are significantly less virulent than the wild-type (Koutsoudis et al., 2006). Therefore, controlling and activating virulence by bacteria is an active process that they can regulate depending on the environment to minimise costs.

1.1.3.4. Spread of plant bacterial diseases

Bacteria can often survive and persist for long periods without access to nutrients in anticipation of favourable conditions. However, their success as plant pathogens would not have come if it was not for humans. Human movement and agricultural practices have allowed the dispersal of seeds and plants and their pathogens worldwide, thus providing an unexpected means for a plant disease to spread (Sotiropoulos et al., 2022). For example, the 2010 vertical oozing canker disease pandemic on kiwifruit would not have been able to spread across the world this rapidly if it were not for human agricultural practices (Colombi et al., 2017). In this case, copper pesticides, growing highly susceptible plant monocultures, and spreading pathogens with our movement across the world were responsible for the pandemic (Colombi et al., 2017).

Pathogenic bacteria have a variety of virulence factors to ensure their survival and fitness inside a plant host. Diversity in these factors creates complex interaction networks between the pathogen and its host. The networks are under constant evolutionary pressure that also depends on environmental variability and, thus, should be able to withstand specific ecological conditions. To understand these networks, it is crucial to understand the patterns of molecular evolution in bacteria. On the plant level, we need to know the drivers of selection among all plant life stages, different parts of the infection cycle, and the trade-offs between different hosts. On a macroevolutionary scale, we need to understand the survival of bacterial species across different habitats over time and the impact of human agricultural practices on the spread and success of the pathogen's survival and evolution.

1.1.4. How do bacterial pathogens evolve to new environments and hosts?

Bacterial populations have big effective population sizes (N_e), allowing genetic variation to accumulate quickly over time, increasing the available space for natural selection and the efficiency of deleterious mutation purging (Lynch and Conery, 2003). This allows for streamlining of the genomes and also observing adaptation to new selection pressures in real-time both in the lab and the environment at the timescale of just several weeks (Lynch and Conery, 2003). Therefore, bacterial study systems have been hugely influential in the testing of evolutionary theory and have provided a wealth of knowledge that can be applied to improving theoretical models of evolution as well as devising control strategies for pathogens (Buckling et al., 2009).

1.1.4.1. Evolution of host-adapted bacterial lineages

A central aspect in the study of pathogen-host interaction is explaining why certain pathogens are associated with specific hosts. Still, it can be hard to determine the genetic factors contributing to bacteria's ability to infect a particular host. Phylogenies of bacterial species and strains can show signals of population structure linked to niche occupation and host specificity. For instance, the phytopathogen *Pseudomonas syringae* has host-specific (specialists) and wide-host (generalist) phylogenetic lineages, however, determining the genes associated with this clustering pattern is not technically or theoretically simple (Hwang et al., 2005; Lindeberg et al., 2009). The presence of a host-associated phylogenetic pattern in a phylogeny can be interpreted within two main theoretical frameworks. The first one is that selection drives the evolution of the ability to infect a specific host and has thus resulted in phylogenetic lineages that are defined by that ability. The second is that the presence of a colonisation barrier between different hosts has, over time, led to the accumulation of mutations and, thus, the separation of lineages by host association. These contrasting explanations of molecular evolution patterns focus on either selection for adapted lineages in different environments (ecotypes) or genetic drift (neutral diversification) (Sheppard et al., 2018). In reality host associated phylogenetic clustering is probably a result of both these evolutionary mechanisms, but it is important to understand how they can differentially affect the distribution of species and genes on a phylogeny in order to study them.

The ecotypes hypothesis stems from the idea that in theory large population sizes such as those seen in bacteria over time lead to a streamlined genome without many deleterious mutations and parasitic gene elements and with fewer total number of genes (Lynch and Conery, 2003). Thus, the stable coexistence of host-specialised lineages implies that there is a selective advantage in maintaining this population structure. Therefore, the genes shared by bacteria infecting a specific host should be providing this selective advantage (Sheppard et al., 2018). The ecotype hypothesis is of scientific interest for genetics and pathologists because we can ask what genes lead to a particular host specialist bacterial lineage. These genes could be associated with virulence and immunity within a host species. Therefore, knowledge of these genes can be essential for practical applications in agriculture and medicine. For example, they can be targeted for drug discovery or used to understand how host jumps occur and potentially prevent them and future spread of a high virulence gene (Sonehara and Okada, 2021). However, one caveat to consider when focusing on host-adapted lineages and genes is the limited number of hosts we can sample in any one host range study (Dallas et al., 2017). Host specialists are hard to identify as a sampling of hosts is usually inexhaustive, mainly because we often need to consider wild animal and plant host populations in our studies. Thus, generalists may be more common in nature than we have observed, and it is hard to determine how the host associated genes contribute to virulence in different environments. Most importantly observed correlation between gene phylogenies and same host associated bacterial lineages could be due to chance and not as the ecotype hypothesis suggests a result of balancing selection maintain population structure of the coexistence of different host-associated lineages.

The second explanation focuses on the null model of evolution (Kimura, 1968). Genetic drift is a mechanism in which allele frequencies of a population change over time due to random chance. Through this process even in the absence of selection and physical or spatial barriers, a degree of genetic variability most of which neutral accumulates over time in a population (Sheppard et al., 2018). By studying the distribution of this variation on a phylogenetic tree of a species or a group of species we should be able to trace it back to the most recent common ancestor of the species. This relationship between time and the accumulation of variation has been used to fit a genetic or molecular clock to phylogenies. The fitting of a molecular clock can help us estimate the evolutionary rates and deviations from them that are signs of

evolution. Dating bacterial phylogenies and many epidemiological techniques are based on this idea and have proved very useful in tracking outbreaks and outbreak development and origin (Neogi et al., 2012). Moreover, genes that do not follow the expected structure of the species' phylogenies and thus deviate from the null model are of interest as this will be a sign of positive or negative selection that has been acting on them and has made their inheritance pattern different (Kuo et al., 2009). Also, the dissemination of a gene multiple times in a horizontal rather than a vertical manner across the branches of the phylogeny could signify that this gene can provide a substantial selective advantage but will not follow the linear expectation of the vertical transmission of clonally propagated genes. For instance, a pathogen population within an agricultural field and the "wild" population of the vegetation surrounding the field are rarely considered together. Still, gene exchange between these populations is not unlikely. For example, two phytopathogenic *Pseudomonas syringae* strains were more closely related to surrounding environmental isolates than isolates from other agricultural fields, suggesting that transfer between the environment and field could be common (Monteil et al., 2016).

Furthermore, there are technical limitations to identifying genes associated with host adaptation. Due to their clonal reproduction, bacterial populations often display strong signals of population structure. Therefore, the genes responsible for adaptation to an environmental niche or a host can be hard to distinguish from genetically linked blocks of genes have moved together horizontally on a phylogeny. Nevertheless, host-specialised genes have been a central theoretical concept for many comparative genomics studies. Many have identified potential genes responsible for host specialisation (Mak et al., 2013; Nowell et al., 2016). A popular technique that tries to associate genes or genetic variants to hosts by controlling for population structure and studying all the genetic variants in a genome is the genome-wide association study (GWAS). It was initially developed for diploid organisms as it relies on linkage disequilibrium. Linkage disequilibrium is a population genetics term that describes the non-random association of alleles at two or more loci in a population. If two loci are more often than expected found together on a genome than expected, then they are said to be linked (Slatkin, 2008). The lack of randomness in the distribution of genes within clonally reproducing organisms makes linkage virtually genome wide. However, it has recently been applied to bacteria, and host-associated genes have been identified through GWAS

sometimes referred to as microbial GWAS (Gori et al., 2020). This approach tries to control for the genetic linkage observed in bacterial populations and associate genes with specific phenotypes or environmental niches. Likely, it will have a considerable impact on research in evolutionary microbiology in the future. GWAS often tries to correct population structure using the signal of the phylogeny (Collins and Didelot, 2018). The phylogeny, however, would be affected by adaptation to specific niches and recombination. Demography inferences based on population structure are based on the neutral evolution model. Still, it has been shown that inferring demographic changes based on genealogies are systematically biased in the bacteria (Lapierre et al., 2016). Therefore, some consider bacterial populations as populations of genes and suggest alternative models where genes rather than clonal lineages inhabit niches (Arnold et al., 2022). Thus, in modern big data biology, both theoretical interpretations are useful and help us focus on different aspects of comparative genomics to understand how pathogens evolve to infect new hosts (Sheppard et al., 2018).

1.1.4.2. Bacterial sex and horizontal gene transfer

Only a minority of prokaryotes are truly clonal, and bacteria often recombine (Bobay and Ochman, 2017). Recombination can be the fastest way to bring beneficial genes together and has been shown to contribute to the evolution of virulence and adaptation to new niches in bacteria, however, it can also disrupt gene combinations. Therefore, there are two interpretations of the presence and maintenance of recombination systems in bacteria. The first one is that damaging selfish genetic entities such as transposons and small parasitic DNA elements are maintaining this system which is not necessarily beneficial for the host but rather a combination of randomness and adaptiveness for the small genetic elements. The second one is an analogy to sex in eukaryotes, where the maintenance of homologous recombination in bacteria is evidence for it playing a role in improving the response time to natural selection. The idea is that recombination can bring together sets of genes that, when combined, provide benefits over and above the benefits that the genes could confer on their own. Simulations show that it is unlikely that gene shuffling evolved in response to selection for increased rates of evolution (Levin and Cornejo, 2009).

Unlike recombination by meiosis in eukaryotes, where two cells exchange DNA, bacterial recombination is single-sided and not linked to reproduction. DNA is taken up by the cells in

three known ways: from the environment through natural transformation, from another cell with the aid of a specialised structure through conjugation or transferred between cells by phage viruses (Vos, 2009). Bacterial recombination can be non-homologous, where foreign genetic material inserts into the chromosome at repetitive regions via mobile genetic elements like bacteriophages or transposable elements (Straub et al., 2021a). Plasmids can also introduce genes into the bacterial genome, but they do not usually integrate directly into the chromosome. Mobile genetic elements include prophages, integrons, integrative conjugative elements, conjugative transposons, integrated plasmids, insertion sequences, and plasmids (Langille et al., 2010). Mobile genetic elements can lead to shared genes in previously unrelated pathogens occupying a similar niche, thus bridging the species barrier. The spread of virulence factors within and between bacterial species is due to the spread of pathogenic regions through horizontal gene transfer (HGT). Entire regions of effector genes can be lost from the bacterial cell if the gene is carried on mobile DNA elements, as they can be lost during cell replication or through rare spontaneous deletions.

Sequences acquired from other organisms often have a different sequence composition or signature. The most common measures of this are the GC content of a sequence and codon bias. Scanning the genome to identify regions of sequence bias is a quick method to identify mobile genetic elements and to estimate a HGT percentage within a genome (Worning et al., 2000). A more accurate method of estimating HGT uses the species' phylogeny (Azad and Lawrence, 2011). Normally, if bacteria are fully clonal, every gene's phylogeny should follow the species' phylogeny. Therefore, identifying genes with alternative phylogeny patterns and whose presence is polyphyletic, points to independent events of gene gain and loss. Sequences within genomic islands, which are areas of the genome associated with mobile genetic elements, have more genes associated with virulence than the rest of the genome. When 631 complete genomes of pathogenic bacteria were compared, on average, 5.1% of the genes in genomic islands were virulence factors, compared to 1.3% of genes outside of the genomic islands (Sui et al., 2009).

The key type III secretion virulence system has been acquired by horizontal transfer in various pathogenic bacteria (Gophna et al., 2003). Comparing the *hrp* gene sequences and re-arrangements within them in various plant pathogens shows two independent gene similarity

groups (Alfano and Collmer, 1996). The discrepancy between the *hrp* gene similarity groups and the taxonomic relationships of the bacteria within them is consistent with the horizontal acquisition of the system by phytopathogens (Alfano and Collmer, 1996). Prophages have been shown to lead to increased virulence when integrated into genes and to contribute to the spread of virulence between pathogens on different hosts. Virulence on cherry plants in *Pseudomonas syringae* strains is a convergent trait, and thus host-specific genes should be easily identifiable from the phylogeny of the phytopathogenic *Pseudomonas* species. Virulence genes can be identified on three phylogenetically separate lineages and amongst them are *hopAR1*, *hopBB1*, *hopBF1*, and *hopH1*, and one of them, *hopAR1*, is in association with prophage sequences, thus, suggesting that virulence on cherry was spread by prophage movement (Hulin et al., 2018).

Furthermore, gene acquisition by horizontal gene transfer has been shown to lead to outbreaks. In 2010 a strain of *Pseudomonas syringae pv. actinidiae* caused a pandemic of vertical oozing canker disease on kiwifruit after acquiring copper resistance genes encoded by integrative conjugative elements and plasmids found in related *P. syrinage* strains (Colombi et al., 2017). The presence of mobile plasmids can also be used as a means of hiding genes from the plant's immune system. For instance, strains of the bacterium *Pseudomonas syringae* can cause halo blight disease in beans, having a clever way to avoid triggering the plant's hypersensitive response. They have a 106-kilobase pair mobile genomic island containing Type III effector-encoding genes needed for the infection of bean plants. The bean plant can specifically recognise this genomic island and trigger the hypersensitive response. To avoid detection, the genomic island can be separated from the main chromosome as a circular plasmid, becoming a separate circular molecule. The molecule is then supercoiled, making the bacteria invisible to the plant's immune system (Neale et al., 2018). Nevertheless, recombination within a bacterial species does not preclude the occasional expansion of clonal lineages, especially when a single gene or mutation leads to a strong selective sweep. Thus, the whole spectrum of population structures is observed in natural populations.

1.1.4.3. Bacterial species and pangenomes

Bacteria's lack of sexual reproduction makes it challenging to define a species, as the reproductive barrier approach used for higher organisms does not apply. Thus, modern

bacteriology often represents bacterial species as ecotypes (Cohan, 2002). To study bacteria in the lab, some practical rules are used to classify them. Traditional microbiology methods include biochemical assays and substrate utilisation to group bacteria into ecotypes based on metabolic capacity. The evolutionary view of a species is an independently evolving entity which is an important construct as we need a unit of evolution and has no strictly defined boundaries in the natural world but rather we use defined boundaries for practical reasons (Stackebrandt et al., 2002). For example, in modern genomics, the ease of sequencing whole genomes has led to the use of average nucleotide identity (ANI) thresholds based on the whole genome similarity or similarity of specific marker genes (Ciufo et al., 2018). A species can be defined as a group of bacterial genomes with more than 95 % ANI within the group, including a type strain, and less than 95 % ANI to any other known type strain (Konstantinidis et al., 2006).

In the ecotype or the genomic species, a definition of distinct units of overlapping ecological and genotypic similarity is required, which will form if microgeographic separation coupled with ecological trade-offs poses barriers to the gene flow (Shapiro and Polz, 2014). The lack of distinct species boundaries in bacteria has led to the development of the pangenome concept (Tettelin et al., 2005). The pangenome of a species consists of the core genome and accessory genome. The core genome is all the shared orthologous genes between 90%-95% of the bacteria under study, which are present in all strains and members of the species. The accessory genome is the rest of the genes in lower frequency and can be considered transient as they are not present in all strains and members of the species. The accessory genes however are not necessarily non-essential as the essentiality of genes concept stems mainly from the transposon libraries used in traditional genomics and have some laboratory evidence backing the term (Rancati et al., 2018). Pangenome is a fluid term that refers to all the genes in a population, a species, a few species, or a group of closely related strains, depending on the research questions.

Pangenomes are now widely used in the study of bacterial genetic diversity, as biologists have needed help explaining the diversity of genes present in a species, considering the long-term effects of enormous population sizes bacterial species tend to have. Therefore, it is thought that the accessory genes in a population should provide a selective advantage in specific

niches to be maintained in the pangenome of the species (McInerney et al., 2017). In general, this is referred to as the niche-specific accessory genome hypothesis. It explains adaptation with gene-specific sweeps where ecologically adaptive mutations or genes can spread within populations independently of their original genomic background in the presence of recombination and moderate selection. This is in contrast to the genome-wide selective sweeps, where a genotype sweeps through the population when selection's effect is stronger than recombination (Shapiro and Polz, 2014). Defining the essential genes in the species is more challenging than representing all the shared genes as it often required for the gene to be necessary for reproductive success (Rancati et al., 2018). Still, the presence of accessory genes and pangenomes suggests bacteria can maintain accessory genes in the population that are likely beneficial in different environments.

Accessory genes are typically acquired by horizontal gene transfer and tend to have a mosaic structure. Horizontal gene transfers can homogenise the genetic diversity between species and bring strains together. For instance, the *Pseudomonas syringae* species complex has been proposed to represent a single species because HGT of virulence genes has been found to occur across the entire complex (Dillon et al., 2019b). HGT and differential gene loss are the main contributors to the formation of pangenomes (Azarian et al., 2020). The maintenance of some accessory genes encoding virulence factors may be subject to positive selection from the plants' immune response. It could be essential for a strain living in a specific condition or host. Depending on the host species, individual host-specific virulence genes within the *Xanthomonas campestris* genome can influence its growth (Gabriel et al., 1994). Therefore, single genes in the bacterial genome can bring enormous selective advantages. The study of niche-specific accessory genomes could help us identify causal genes and potential gene targets for controlling bacterial disease. We could link variations in virulence factor repertoires such as toxins, phytohormones and secreted substances to disease severity and host resistance. The accessory genome can vary independently of the core genome and can therefore be useful and important in predicting the evolution of a population (Lees et al., 2019).

1.2. *Ralstonia Solanacearum* Species Complex (RSSC)

Ralstonia solanacearum species complex (RSSC) strains are a group of bacterial plant pathogens commonly referred to as simply *Ralstonia solanacearum*. They are aerobic, motile gram-negative β-proteobacteria with pili, one to four polar flagella, and 0.5-1.5 µm long rod-shaped cells (Figure 1 a,b) (Corral et al., 2020; Guarischi-Sousa et al., 2016). *R. solanacearum* cells can be found free living in the soil or in association with a plant host (Jiang et al., 2021a). The complete sequence of the whole genome of the *R. solanacearum* type strain GMI1000 shows that it possesses a two replicon genome with a total length of 5.8 Megabases (Mb), organised into two replicons: a Chromosome: 3.7 Mb and a Megaplasmid: 2.1 Mb, the two molecules have nearly identical G+C content (67.04% and 66.86%, respectively) (Salanoubat et al., 2002). The strains within the species complex are very diverse phenotypically and genetically, so the genome length can vary greatly between them, with up to 1 Mb size differences observed (Genin and Boucher, 2004).

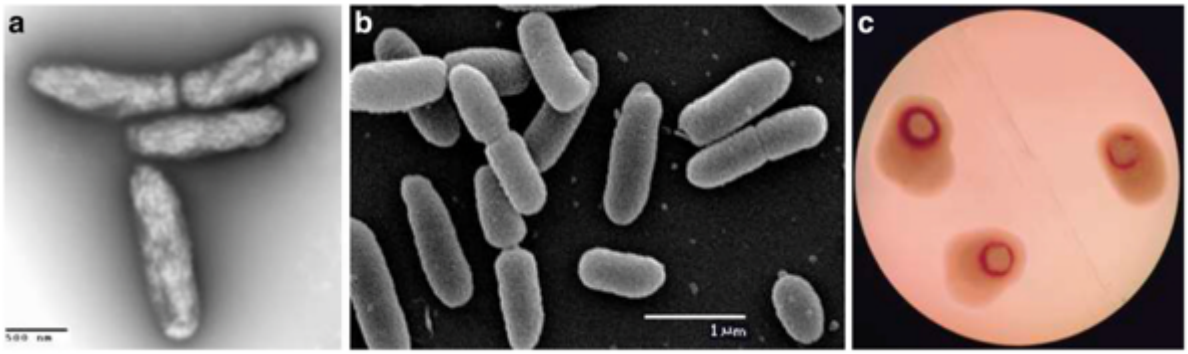


Figure 1. ***Ralstonia solanacearum* cells.**

Ralstonia solanacearum pictures of the strain UY031. a and b showing electron microscopy photos. c showing the Nile Blue test used to identify R. solanacearum cells - virulent colonies develop pearly cream-white, flat, irregular, and fluidal colonies, often with characteristic whorls in the centre. In contrast, avirulent colony forms are small, round, non-fluidal, and entirely cream-white (Hayward, 1960). Figures are taken from: (Guarischi-Sousa et al., 2016, p. 031).

Figure 2. **Disease symptoms caused by *Ralstonia solanacearum* infections.**
A) Tomato (*Solanum lycopersicum*) plants infected with *Ralstonia solanacearum* show wilting symptoms and left infected plant right: not-infected (Photography: Lauri Mikonranta at The University of York) B) Potato (*Solanum tuberosum*) showing early symptoms of potato brown rot disease caused by *Ralstonia solanacearum.* Oozing of the milky white bacterial colonies can be seen in the picture—The photograph was taken from the European Plant Protection Organisation (EPPO) website (https://gd.eppo.int/taxon/RALSSL/photos).

1.2.1. Agricultural importance

Ralstonia solanacearum species complex (RSSC) strains are important agricultural pests as they cause devastating plant diseases, often leading to characteristic wilting of the above-ground plant tissue and rotting of underground tubers or fruits (EPPO, 2018, p. 21; Zulperi et al., 2014) (Figure 2). *R. solanacearum* was ranked 2nd in the top 10 most scientifically and economically important bacterial plant pathogens in 2012 (Mansfield et al., 2012). Similar to other plant pathogens, *Ralstonia solanacearum* species complex has the greatest impact on the world's tropical and subtropical climate zones (Peeters et al., 2013b). However, there are recorded cases in multiple countries worldwide, and it is assumed that the bacterium is present in the soils everywhere around the tropical and temperate regions. The European and Mediterranean Plant Protection agency continuously updates information on recorded infections worldwide and the number of confirmed plant hosts. The current number of

countries in their database where *R. solanacearum* is considered present is 77 (12/10/2022) (Figure 3). In Southeast Asia, RSSC strains cause wilt on banana plants, leading to substantial economic losses. However, infections, especially in potatoes, are also present in temperate regions in North America and Europe. Therefore all RSSC strains are listed as quarantine pathogens by EPPO within the European Union because of significant potato crop yield losses (Bragard et al., 2019). *Ralstonia solanacearum* is a pathogen without adequate chemical

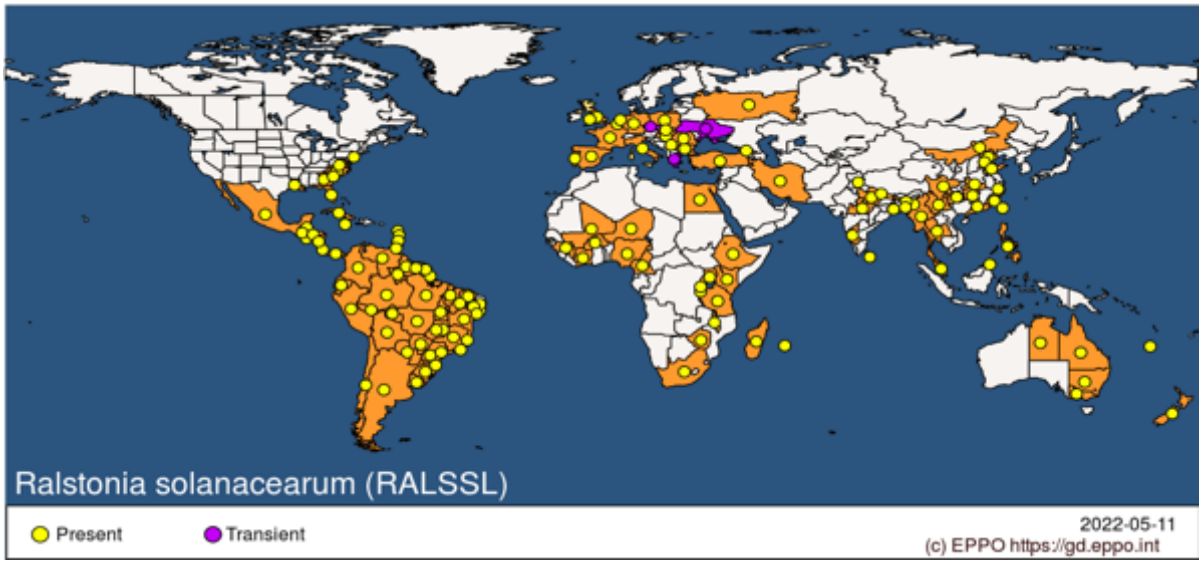


Figure 3. ***Ralstonia solanacearum sensu lato* distribution.**
Map of the distribution of *Ralstonia solanacearum* species complex recorded cases according to the European and Mediterranean Plant Protection Organization (EPPO). Present means the pathogen is an immediate risk to agriculture, and transient means it is being controlled effectively.

control, and tremendous efforts are required to control and prevent its spread. This includes controlling and monitoring contaminated fields and water systems and rotations with resistant crops such as strawberries (Yuliar et al., 2015). Unfortunately, there is no accurate and up-to-date account of the worldwide economic losses caused by the RSSC. Still, in the early 2000s, bacterial wilt of potatoes was estimated to affect 3.75 million acres of agricultural land in 80 countries, with global damage costs exceeding $950 million/year (Yamada, 2016).

1.2.2. Plant hosts

RSSC strains cause disease on a variety of plants, including essential food crops such as potatoes, fruit crops (banana, tomato), oilseed crops (sunflower, groundnut), spices, fodder, forest trees (ironwood and eucalyptus), and many ornamentals such as rose (Genin and Denny, 2012a). The disease is often fatal to the host and can have a devastating effect on the agricultural harvest. The host range of the RSSC strains is vast and not well defined, but it is thought to include over 200 plant species within at least 50 different families (Hayward, 1991a). Moreover, the host range is ever-expanding, with new hosts being added to the list every year. The most hosts are dicotyledons belonging to Solanaceae family, such as potatoes, leading to brown rotting of the tubers, (Cellier et al., 2012). However, there are also a few monocot hosts, such as ginger, or members of the *Musa* genus like banana, in which RSSC is responsible for the Moko disease (Hayward, 1964). The ability to infect both monocots and dicots is unusual and highlights the diversity within the *Ralstonia solanacearum* species complex (Genin and Denny, 2012a; Zulperi et al., 2014). Moreover, *Ralstonia solanacearum* was recently reported to be infectious in the plant pathogenic fungus *Fusarium oxysporum* (Tsumori et al., 2022). In theory infecting fungi should require a whole array of different genes as the immunity of a fungus is very different from a plant. Therefore, these examples show the tremendous genetic diversity of this pathogen, allowing infections outside the plant kingdom. Overall, the genes responsible for pathogenicity in different hosts are assumed to be highly variable, and in the past, studies with potatoes and bananas have not found genes associated with pathogenicity in these hosts (Cellier et al., 2012; Guidot et al., 2009). Therefore, establishing precise classification within the species complex and experimental validations of infectivity will be of great value for correctly evaluating this pathogen's host diversity.

1.2.3. Taxonomy

Ralstonia solanacearum's taxonomy and classification have been through multiple revisions before reaching its status of a species complex, consisting of three phylogenetically distinct species: *Ralstonia solanacearum, Ralstonia pseudosolanacearum* and *Ralstonia syzygii* (Safni et al., 2014a). The pathogen was first described as *Burkholderia solanacearum* in 1896, later

renamed *Pseudomonas solanacearum*, until finally being assigned to the *Ralstonia* genus. Multi-locus sequence typing (MLST) and whole genome comparative studies agree with the current separation of the species complex into three species, each composed of divergent and deep-rooted monophyletic phylotypes (Guidot et al., 2009; Remenant et al., 2010). However, the historical RSSC classification splits the species complex into phylotypes based on the phylogeny of the sequences of few genes: internal transcribed spacer (ITS) region, the *hypersensitive response and pathogenesis B* (*hrpB*) and *endoglucanase* (*egl*), which are still used widely. (Mark Fegan and Prior, 2005). There are four main phylotypes (I-IV), and phylotype II is further subdivided into IIA and IIB. Currently, phylotypes IIA and IIB make up *Ralstonia solanacearum* species *sensu stricto*, phylotypes I and III are *Ralstonia pseudosolanacearum*, and phylotype IV is *Ralstonia syzygii* which has three sub-species: *R. syzygii subs. syzygii, R. syzygii subsp. indonesiensis, R. syzygii. celebesensis.* According to phylogenetic coalescence analysis studies, the four phylotypes have distinct geographical origins: Phylotype I - Asia; phylotype II - the Americas; phylotype III - Africa; and phylotype IV – Indonesia (Wicker et al., 2012a). However, the current geographical distribution of the individual species and strains within the complex is not well defined. Phylotypes I-II extend much further than their place of origin, with strains that can be found across six continents. There is also an early phylogeny-based method to divide the strains within each phylotype; strains can be subdivided into multiple sequevars based on the similarity of a 750-bp fragment of a single DNA marker - the endoglucanase (*egl*) gene, and strains with similar sequences are assigned to the same sequence variant group (sequevar) (Mark Fegan and Prior, 2005)**.** To further complicate the RSSC taxonomy, old nomenclature based on biochemical assays regarding carbon utilisation and a limited number of host infectivity assays where the strains are split into biovars and races, respectively, is still sometimes used; however, this system does not reflect the phylogenetic relationship between the strains and many of the biovars and races are polyphyletic (Mark Fegan and Prior, 2005). The long history of research into *Ralstonia solanacearum* species complex has led to a complicated taxonomic picture. However, the relevance of the species complex term is still useful for plant pathology as the disease symptoms observed on plants by all strains are highly similar (Sharma et al., 2021a). Moreover, all *R. solanacearum* species complex strains from the four phylotypes can infect tomato plants and cause the same disease symptoms (Remenant et al., 2010). Therefore, the

species complex and the separate phylogenetic species within it are both commonly used in literature today.

1.2.4. Disease aetiology

The diverse *Ralstonia solanacearum* species complex strains are unified by their common disease aetiology. All strains cause some form of bacterial wilt disease on plants characterised by colonisation of the plant xylem, where the bacteria reach high cell densities (10^9 – 10^{10} CFU/ml xylem fluid) (Buddenhagen and Kelman, 1964). The exopolysaccharides (EPS) produced by the cells block the water supply of the plant and cause it to wilt and often tubers or fruits to rot (Denny, 2006). Symptoms include stunting, browning of the xylem, chlorosis, wilting, and often rapid death of the plant host. Bacterial cells spread through the soil and enter the plant vasculature through the roots. After reaching the xylem, the cells form biofilms increasing the viscosity of the xylem liquid and leading to the drying out of the plant (Figure 4).

Ralstonia solanacearum spreads through the soil, where water and chemical signals help guide the bacterial cells' movement in the soil towards a susceptible host using chemotactic polar flagella (Corral et al., 2020). Energy taxis contribute to the ability of *R. solanacearum* to locate and interact with its host plants, and the cells need aerotaxis for normal biofilm formation (Yao and Allen, 2007). Impairment in swimming motility leads to decreased virulence which is one of the reasons RSSC strains are thought to be incapable of surviving in arid conditions (Tans-Kersten et al., 2001). After encountering the host, the bacterium uses root hairs and lesions at secondary root emergence or elongation sites to aid infection. It colonises the plant's root cortex, evading the plant host immunity and utilising myriad excreted effectors and, notably, the type III effector family (T3Es) of proteins (Coll and Valls, 2013). Subsequently, the bacteria move to the xylem, quickly spreading through the vasculature to the plant's leaves (Lu et al., 2018). The wilting symptoms arise as the bacteria invade the plant xylem, proliferate, and form biofilms excreting exopolysaccharides (EPS) which increase the viscosity of the xylem fluid leading to blockage of the vasculature (Figure 4).

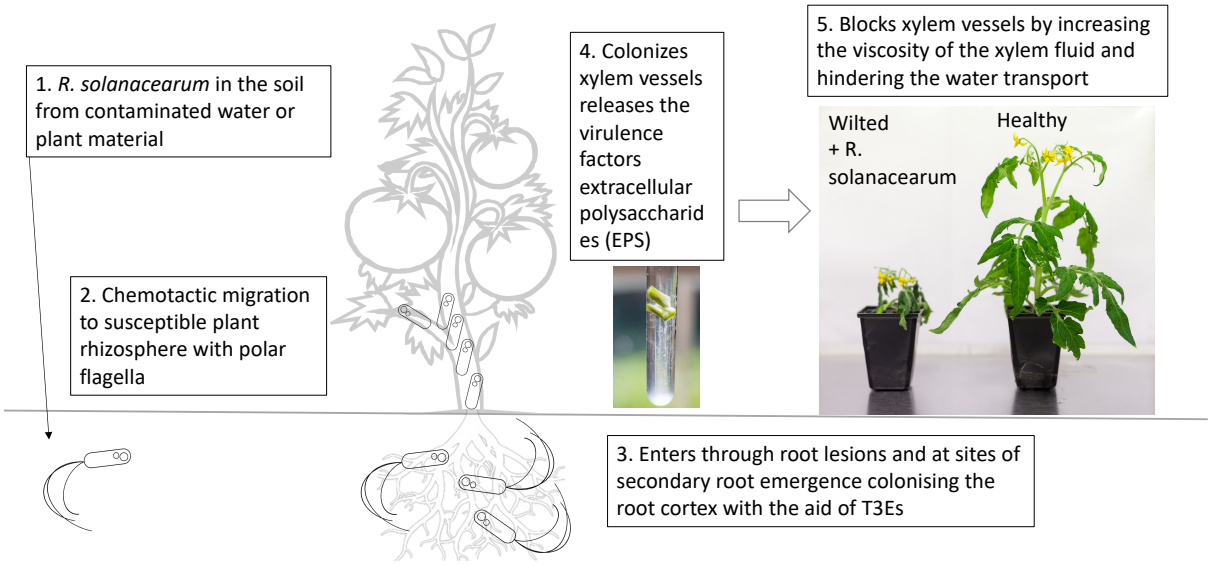


Figure 4. **Schematic of plant infection by *Ralstonia solanacearum***

1. *Ralstonia solanacearum* bacteria reside in the soil. 2. Cells sense signals from susceptible plants and move to the rhizosphere of the host using chemotaxis. 3. Cells enter the plant through root lesions or sites of secondary root emergence. After successfully evading the plant's immune system defences they colonise the roots. 4. The bacteria migrate to the water-conducting xylem tissue where they settle form biofilms and excrete exopolysaccharides (EPS) that can block the vasculature. The photograph in the middle shows white liquid due to EPS produced by RSSC cells oozing out of an infected plant's xylem and collecting at the bottom of a tube full of water. 5. After RSSC has colonised the whole plant and blocked its water-conductive tissues the plant wilts and dies. The photograph on the right shows a tomato plant infected with Ralstonia solanacearum (left) and a healthy one without bacteria (right).

1.2.5. Pathogenomics and virulence

Many cellular processes and functions that contribute to virulence and pathogenicity in *Ralstonia solanacearum* species complex have been characterised. Like in many other gram-negative pathogens Type III secretion system (T3SS) is a central component of the RSSC

virulence (Coll and Valls, 2013). Also, the quorum-sensing regulatory system controls the expression of the master regulator *Phc* and multiple downstream virulence factors (Ujita et al., 2019) (Figure 5). Overall, swimming, and twitching motility, biofilm formation, extracellular polysaccharides, plant cell wall-degrading enzymes, and Type I-IV secretion systems all play an essential role in the RSSC's master-controlled virulence system.

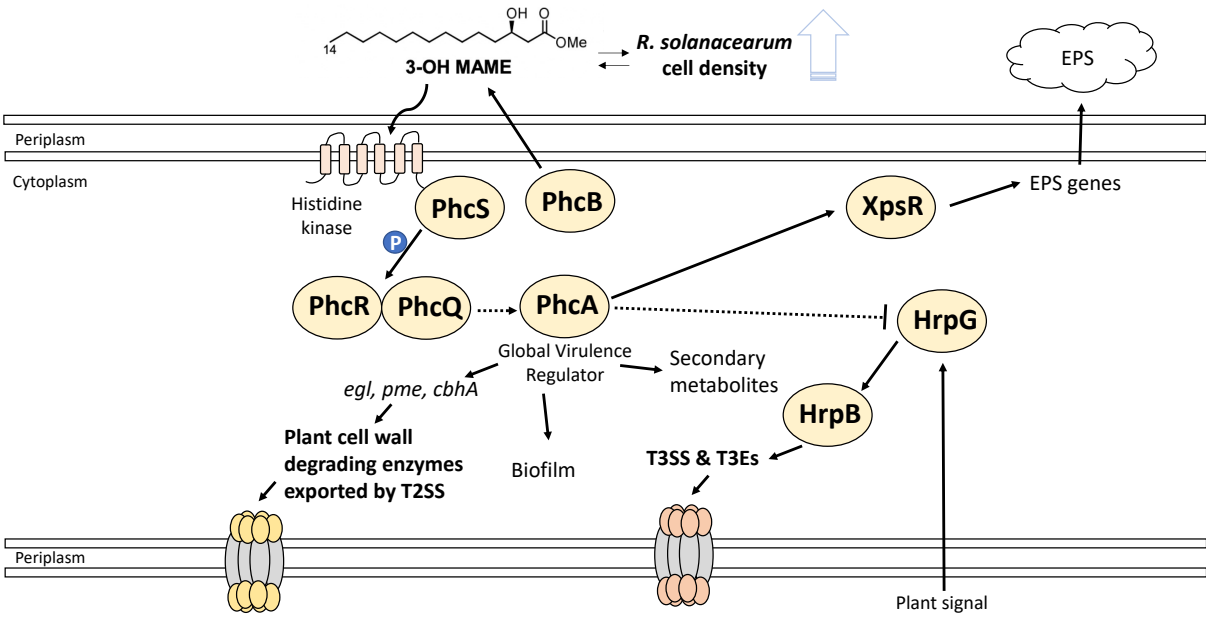


Figure 5**. Master virulence gene regulators PhcA and HrpG**

Simplified gene schematic showing the importance of PhcA and hrp genes in the virulence regulatory network of *Ralstonia solanacearum* species complex. The master regulator PhcA is activated by quorum sensing and then leads to activation of an array of virulence pathways including Type II secretion system (T2SS), biofilm formation, exopolysaccharide (EPS) production and the situational suppression of hrp genes that control Type III secretion system. Figure adapted from (Genin and Denny, 2012a; Kai et al., 2015).

The T3SS in *R. solanacearum* is functional throughout the disease development from infection to the end when the cells have colonised the xylem (Meena et al., 2019). The T3SS system enables the secretion of virulence factors and their injections into the host cell. The effectors modify the host cell environment by interacting with the immune system and altering the host's metabolism in favour of the pathogen. This molecular interaction is often highly specific

between a host and a pathogen. The RSSC type strain GMI1000 is known to form 60-80 type III effectors (Genin and Boucher, 2004). Type III effector (T3E) pangenome of RSSC was estimated to be 102 genes and 16 hypothetical genes (Sabbagh et al., 2019). This analysis included all known diversity in the RSSC strains and phenotypes from 155 genomes, but the variability of T3Es across the full pangenome of the species complex is probably greater. The diversity of effectors is correlated with the host range, the extent of which is yet to be fully characterised in RSSC. Due to the importance of the T3Es and their early discovery in RSSC, the effectors are also sometimes called Rips ("Ralstonia injected Proteins"). The gene regulatory cascade controlling T3SS has been well-characterised in RSSC strains. The activation of transcription is controlled by the two-component response regulator *HrpG* and the *hrp* gene cluster, which spans a 23-kb region on the Megaplasmid (Wu et al., 2015). The transcription of T3SS genes is induced in response to the bacterium-plant cell contact and negatively regulated by the *PhcA* master regulator pathway (Aldon et al., 2000; Genin and Denny, 2012a) (Figure 5).

Ralstonia solanacearum has a master regulator system of the virulence gene called *Phc* (phenotype conversion). The name comes from the small colony variants that can be observed in a bacterial culture which lack the production of exopolysaccharides (EPS) and thus appear small and lack milky-white exudates (Figure 2). The master regulator gene *PhcA*, a *LysR*-type transcriptional regulator, is the most important in activating and deactivating the virulence of the bacterial cell (Figure 5). *PhcA* is regulated by a unique volatile 3-OH MAME that acts as a quorum-sensing signal, activating the cascade when 10^7 CFU/ml cell density is reached (Flavier et al., 1997). Therefore, bacterial cell–cell communication is used to regulate the production of exopolysaccharides (EPS) and degradative enzymes such as endoglucanases (*egl*) and polyglucanases (Roberts et al., 1988). Mutants can spontaneously appear in culture and not produce EPS (Poussier et al., 2003). The mutants, however, have increased motility and siderophore production, which is needed for out-of-host survival of RSSC strains in the soil rhizosphere (Kai et al., 2015; Perrier et al., 2018). Motility, another virulence-associated trait in RSSC, is positively regulated by *PehSR*, which also positively regulates plant cell wall degradation and is ultimately regulated by *PhcA* (Tans-Kersten et al., 2001).

In addition, there is a link between virulence and metabolic capacity in *Ralstonia solanacearum.* The trade-off between growth and virulence is well-known and predicted by the evolutionary dynamics theory (Polz and Cordero, 2016). This negative correlation was experimentally shown in *R. solanacearum* mutants with progressive loss of virulence (mutant *eps*, mutant *xpsR*, and mutant *phcA*), corresponding to additive increases in their growth rates (Peyraud et al., 2016). The production of Type III Effectors is expensive; thus, they need to be tightly regulated so that the trade-off between proliferation and defence against immunity can be managed during infection. The nutrient-rich xylem niche RSSC occupies must sustain this complex behaviour (Baroukh et al., 2022).

1.2.6. Transmission and spread

The initial spread of the *R. solanacearum* species complex worldwide and its differentiation into three species is thought to have happened before the continents' split and Gondwana's fragmentation (Wicker et al., 2012a). *R. solanacearum* species complex is believed to have originated from Indonesia, where currently all strains of *R. syzygii* (phylotype IV) are found (Remenant et al., 2011). It is believed that one sub-group from that region migrated to Madagascar and East Africa and gave rise to *Ralstonia pseudosolanacearum* (phylotypes I and III), and another sub-group to have spread to Brazil and South America and given rise to *Ralstonia solanacearum* (phylotype II) (Wicker et al., 2012a). As more genome data becomes available for the whole genome of RSSC strains, coalescence dating will give more accurate estimates of the spread of the bacteria, the date of the root, and the most recent common ancestor of these Solanaceae infecting pathogens.

Chemotaxis makes it relatively easy for *Ralstonia solanacearum* species complex strains to spread rapidly across an agricultural field. RSSC strains can survive in the soil for years as a saprophytic bacterium waiting for favourable host signals and conditions (Van Elsas et al., 2000). However, the spread of the pathogen across different fields and countries is not currently understood well and is thought to be generally associated with humans and agricultural practices. In the temperate regions, overwintering cold temperature infective phylotypes I and II are believed to rely on infection of perennial reservoir host plants because its persistence in the soil is limited, and detection from the soil in temperate climates is rare.

One known means of transmission for the phylotype IIB-1 (Race 3 Biovar 2) strain known as the “potato race”, causing brown rot disease in temperate regions, is infected seed potatoes. They do not show symptoms at colder temperatures, but the bacteria proliferate when potatoes are sown in the Summer when temperatures are higher and favourable for the pathogen. In the United Kingdom specifically, contaminated river water used for irrigation has led to multiple outbreaks (Elphinstone et al., 1998b, 1998a; Tomlinson et al., 2009). Overall, little is known about the role of asymptomatic hosts in the spread of RSSC strains around the world. However, it is thought to be important in the UK, where there is a recognised but not well-understood link between the riparian plant *Solanum dulcamara,* which can sustain an asymptomatic infection, and the persistence of the *Ralstonia solanacearum* phylotype IIB-1 (Race 3 Biovar 2) strain in the river water system (Elphinstone and Matthews-Berry, 2017). *Solanum dulcamara* is a woody plant from the Solanaceae family that lives along riverbanks, and it is thought that it can provide refuge for *R. solanacearum* during the winter in the UK. In summer, the cells can be detected in the river water, and the water used for irrigation of potato fields has led to several outbreaks.

1.3. THESIS AIMS AND OBJECTIVES

The overarching aim of this project was to analyse the genetic diversity in *Ralstonia solanacearum* species complex based on whole genome sequences. In this research, I aimed to answer the following questions:

1. **What is the global pangenome diversity in the *Ralstonia solanacearum* species complex, and what is the size of the core genome of the complex?**

To study this, I constructed a pangenome of the *Ralstonia solanacearum* species complex based on a worldwide collection of bacteria obtained from the Fera Ltd. collection. I sequenced 384 bacterial isolates representing the whole distribution and all the known species of the complex. I aimed to obtain a comprehensive overview of all the gene diversity in the species complex and construct the pangenome with the latest sequencing technology and software.

2. **What is the genomic diversity of *Ralstonia solanacearum* phylotype IIB-1 or the "potato race" in the United Kingdom?**

I constructed a pangenome of *Ralstonia solanacearum* from the UK to study this. I studied a subset of 170 isolates from the species complex wide sample set obtained from Fera Ltd collection. The samples originated from a 30-years long annual river water survey performed by Defra in the UK to screen for contamination and prevent the spread to potato crop fields through irrigation. I studied this time-series genome sample by analysing the fine-scale single nucleotide changes that occur through time in the pangenome and the accessory genome variation in this temperate river water environment.

3. **What is the diversity of *Ralstonia solanacearum* within agricultural fields in China?**

In this study, I aimed to investigate the spatial distribution of genetic and phenotypic diversity of 96 *Ralstonia pseudosolanacearum* isolates from four tomato fields in four different provinces in China. Specifically, I asked if each tomato field harbours a unique genetic strain or if strain genotypes are mixing between fields for example due to human-mediated transmission. To achieve this, I constructed a phylogeny based on the whole genome sequences of the 96 isolates and compared the presence/absence of important virulence-associated genes such as type III effectors and prophages. In addition, we looked at missense mutations in genes associated with key virulence pathways such as quorum sensing, motility, exopolysaccharide production., etc.

1.4. THESIS CHAPTER OUTLINE

This thesis includes the following chapters, presented in the form of research papers:

2nd Chapter: *Ralstonia solanacearum* species complex pangenome

In this chapter, we compared the genomes of 384 *Ralstonia solanacearum* species complex (RSSC) strains from 55 countries including all four major phylogenetic groupings of the species complex. We found 18,080 gene clusters in the pangenome and estimated a core genome size of 1,704 genes. In the accessory genome network, we found that there are 210 separate components of genes that are cooccurring together. These genes included important

virulence and defence genes and in 10% of the components, there were transposases. We also showed that recombination happened five times more often within the diversification of the RSCC phylogeny than mutation.

3rd Chapter: Three decades of survival in environmental reservoirs within the UK led to little genetic diversification within *Ralstonia solanacearum*

Following Chapter II, we investigated the genetic variation in time within a subset of 170 *Ralstonia solanacearum* species complex genomes originating from the United Kingdom. The population was highly clonal and belonged to the strain phylotype IIB-1 with mutation rates as low as one nucleotide change per year. However, the accessory genome of the country-wide sample studied had 55 genes that varied in intermediate frequency. The variation was linked to transposable elements and genes in their proximity in the genome. Using molecular clock dating techniques, we show that the population in the UK originated between 1958 and 1988, just several years before the first recorded outbreak in 1992. In addition, microbial GWAS identified links between a gene associated with antimicrobial resistance in a *Bacillus* spp. to be associated with time. Time was correlated with the increase in frequency of this gene.

4th Chapter. Multiple *Ralstonia solanacearum* lineages coexist in agricultural monocultures in China

In this chapter, we studied the phenotypic and genetic diversity of *R. solanacearum* within and between four tomato fields in China using a combination of genomics, phenotyping and physicochemical metadata collected at the level of individual plants' rhizosphere. We show that the plant rhizosphere between the fields is associated with each field on the biological and physicochemical levels. Furthermore, comparative genomics of 96 isolates (24 x 4 fields) shows that there are two genotypes of *Ralstonia solanacearum* phylotype I cooccurring per field and in total 8 different genotypes across the fields. The cooccurring genotypes differ regarding virulence traits and genes associated with type three effector proteins, quorum sensing, motility and iron-scavenging siderophores. In addition, microbial GWAS identified links between siderophore production and the IS5 family transposase IS1420, between growth and type III effector protein, and between *in planta* virulence and a RidA family protein.

2. *Ralstonia solanacearum* species complex pangenome

2.1. ABSTRACT

Ralstonia solanacearum species complex (RSSC) is a group of phylogenetically and phenotypically diverse plant pathogenic bacteria. The strains within RSSC are found in multiple countries around the tropical and climatic zones and cause wilting and rot diseases on several economically important crops. Here we performed a pangenome study using 404 RSSC genomes, including three species from four phylotypes originating from 55 countries. We found 18,080 gene clusters in the open pangenome and estimated that the core genome consists of 1,704 genes. Across the RSSC phylogeny, we show that recombination has happened five times more often compared to mutation, indicative of its importance for RSSC adaptation. Moreover, we found 2,437 associating accessory genes within 210 independent components in the pangenome, which suggests multiple links or epistatic relationships between accessory genes. The genes within these components included type I-IV effectors, heavy metal resistance genes, prophage integrases and transcription factors. Moreover, we found that transposons accounted for 10% of the genes in the associating modules, which suggests that mobile genetic elements play an important role in the dissemination of genes related to virulence and environmental persistence in the RSSC pangenome. Together, these findings suggest that horizontal gene transfer is a major driver of RSSC accessory genome diversity, allowing it to adapt across a variety of ecological niches.

2.2. INTRODUCTION

Recombination, gene gain, and gene loss create variation in the gene content between individuals of the same species. This is the basis for the field of pangenomics studies (Golicz

et al., 2020), an extension of the idea of gene essentiality (Rancati et al., 2018). Bacterial species have a core genome shared by all or almost all the strains in the species and an accessory genome, which is shared by some strains and is beneficial only in certain environments or ecological niches. Pangenomics allow us to identify the frequency of genes of interest in a species, or even frequency of homologous genes in a whole database of gene sequences over wider taxonomy. This allows us to infer the essentiality of genes based on how universal or unique they are within certain environments. The pangenome of a species can be classified as open or closed, based on a gene accumulation curve. In other words, a pangenome is open if the number of new genes constantly increases when new genomes are included in the dataset. In contrast, if the number of genes plateaus after a certain threshold the pangenome is considered closed (Mira et al., 2010). However, this is a concept that can be biased by sampling and classification. Generally, species with larger and open pangenomes are expected to have wider distribution, more migration, and bigger long-term effective population sizes compared to species with closed pangenomes (McInerney et al., 2017). Pangenomics studies on model organisms such as *Eschericia coli* and *Pseudomonas aeruginosa*, have been very successful, for example, in identifying horizontal gene transfers, exploring accessory gene variation, and investigating outbreaks (Freschi et al., 2019; Tantoso et al., 2022). Pangenomics approaches have also been utilised with plant pathogenic bacteria. For example, in the cassava blight pathogen *Xanthomonas axonopodis,* sequencing a large number of diverse isolates helped to identify a core set of Type III effectors (Bart et al., 2012). Although pangenomics has become a rapidly expanding field of research, there is still a big gap in our understanding about how the immense gene diversity is maintained in bacterial populations.

Ralstonia solanacearum is an example of a bacterial species complex whose genome has been described as open (Geng et al., 2022). Due to historical reasons, the name *R. solanacearum* is still used to refer to a species complex of bacteria currently classified as three phylogenetic species: *R. solanacearum*, *R. pseudosolanacearum*, and *R. syzygii* (Safni et al., 2014a). In the past, the species complex was divided into phylotypes based on a few marker gene sequences which now roughly represent the species clades in the phylogeny: *R. solanacearum* is phylotype II, *R. pseudosolanacearum* phylotype I and III, and *R. syzygii* phylotype IV (Wicker et al., 2012a). In this *R. solanacearum* species complex (RSSC), all the strains are

phytopathogenic and can infect important crops such as tomatoes, potatoes, and bananas. The RSSC is notoriously diverse genetically and phenotypically with many strains capable of infecting multiple plant host species (Hayward, 1991a). Although RSSC consists of three species, the capability of infecting solanaceous plants is a dominating feature in the group, making the species complex a meaningful unit of bacteria that are a causal agent of wilting disease in plants (G Cellier and Prior, 2010). Thus, this capability is likely to be part of the core genome of RSSC and an ancestral trait that can be traced back to their most recent common ancestor. Alternatively, the presence of the same hosts within all three phylogenetic species clades could be indicative of convergent evolution where RSSC strains have independently evolved to infect tomato plants on three different occasions for the three clade split of the species phylogeny (G Cellier and Prior, 2010). Another interesting feature of RSSC is their ability to infect plants in warm and temperate climates. In general, RSSC strains prefer warmer temperatures and climatic zones, and all known strains can grow at 28 °C (Bocsanczy et al., 2014). However, *R. solanacearum* and *R. pseudosolanacearum* can be considered cosmopolitan as they include strains that can grow at 10 °C (Personal communication E. Farnham, (Bocsanczy et al., 2017; Caruso et al., 2005a)). In contrast, *R. syzygii* and its three subspecies have only been found in Indonesia where it causes diseases on several plants, including economically important banana. The *R. syzygii* subsp*. syzygii* causes Sumatra disease on clove and has an insect vector, *R. syzygii* subsp. *indonesiensis* is a broad host range pathogen, and *R. syzygii* subsp. *celebesensis* affects plants from the monocot genus Musa (Remenant et al., 2011; Safni et al., 2018). Collectively, the broad host range and width of occupied niche space highlights the RSSC ecological diversity, but how this reflects to the pangenome level is still unclear.

To an extent, the capacity of RSSC to thrive in a repertoire of hosts and environments can be accounted to its relatively large genome size (~5–6 Mb), which is made of two replicons: chromosome (~3.7 Mb) and megaplasmid (~2.1 Mb) (Genin and Boucher, 2004). The two replicons share evolutionary history, carry essential genes, and are always inherited via binary fission but are called the Chromosome and the Megaplasmid, rather than two chromosomes, for historical reasons (Genin and Denny, 2012a). Previous studies have shown great diversity within RSSC in virulence-associated genes such as the type 3 effectors (Sabbagh et al., 2019). These effectors play an extremely important role in the bacterial infection mechanisms and

in the recognition by the plant immune systems (Straub et al., 2021b). In addition to secretion system effectors, RSSC genomes are notably diverse in prophages (Greenrod et al., 2022a; Souza et al., 2021), phage defence systems (Castillo et al., 2020), insertion sequences (Gonçalves et al., 2020a; Greenrod et al., 2022b), which can also be indirectly linked with virulence. The relative importance of different genes have been studied using transposon insertion libraries, where effects of different genes can be assessed across different environments (Su et al., 2020). In one RSSC gene essentiality study, where a near-saturated transposon insertion library of the type-strain for RSSC GMI1000 was created, researchers identified 465 essential genes, 354 of which were previously identified in other bacteria, and 34 new genes potentially critical only for RSCC bacteria (Su et al., 2020). This number is much smaller than the previously estimated core genome size of the RSSC although to our knowledge no study has performed comprehensive core genome size estimates (Geng et al., 2022; Sharma et al., 2022). There are known differences in horizontal gene transfer rate and accessory gene content between the phylotypes (Sharma et al., 2022; Wicker et al., 2012a), for instance, virulence factors and mobile genetic elements that are specific to strains (Geng et al., 2022). Therefore, movement of genes within each species could be limited and there may be no horizontal gene transfer happening across the species complex due to genetic limitations or simply due to geographic separation.

In this study, we collated an extensive worldwide strain collection of RSSC species complex from Fera Ltd. cryo stock samples. We aimed to acquire isolates from all the sampled countries with a variety of hosts of isolation to get the wealth of diversity of the species complex. This resulted in a collection of 195 isolates from 55 countries and 189 isolates from a time series over 27 years in the United Kingdom for a total of 384 bacterial isolates. The genomes of these isolates were sequenced and after quality checks, 356 high quality genomes were kept for analysis of their pangenome. Moreover, we downloaded 46 complete genomes available in NCBI to represent published genome diversity of the species complex. Based on this extensive genome collection we constructed a pangenome for RSSC that consisted of 18, 080 gene clusters and estimated that the core genome consists of 1,704 genes. Furthermore, we found 210 associating components within an association network constructed based on the accessory genome. Moreover, 10% of these components had transposal elements within

them suggesting insertion sequence movement is associated with lineage independent accessory genes in RSSC.

2.3. METHODS

2.3.1. DATA

2.3.1.1. STRAIN COLLECTION

384 isolates belonging to the *Ralstonia solanacearum* species complex (RSSC) were chosen from the cryo collection at Fera Ltd. based on metadata available in archives. The whole set was confirmed to belong to the RSSC complex by real-time PCR protocol designed to identify isolates within any of the four phylotypes of the RSSC (Weller et al., 2000). All strains were preserved at -80 °C in a cryoprotectant system (Protect) at Fera Ltd. Some of the strains are part of the commercially available stocks within the National Collection of Plant Pathogenic Bacteria (NCPPB) and were available as freeze-dried cultures. Working cryo stock libraries of the isolates and genome database of RSSC were created at The University of York, and a YO number from 1-384 was given to the isolates (Appendix 1). The *R. solanacearum* species complex isolates were chosen to represent a UK and a World collection: 176 isolates belonged to the UK *R. solanacearum* population and 208 representing the known worldwide diversity across the RSCC.

2.3.1.2. SEQUENCING

The isolates were grown on SP and agar media overnight (EPPO, 2018; Lelliott and Stead, 1987), and a single colony was chosen for DNA extraction with Qiagen Blood and Tissue kit. The DNA quality was checked with Nanodrop and Quantit. The DNA was sent for Illumina MiSeq sequencing at the Earlham institute in the UK, and we received raw untrimmed paired FASTQ files for all 384 genomes. In addition, 24 of the isolates were re-sequenced with Nanopore MinIon technology at the University of York by the technological facility on site. The data received was base called by the technology facility staff using Guppy and raw FAST5 files, and raw FASTQ files were received. All the raw FASTQ files from the Illumina and Nanopore runs are publicly available at the NCBI SRA and metadata under project number PRJNA823737.

The raw FASTQ short read files were adapter trimmed with TrimGalore, and quality was checked with FASTQC and MultiQC (Ewels et al., 2016). *De-novo* assemblies were made using Unicycler (v 0.4.7) on either short read mode only or combined mode with the long reads for the 24 isolates available (Wick et al., 2017a). Contigs were filtered based on GC content (>=0.66<0.67) and minimum length (5Mb). As a result, 356 draft genome assemblies were deemed good quality out of the initial 384. Mash (v2.3) was used to align contigs against a local RefSeq installation to identify potential contaminants in the assemblies (Katz et al., 2019). In addition, 24 of these genomes were re-made by a hybrid assembly of MinIon and Illumina reads using Unicycler. Furthermore, 48 assemblies for *Ralstonia solanacearum*, *Ralstonia pseudosolanacearum* and *Ralstonia syzygii* were downloaded from NCBI GenBank FTP on 10/06/2020 along with 1 *Ralstonia pickettii* 12J representative genome strain for the species and type strain for the *Ralstonia* genus. The 47 genomes were chosen if they were labelled as *Ralstonia solanacearum* (taxonomy number 490) and had full and complete assemblies in NCBI (See Appendix 2 for accessions). All assemblies were annotated with prokka (v1.14.6) (Seemann, 2014a).

2.3.2. PANGENOME & PHYLOGENY

Panaroo (v1.3.0) (Tonkin-Hill et al., 2020) was used to construct a pangenome for 405 RSSC genomes (48 NCBI genomes + 356 from the University of York). Panaroo was run on default mode with the option "–strict" and without merging the paralogs and with the MUSCLE aligner. The phylogeny was constructed with IQ-tree and GTR+G model with two bootstrap methods, FastBoot and ALrT boot, from the core genome alignment produced by panaroo (Nguyen et al., 2015a). The phylogeny was dated using BactDating (Didelot et al., 2021). Pangenome size was estimated using a helper script with panaroo for The Infinitely Many Genes (IMG) (Collins and Higgs, 2012).

2.3.3. ASSOCIATION OF ACCESSORY GENES NETWORK

Panaroo was run a second time with the same command but a subset of highest quality genomes for further downstream analysis: 166 (47 from NCBI + 119 from the York Collection) high-quality genomes were chosen by removing most of the genomes that belong to phylotype IIB-1 clonal branch and leaving only the nanopore and Illumina hybrid assemblies

for this branch and removing genomes that were less than 4Mbp total length and were producing very short alignments after gap trimming. The core gene alignment was trimmed using trimAL and -gt 1 to remove the gaps from the alignment (Capella-Gutiérrez et al., 2009). The phylogeny was constructed using the same method as above. The output was used to put into Coinfinder (v1.2.0) with Bonferroni correction and a cut-off value for the D statistic of >=-0.25 (Whelan et al., 2020). Eggnog mapper was used with the DIAMOND algorithm (Buchfink et al., 2021) to annotate the significant accessory genes with KEGG and GO terms (Cantalapiedra et al., 2021; Jensen et al., 2008).

2.3.4. RECOMBINATION ANALYSIS

Recombination in the genome sequences was detected using 100 runs of prokaryotic recombination estimation software ClonalFrameML (Didelot and Wilson, 2015). Recombination calculations were done using the following formula for the ratio of the relative effect of recombination and mutation (r/m) for each simulation run:

$$\frac{r}{m} = \frac{R}{\theta} \times \delta \times \nu$$

$\frac{R}{\theta}$ - The ratio of recombination and mutation rates

δ - the mean length of imports

ν - the average distance of the imports

2.3.5. VARIANT CALLING

Snippy (v4.6.0) was used on all the 356 short-read genomes against type strain for the species complex GMI1000. Snipy-core accessory script available on the Snippy Github page was used to generate a core genome alignment and the recombination was cleaned from the alignment using Gubbins (v3.2.1) (Croucher et al., 2015). Using the BAMs from Snippy freebayes was run on an alignment of the best 153 genomes against the GMI1000 reference genome with the following command: freebayes --min-alternate-count 5 --min-coverage 20 --use-best-n-alleles

4 -q 30 -m 60 -L bams.list -p 1 . This produced a vcf with 500,118 variable sites which was further filtered for missingness allowing maximum of 40 individuals to have a missing site with vcfkit using the following command: vk filter MISSING --max=40. This left 374,706 core genome variants in the variant calling file.

2.3.6. GRAPHICS & DATA ANALYSIS

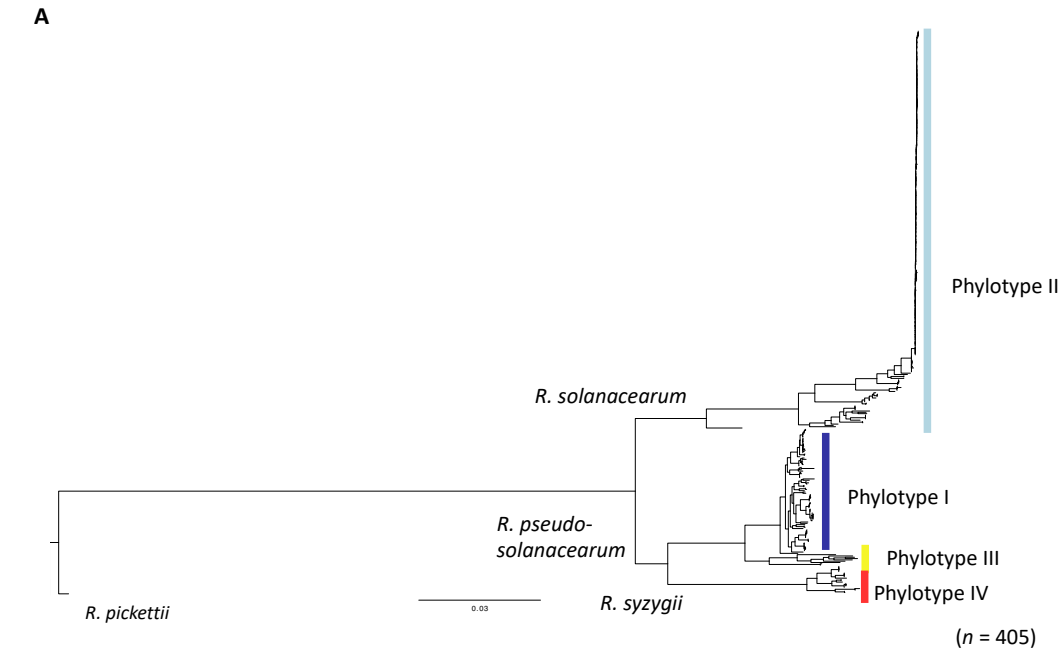
Additional graph plotting, tree visualisation and data analysis were done in R (R Core Team, 2017) with R studio (R Studio Team, 2020) and the packages ggtree (Yu et al., 2018) and ggplot (Wickham, 2009).

2.4. RESULTS

2.4.1. Phylogenetics of the *Ralstonia solanacearum* species complex

We ran a pangenome analysis with the pangenome software panaroo using 404 *Ralstonia solanacearum* species complex genomes and 1 *Ralstonia pickettii* 12J genome as an outgroup. The phylogeny was used to assign the *R. solanacearum* species complex genomes to a species and phylotype according to the phylogenetic clustering of the strains. To achieve this, NCBI genomes with known phylotype assignments were used as a reference for each phylotype and branch assignment as well as type strains present in our dataset from the NCPPB collection (Figure 6 A). The overall clustering of the phylogeny and the placement of the root for the species complex aligned with the previously observed phylotypes I to IV. However, the splitting of phylotype II, also known as *R. solanacearum* species *sensu stricto*, into the sub-phylotype IIA and IIB was not clearly observed, and for this reason we assigned the strains belonging to this branch as simply phylotype II (Figure 6 A). We show the distribution of the assigned phylotypes in our dataset across the world map based on the country they originated from. The map shows the clear bias in our dataset towards phylotype II and the overall very small number of isolates from phylotypes III and IV. Moreover, most of the isolates in the dataset (n=250) were placed within a clonal branch of phylotype II known as the “potato race” or phylotype IIB-1. The isolates within the clonal branch were from South America, Africa and Europe but had hardly any genetic diversity among them. Therefore, our dataset represents

a good global distribution of the RSSC strains but fails at representing replication and diversity within locations and especially diversity within different locations in in time.



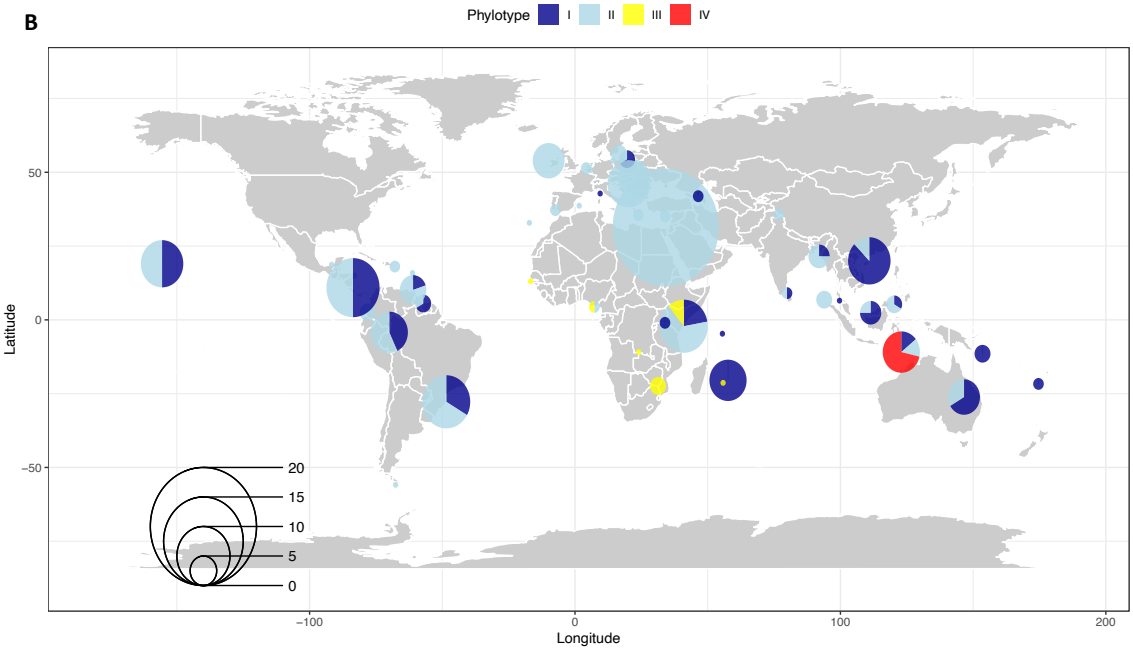


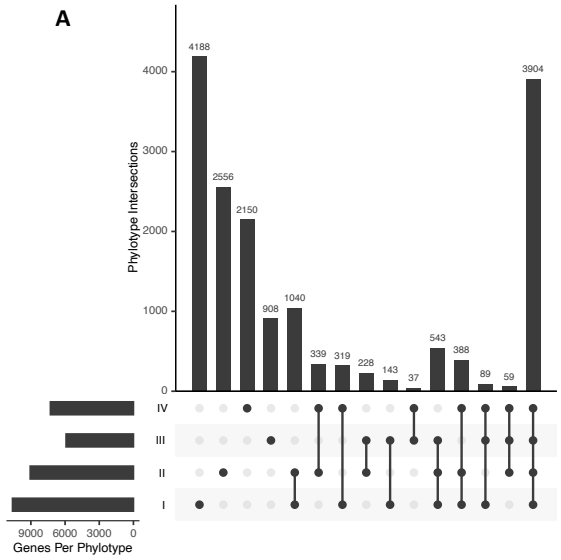
Figure 6. **Core genome phylogeny of *Ralstonia solanacearum* species complex.**
A. Phylogeny highlighting the UK isolates belonging to clonal strain phylotype IIB-1 (race 3 biovar 2). Maximum Likelihood (ML) phylogeny based on 405 RSSC genomes: 332 draft short read data only assemblies + 24 hybrid long read and short read assemblies + 48 *Ralstonia solanacearum* species complex NCBI genomes and *Ralstonia pickettii* 12J used as an outgroup. B. Map showing the country origin of the samples within a known country of isolation and split by phylotype assignment. The isolates originating from the UK were removed for ease of visibility.

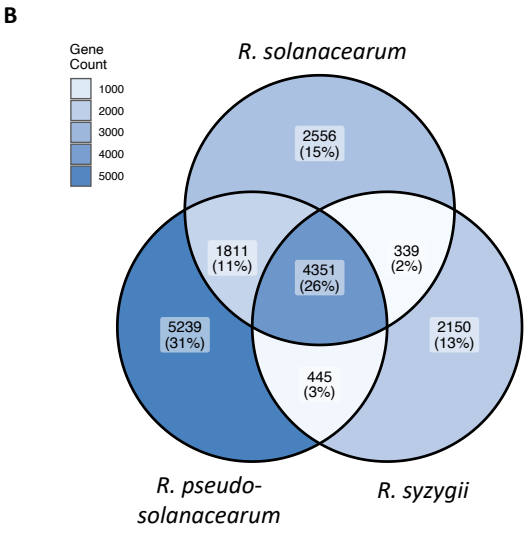
2.4.2. Pangenome size of the *Ralstonia solanacearum* species complex

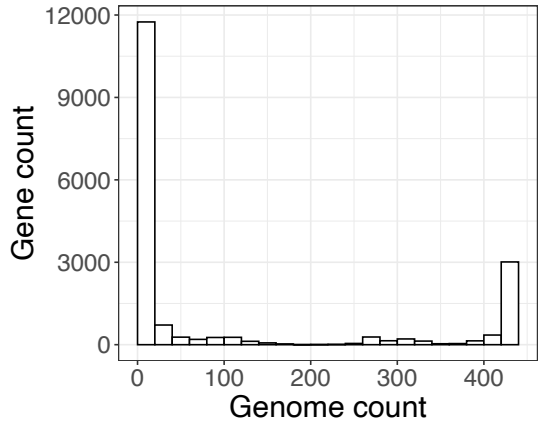
We estimated the size of the pangenome using the output from panaroo and the IMG model implementation provided as an accessory script. The total number of gene clusters found in the pangenome was 18,080. The core genome was defined as gene clusters shared by 90% of the genomes in the sample set and equalled 19% of the total pangenome clusters (n = 3,463). The rest of the gene clusters shared by less than 90% of the genomes made up the accessory genome consisting of 14,617 gene clusters. We observed large numbers of accessory genes unique to each species and phylotype in the complex (Figure 7 **A**&**B**). Moreover, the majority of the gene clusters were at low frequency as 63% of the total gene clusters were found only in 1-20 genomes (Figure 7 **C**). This result was unsurprising due to the large diversity of strains and location of isolation in the dataset and especially because we lacked replication within the locations studied apart from within the UK (Figure 6 **B**). Moreover, a total of 1,324 gene clusters or seven percent of the RSSC pangenome were annotated as transposons or transposition-related genes, which highlights the potential importance of mobile genetic elements for the evolution of RSSC pangenome.

To estimate the size of the pangenome we dated the phylogeny with the Bayesian molecular clock model estimator Bactdating and ran the IMG model provided as a helper script in panaroo. This analysis was performed to correct for an ever-decreasing core genome issue encountered with simple rarefaction curve estimates. The IMG model requites a dated

phylogeny. However, the parameters of the Bayesian model used for dating the phylogeny here did not reach sufficient sampling. Therefore, the results of the pangenome size analysis will be affected by the lack of accurate dating but we still deemed the model approach better than the simple rarefaction curve analysis. The result is a core genome size estimate of 1705 gene clusters (Figure 7 **D**). Overall, the core genome size did not change much with the addition of genomes after 100 showing that the core genome size remains fairly constant in the RSSC.


A
Phylotype Intersections
4188
2556
2150
908
1040
339
319
228
143
37
543
388
89
59
3904
IV
III
II
I
9000 6000 3000 0
Genes Per Phylotype


B
Gene Count
1000
2000
3000
4000
5000
R. solanacearum
2556 (15%)
1811 (11%)
4351 (26%)
339 (2%)
5239 (31%)
2150 (13%)
445 (3%)
R. pseudo-solanacearum
R. syzygii


C
Gene count
12000
9000
6000
3000
0
0
100
200
300
400
Genome count

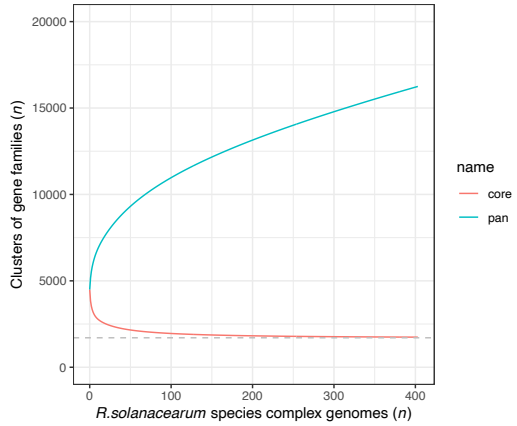

D
Clusters of gene families (n)
20000
15000
10000
5000
0
0
100
200
300
400
R.solanacearum species complex genomes (n)
name
core
pan

Figure 7. **Pangenome diversity and size**

The pangenome of the 431 genomes used in the pangenome analysis is broken down in frequency groups. The pangenome (total number of genes) = 18,080. The core genome (genes shared by 90% of the isolates) = 3,463 genes. The accessory genome (genes shared by less than 90% of the genomes) = 14,617. A) Gene clusters per phylotype. Combinations of gene cluster size depending on grouping of phylotypes shown as bar size and number on top of the bar. B) Gene clusters per species; C) The number of genomes that carry each gene cluster in the pangenome. x-axis = Number of genomes with a gene present.; y-axis = Number of occurrences of a certain number of genes present in the genome set (population). The number of bins in the histogram was determined by Sturge's rule. D) Rarefaction curve of the pangenome. The dashed grey line shows the size of the core genome for RSSC based on IMG model = 1705. Solid lines show the increase in gene cluster number as genomes are added and the effect on core size in red and pangenome size in blue.

2.4.3. Recombination has happened five times more often than mutation in the *Ralstonia solanacearum* species complex phylogeny

To investigate the source of the sequence diversity within our genome sample set, we estimated the homologous recombination rates relative to mutation rates using the whole genome alignment of the genomes. We used ClonalFrameML software as it can estimate recombination from outside the sample set of genomes studied. We ran the model for the whole species complex phylogeny to get the overall estimate of the relative effect of recombination and mutation for the diversification of the *Ralstonia solanacearum* species complex (Supplementary Figure 2).

We saw a total number of 1,290 predicted recombination events in the alignment. According to the bootstrapped whole phylogeny model the ratio of the relative effect of recombination and mutation was predicted to be r/m = 5.18 [5.07 - 5.32] and the ratio of the frequency of recombination and mutation R/theta = 0.02 [0.02 - 0.02]. The average length of recombined fragments was estimated to be δ = 1,864 [1,832 – 1,894] bp and the average divergence

between donor and recipient was ν = 0.12 [0.12-0.12]. This result shows that mutation happened five times more often than recombination in the diversification of the *Ralstonia solanacearum* species complex, however, recombination contributed on average over 200 base pairs (δν = 221 bp) of sequence which is a lot longer sequence change compared to point mutation. Therefore, recombination could transfer whole or parts of genes and change the evolutionary trajectories faster than mutation.

2.4.4. Coincident gene relationship in the *Ralstonia solanacearum* species complex accessory genome

For the associating and dissociation analysis genes, we excluded genomes shorter than 3.5 Mbp and kept only a few isolates from the phylotype IIB-1 clonal branch as the software used Coinfinder cannot deal with clonal branches in a phylogeny. This left 166 high-quality genomes (118 from York Collection and 48 from NCBI) for the cooccurrence analysis. However, this only reduced the total number of genes present in the pangenome by 192 gene clusters (from 18,080 to 17,888) (Supplementary Figure 1).

In the software Coinfinder, coincident gene relationships are expected within the accessory genome, where the presence or absence of one gene is influenced by the presence or absence of another (Whelan et al., 2020). Using Coinfinder, we identified a gene co-occurrence network that consisted of 2,437 nodes joined by a total of 68,028 (Figure 8) and an avoidance network composed of 1,998 nodes joined by 222,571 edges (Figure 9). Interestingly, the co-occurrence network is more extensive in terms of the number of nodes but smaller in terms of the number of interactions compared to the dissociation network. This could be due to the small number of closely related strains present within each species in the phylogeny, creating very few connections between the genes from different clades that are not in the core genome that will have the required D value. The co-occurrence network has 210 components (clusters of nodes), and the number of components in the avoidance network is 6. The co-occurrence network had a lot of small components and the avoidance network had very few with most of the nodes (n=1,937) belonging to a single component (Supplementary Figure 3).

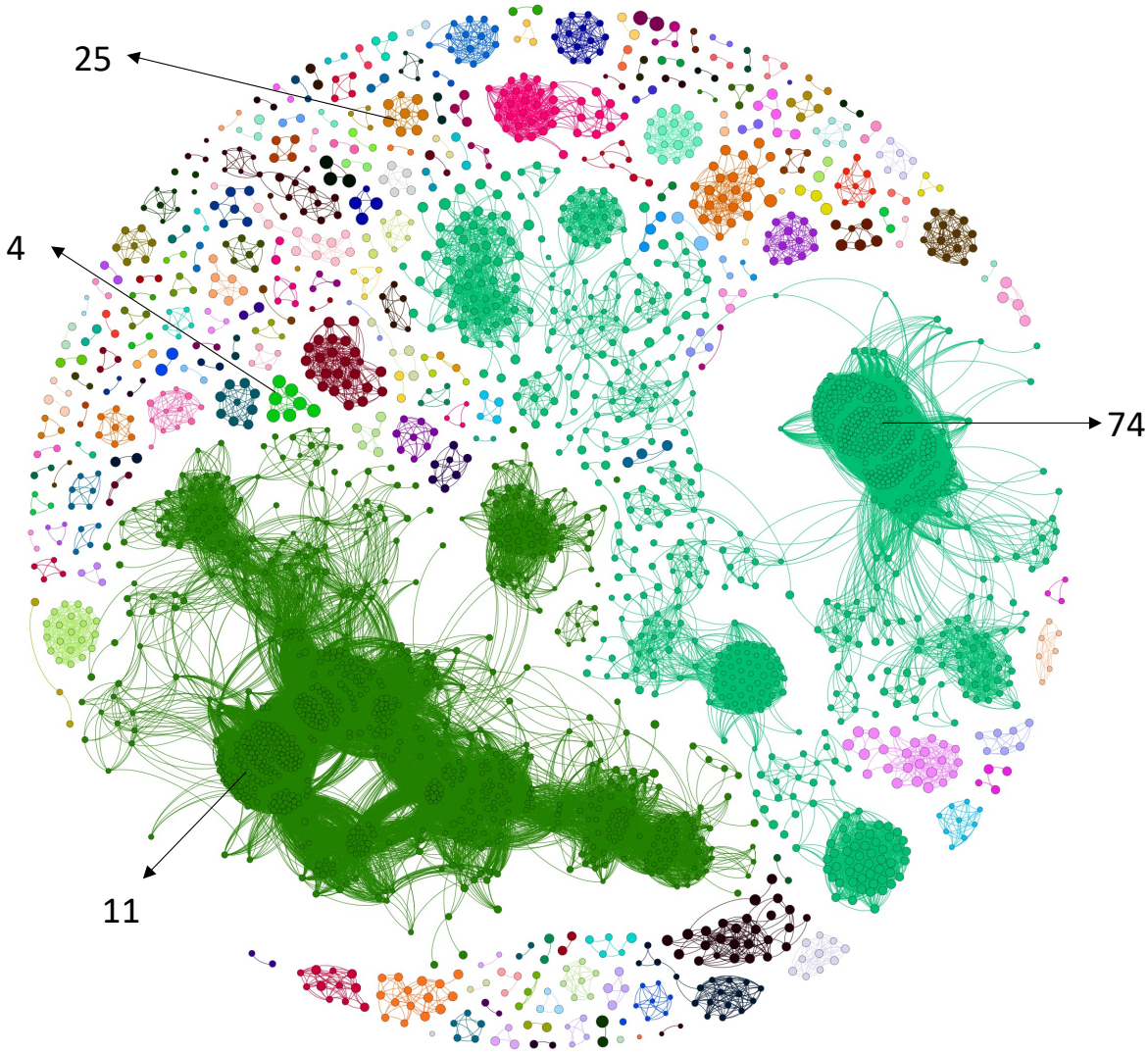


Figure 8. **Association (cooccurance) network of associating accessory genes**

Each gene cluster from the pangenome analysis is shown as an individual node. The edges that connect each node indicates a significant gene relationship between two nodes after Bonferroni correction and D-value filtering. The nodes are coloured by connected component (cluster), defined as a group of genes that form relationships with one another and not with the rest of the network. The size of the nodes indicates the D-value (large means larger D-value or more lineage independence).

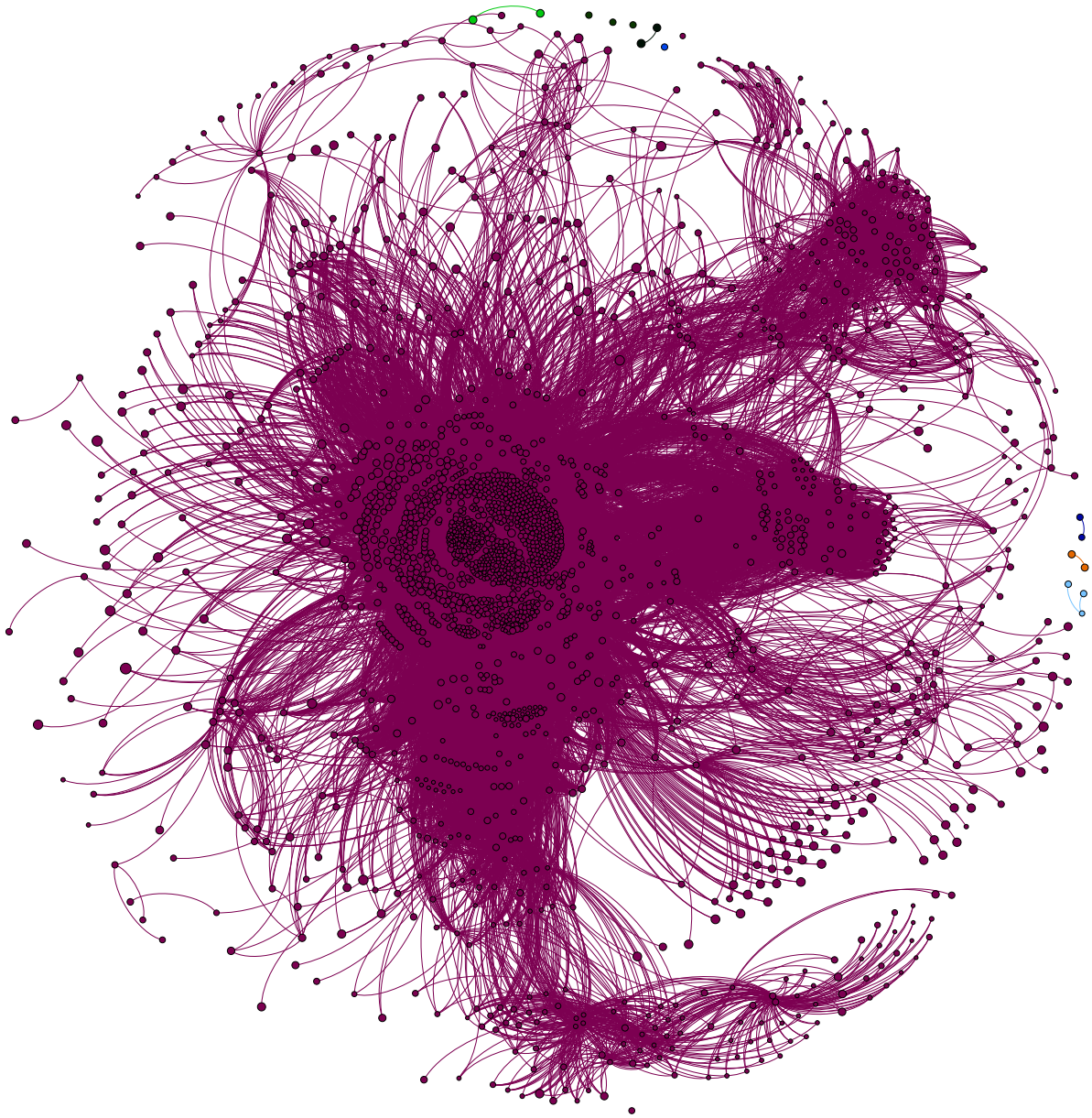


Figure 9. **Dissociation (avoidance) network of associating accessory genes**
Each gene cluster from the pangenome analysis is shown as an individual node. The edges that connect each node indicate a significant gene relationship between two nodes after Bonferroni correction and D-value filtering. The nodes are coloured by connected components (clusters), defined as a group of genes that form relationships with one another and not with the rest of the network. The size of the nodes indicates the D-value (large means larger D-value or more lineage independence).

2.4.5. Coincident gene relationship in the *Ralstonia solanacearum* species complex accessory genome associated with virulence and mobile genetic elements

The co-occurrence gene network is based on the idea that components of interacting genes share functional interactions and dependencies and therefore we report some known and potentially novel gene interactions within the components we found. Out of the 573 unique annotation terms, we found the most common term was “hypothetical protein” found 3,168 times across the association network. Most other terms were found in low frequency across the network with less than five occurrences per annotation term (See Appendix 3 for all components and gene annotations within them). Apart from unknown function proteins transposases take the lead with 12 transposases present in the most frequent terms (>=5 occurrences across components) in the association network (Figure 10 (2)). Among the most frequent annotations were also porins (Figure 10 (1)) and HTH-type transcriptional regulators (Figure 10 (3)). Other terms found in lower frequencies in the network were drug resistance, type I, II, III, IV system protein, nickel, cobalt and copper resistance, flagellar synthesis associated proteins, prophages, many transcription regulators and DNA-binding proteins.

We found multiple type III and II effector genes within the same components. Component number 86 consisted of ten gene clusters four of them annotated to be within the type II secretion system (*epsF, epsG, epsE* and *xcp*T) and one within the type III secretion system (*sctC*) (Supplementary Table 1). The production of exopolysaccharides (EPS) is a major component of the RSSC strains' virulence and contributes to plant host recognition and defence system activation. Avirulent variants do not produce EPS and have been found to be resistant to some phages (Milling et al., 2011). It has been shown that spontaneous avirulent mutants appear in *Ralstonia solanacearum* cultures that cannot produce EPS. This suggests a cluster of EPS production-related genes can move around the genomes of the *Ralstonia solanacearum* species complex.

Transposable elements are known to contribute to *Ralstonia solanacearum* species' complex genome plasticity (Gonçalves et al., 2020b). We found that transposases were the most frequently found annotation term in the association network, so we investigated how many

components in the network included transposases. A total of 21 of the associating components had genes annotated as transposases or transposition related genes. This suggests that 10% of the total 210 components in the network belong to transposases, highlighting the importance of mobile genetic elements and horizontal gene transfer for the accessory genome size and diversity. The 21 components that included transposases (11, 74, 5, 42, 33, 175, 30, 80, 18, 41, 57, 110, 2, 9, 22, 8, 48, 168, 199, 196, 208) had a total of 38 unique transposases (Supplementary Table 2). Two of these components were the two biggest components found in the association network 11 and 74 (Figure 8). The component 11 had 790 nodes and 72 of the nodes were for transposase genes. The most common one was IS5 family transposase IS1405. The component 74 had 730 nodes and 16 of the nodes were for transposase genes with the two most common being IS5 family transposase IS1021 and IS5 family transposase ISAzo23. However, we also found a lot of gene components that were not associated with transposable elements and have hence used a different means of moving across the phylogeny.

Overall, most of the genes in associating with the transposons were hypothetical. However, component number 33 had a transposal element associated with a molecular transport protein. The component had 15 nodes and thirteen of them were hypothetical genes but one of them group_ 10388 is annotated as *cycA* a D-serine/D-alanine/glycine transporter (Robbins and Oxender, 1973) (Supplementary Table 1). Five of the transposase-associated components represented multiple genes all encoding the same transposase (5,42, 80, 22, 8). For example, component number 5 is composed of 23 genes all of which encode IS5 family transposase IS1021. Component 42 consists of two genes both encoding IS3 family transposase ISAisp2. Component 80 is two genes both encoding IS3 family transposase IS401. Component 22 is two genes both encoding IS3 family transposase IS401. Component 8 is two genes both encoding IS701 family transposase ISRso17. One component, number 41, had two genes a transposon Tn3 family transposase ISPa43 and a recombinase Tyrosine recombinase XerC associated with it.

We also examined the components that had the highest D-value which meant they were most lineage independent. Some nodes were highly dispersed across the phylogeny and therefore had a high D-value score which indicates high lineage independence as the gene is present in

a random not lineage-associated pattern in the phylogeny (Fritz and Purvis, 2010). The components that had the highest average D scores across the nodes they were made of were number 4 and number 25. The top component was number 4, made of seven genes, one of which was the gene *capV* - a gene known to be part of a two-component complex *capV* and *dncV* in *Escherichia coli* where it provides 1000-fold protection against phage P1 (Cohen et al., 2019). The second highest D value was in component number 25 which contained eight nodes, seven of the genes were hypothetical and one was associated with flagellin production control (hin_1~~~hin~~~hin_2). The *hin* gene encodes a DNA-invertase that regulates the synthesis of phase-2 associated with the gene *fljB* in the animal pathogen *Salmonella typhimurium* (Kutsukake et al., 2006). Flagella swimming motility is a key component of the *Ralstonia solanacearum* virulence (Corral et al., 2020). This result suggests that phage defence and virulence proteins are highly mobile as they were found within components associated with lineage independence. Moreover, a prophage integrase and a tail sheath protein were found in the association network within seven components (11, 43, 31, 70, 26, 178, 65). The prophage integrase *intA* was the most common phage-associated gene found and it was in five components but there were also two occurrences of a putative prophage major tail sheath protein *gpFl*. All the rest of the genes in these seven components were annotated as "hypothetical proteins". Prophages are known to provide genomic plasticity in *Ralstonia solanacearum* species complex (Gonçalves et al., 2021; Greenrod et al., 2022a).

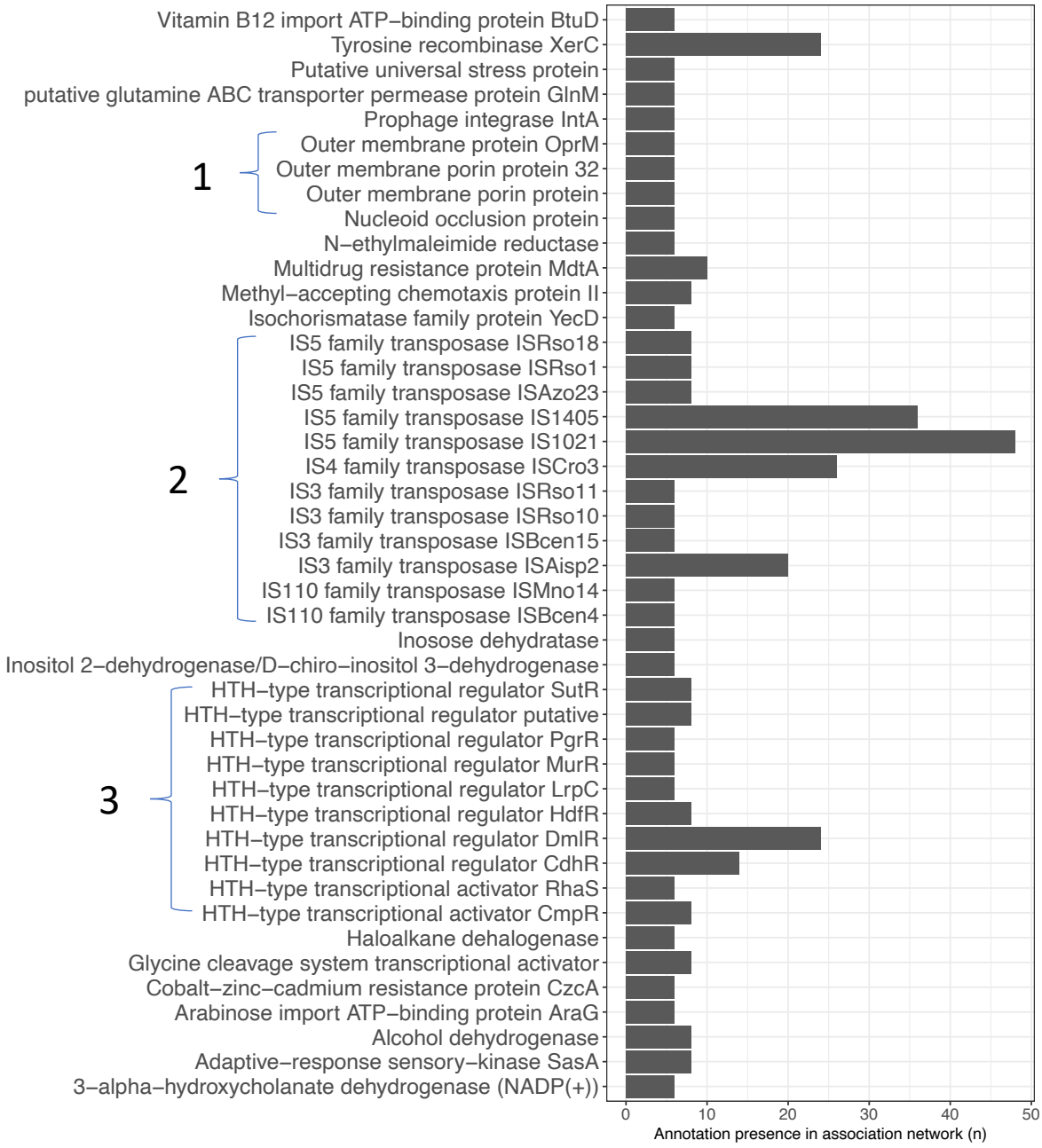


Figure 10. **Top annotation terms in association coincidence network**

The annotation terms for all components in the association network were pooled and most frequent ones are plotted (>= 5 occurrences across components). The term "hypothetical protein" that occurred 3,168 times is excluded for ease of visualisation. Labels 1-3 added for 3 groups of similar annotations as follows: outer membrane porins, transposases and HTH-type transcriptional regulators.

2.5. DISCUSSION

Here we performed a pangenome study of an extensive sample set *of Ralstonia solanacearum* species complex from all four phylotypes and from 55 different countries. We show that the species complex has a large pangenome with multiple unique genes in each phylotype. Moreover, we show that multiple genes associated with virulence, defence, prophages, and transposases are present in association modules within a cooccurrence network we constructed. This indicates transposal elements and prophage integrases are associated with virulence and defence and potentially with the transport of these genes within the species complex. This result aligns with the other finding that recombination occurred five times more often than mutation.

We estimated the size of the core genome of the *Ralstonia solanacearum* species complex using the IMG model based on a pangenome of 18,080 gene clusters to be 1704 (Collins and Higgs, 2012). However, the input for the IMG model is a dated phylogeny, and we were unable to reach sufficient sampling of the Markov Chain Monte Carlo (MCMC) model for the dating. Therefore, there is some uncertainty in this result as when the dating of the phylogeny is not sufficient problems associated with simple rarefaction estimates of core genome size apply (Collins and Higgs, 2012). Moreover, our pangenome size was smaller than a previous pangenome study on a smaller number of genomes of RSSC (n=131) that found an open pangenome with 32,961 gene families (Geng et al., 2022). This result is probably due to the overrepresentation of phylotype I and II isolates in our sample and the underrepresentation of phylotype III and IV. Moreover, here we constructed a whole genome phylogeny based on the core genome alignment of the core gene clusters. Our phylogeny clustering agreed with species delineation within RSSC, however, we failed to separate phylotype II into the previously estimated IIA and IIB (Safni et al., 2014a). However, problems with the separation of phylotype II into subgroups were observed before in a recent metanalysis that aimed to unify the nomenclature and classification across RSSC studies (Kurm et al., 2021). The researchers suggested a new phylotype IIC in needed to account for diversity based on whole genome studies (Kurm et al., 2021). These discrepancies mainly indicate the change from several gene phylogenies which were used to create the phylotype phylogeny and the new

whole genome sequencing era where phylogenies are based on the whole set of core genes of the species. However, overall, the phylotype concept holds well with the whole genome phylogeny performed here.

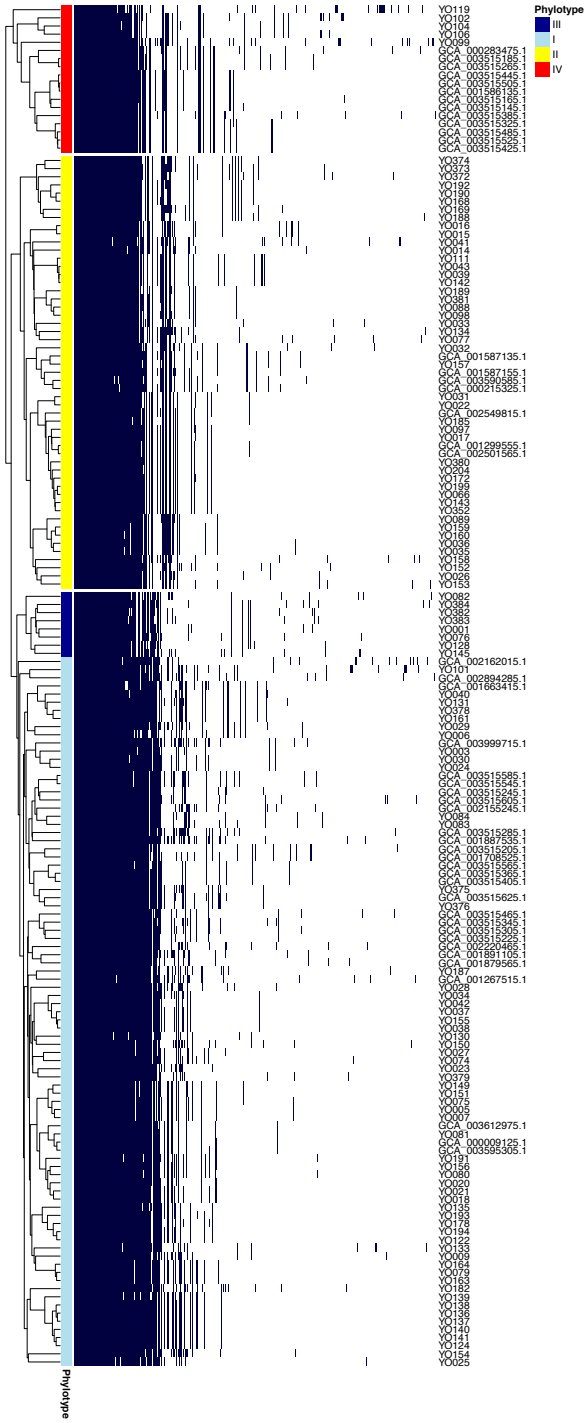
The presence of cooccurring genes suggests these genes exhibit a positive reinforcement (Whelan et al., 2020). Here we found 210 components of significantly associated genes in the accessory genome and amongst them there were transposal elements, prophage integrases, transcriptional regulators, and genes associated with abiotic defence. We previously reported that prophages are phylotype specific in the *Ralstonia solanacearum* species complex but here we are also reporting prophage integrase genes being sufficiently lineage-independent (Greenrod et al., 2022a). Prophages have been shown to directly influence the virulence of *R. solanacearum* in tomato (Addy et al., 2012). Therefore, the presence of integrase in associating components across the phylogeny suggests their role may be universal and associated with virulence and survival mechanisms in RSSC. Moreover, transcriptional regulators and IS element overrepresentation in the accessory genome association network indicate that diversification in RSSC happens with transposable elements moving around transcriptional regulators. This agrees with the previously proposed model for genetic diversification of RSSC with master regulators controlling virulence traits (Genin and Denny, 2012a). The huge genome of *R. solanacearum* can be regulated by introducing variation through transcriptional regulators. However, the massive number of hypothetical proteins found in the association network suggests there are a lot of highly niche-specific genes that we cannot infer a function of within the accessory genome. In addition, a previous study of gene cooccurrence in *E. coli* also identified transposable elements in the association network (Hall et al., 2021). Moreover, they found that one transposase *tnpA* was a hub gene making connections to around eight hundred other genes (Hall et al., 2021). These results show the importance of transposases in the accessory genome and show that they are not only part of it but a central component. Together the results indicate that transposable elements may be common within lineage independent accessory gene components not only in *R. solanacearum* but broadly across bacteria.

Here, we found 38 unique transposal elements within the association network, but under-sampling probably affects this diversity, and more sampling would have provided a greater

diversity of genes. Moreover, transposable elements move frequently therefore if multiple samples from the sample location and time, could provide better detection of transposases and show their dynamic movement across the genomes. For instance, isolate YO119 is the only *Ralstonia syzygii* (phylotype IV) isolate from clove from Indonesia isolated in 1983 we have. We found that an IS5 family transposase IS1420 (group_13645) is in proximity with genes Toxin *HigB-2* and Antitoxin *HigA-2*, DNA-invertase *hin,* modification methylase *DpnIIB*, LexA repressor and Tyrosine recombinase *XerC* in that genome. But in another 3 genomes from *R. syzygii* (YO102, YO104, YO106) all isolated from banana, again in Indonesia, but collected a few years later in 1987, we found the same toxin gene (Toxin HigB-2) on a DNA fragment close to putative HTH-type transcriptional regulator but no IS element. The lack of replication makes it hard to draw conclusions from comparisons like these and understand the dynamics of the accessory genome in RSSC. It is possible that the toxin-antitoxin systems have moved across strains of RSSC with the aid of an insertion sequence, but it is very hard to disentangle.

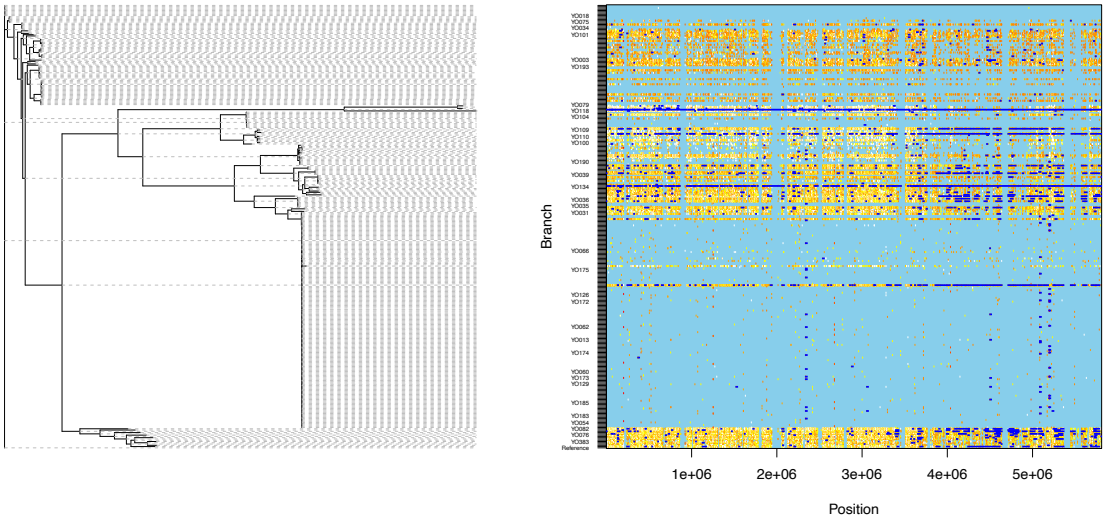
Overall, the comparative power of the analysis is limited here by the lack of replication for some isolates. Strains were under-sampled if they are from outside Europe, North America, or China. Phylotype IV (*R. syzygii*) strains are particularly badly represented in the dataset as it has only been found in Indonesia. Here, we aimed to include more countries and show the diversity across many countries which limits the analysis if a phylotype is only found within one country. The only country where we have multiple samples is the United Kingdom. Moreover, potato agriculture bias has driven the oversampling of the “potato race. However, the existence of a clonal branch is not exclusive to phylotype II and in phylotype I where more samples are available there are also clonal branches in the phylogeny. The phylotype I isolate which have been sampled from the same place over the space of several years shows clonality. For instance, several Mauritius samples over a few years from different crops seem to show high similarity. Therefore, we believe more clonal lineages will be present across the RSSC phylogenies if sampling is performed over time in different locations.

2.1. SUPPLEMENTARY

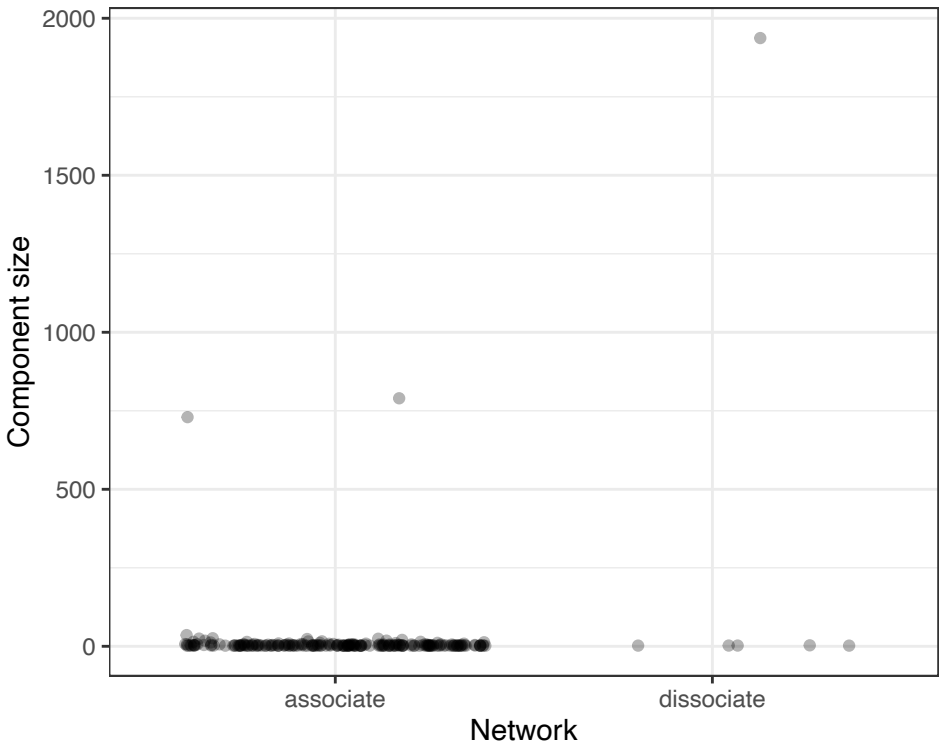


Supplementary Figure 1. **Panaroo pangenome gene presence-absence**

The hierarchical clustering shows the clustering of the bacterial isolates based on their gene presence and absence. Gene names on the x-axis are removed for ease of visualisation. 17,888 gene clusters are shown representing the pangenome.



Supplementary Figure 2. **ClonalFrameML default output graph showing recombinant regions in the genome.** The dark blue regions represent recombination. Sites that are non-polymorphic for a given branch are shown in light blue. Polymorphic sites are shown in a colour indicating their level of homoplasy: white is no homoplasy and the range from yellow to red is increasing degrees of homoplasy.



Supplementary Figure 3. **Coinfinder graphs component size**

The number of nodes in each component in the network graphs for association and dissociation shown. Points are jittered and slightly transparent for ease of visualisation.

N	**Gene clusters in component**	**Gene cluster annotation**
86	epsF_5...epsF_4...epsF_3	Type II secretion system protein F
	group_552	hypothetical protein
	epsE...hxcR_2	Type II secretion system protein E
	epsG...xcpT_3...xcpT_1...xcpT_2	Type II secretion system protein G
	group_53	hypothetical protein
	group_508	hypothetical protein
	xcpT_3...xcpT_5...xcpT_2...xcpT_4	Type II secretion system protein G
	group_11633	hypothetical protein
	sctC_6...sctC_7...sctC_4...sctC_5...sctC_8...sctC_2...sctC_1	Type III secretion system secretin

	group_11904	ATP-dependent zinc metalloprotease FtsH
4	group_11586	hypothetical protein
	group_5626	hypothetical protein
	group_4325	hypothetical protein
	group_10342	hypothetical protein
	group_11873	hypothetical protein
	group_4197	hypothetical protein
	capV	cGAMP-activated phospholipase
25	hin_1~~~hin~~~hin_2	DNA-invertase hin
	group_11821	hypothetical protein
	group_1447	hypothetical protein
	group_10227	hypothetical protein
	group_9261	hypothetical protein
	group_7173	hypothetical protein
	group_7172	hypothetical protein
	group_2264	hypothetical protein
	group_1243	hypothetical protein
33	group_3230	hypothetical protein
	group_132	hypothetical protein
	group_2387	hypothetical protein
	group_599	IS630 family transposase ISCARN39
	group_11985	hypothetical protein
	group_11756	hypothetical protein
	group_11422	hypothetical protein
	group_407	hypothetical protein
	group_10389	hypothetical protein
	group_4318	hypothetical protein
	group_3457	hypothetical protein

	group_10388	D-serine/D-alanine/glycine transporter
	group_8091	hypothetical protein
	group_6126	hypothetical protein
	group_10313	hypothetical protein

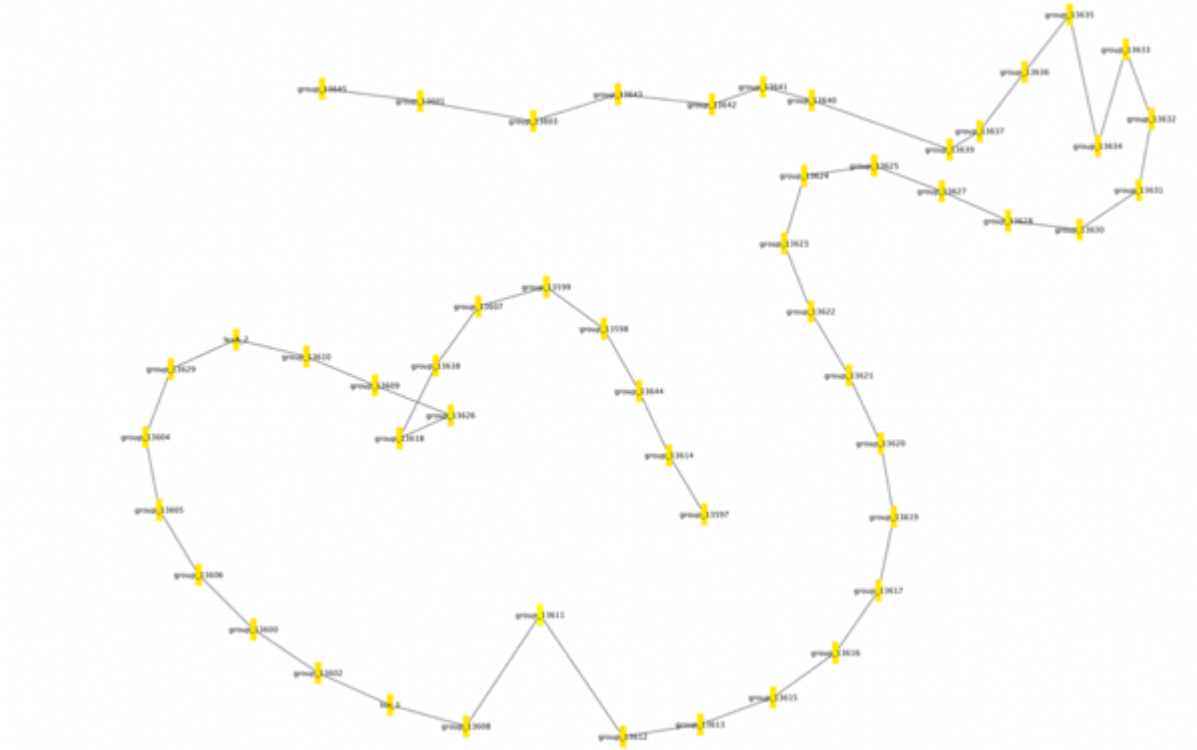
Supplementary Table 1. **Coinfinder graphs example components**

The gene clusters and their annotation as taken from panroo as presented in the table.

N	**Transposase Annotations**
1	IS21 family transposase ISRso6
2	IS21 family transposase ISRso19
3	IS3 family transposase ISAisp2
4	IS701 family transposase ISRso17
5	IS3 family transposase IS401
6	IS66 family transposase ISPa82
7	IS66 family transposase IS1313
8	IS3 family transposase ISRso11
9	IS3 family transposase ISButh1
10	Tn3 family transposase ISPa43
11	IS66 family transposase ISAeh1
12	IS3 family transposase ISBam2
13	IS30 family transposase ISHar4
14	IS630 family transposase ISCARN39
15	Putative transposase InsK for insertion sequence element IS150
16	IS5 family transposase IS1405
17	IS5 family transposase IS1021
18	IS5 family transposase ISRso18
19	IS5 family transposase IS1421
20	IS5 family transposase ISAzo11

21	IS3 family transposase ISRso10
22	IS5 family transposase ISAzo23
23	IS3 family transposase ISRso20
24	IS3 family transposase ISBcen15
25	IS3 family transposase ISRso14
26	IS3 family transposase IS222
27	IS3 family transposase ISPsy24
28	IS5 family transposase ISRso1
29	IS110 family transposase ISBcen4
30	IS4 family transposase ISCro3
31	IS30 family transposase IS1382
32	IS1182 family transposase ISBusp4
33	IS256 family transposase ISRso7
34	IS3 family transposase ISRso16
35	IS110 family transposase ISBma3
36	IS110 family transposase ISMno14
37	ISNCY family transposase ISBcen27
38	IS110 family transposase ISPye16

Supplementary Table 2. **Transposases found within the association network of coincident accessory genes.** 21 components in the association network of coincident accessory genes were found to contain a transposase within them. This table shows the 38 unique transposases identified.



Supplementary Figure 4. **Panaroo pangenome graph for IS1240 and Toxin – antitoxin associated gene fragment.** Network graph produced by panaroo.

Supplementary Figure 5. **Phylogeny used for Coinfinder.**

Highlighted in blue is the clonal branch that contains samples from Mauritius sampled over several years and one from Kenya. The clonal branch is within phylotype I showing that phylotype I can also harbour clonal lineages.

3. Three decades of survival in environmental reservoirs within the UK led to little genetic diversification within *Ralstonia solanacearum*

3.1. ABSTRACT

In temperate climates, a cold-infective strain of *Ralstonia solanacearum,* known as phylotype IIB-1 (Race 3 Biovar 2), causes the devastating brown rot disease in potatoes. In the UK, the presence of *R. solanacearum* in the river water and its association with a secondary host, *Solanum dulcamara,* has been monitored since the first recorded outbreak of the disease in 1992. Here, we present an in-depth analysis of the genetic diversity of the UK *Ralstonia solanacearum* population using a time series sample spanning 27 years (1992-2018) representing the UK-wide spatial distribution. Our analysis shows that very little genetic variation exists in the UK population of *R. solanacearum,* and strains show no clear differences regarding their isolation source or geographic location. Specifically, we find minimal core genome variation with mutation rates as small as one nucleotide change per year and estimate that the population originated between 1958 and 1988. The accessory genome comprised 55 intermediate-frequency genes with a total pangenome size of 4,725 genes. Temporal accessory genome GWAS identified a gene related to *Bacillus brevis* antibiotic synthesis whose presence was associated with time of isolation. Interestingly, three distinct populations with unique sets of accessory genes could be identified. Population 1 was only present during the first sampled outbreak and could not be detected ever since. In contrast, populations 2 and 3 were observed until the end of the sampling. The two populations differed regarding a large genetic region containing 20 genes associated with metabolic functioning such as maltose and trehalose assimilation. Our results show that after a reduction in accessory genome size during the first recorded brown rot disease outbreak, there has been only little genetic diversification in the *R. solanacearum* UK population indicative of punctuated equilibrium theory.

3.2. INTRODUCTION

Ralstonia solanacearum species complex (RSSC) is a group of plant pathogenic Beta-proteobacteria causing wilt disease on over 200 plant species (Hayward, 1991a). Currently, the species complex is divided into three bacterial species or four phylogenetic groups called phylotypes (I-IV) making up the *R. solanacearum* species complex phylogeny: *Ralstonia pseudosolanacearum* or phylotype I and II; *Ralstonia solanacearum sensu stricto* or phylotype IIA and IIB; and *Ralstonia syzygii* or phylotype IV (Safni et al., 2014a). RSSC genome consists of two replicons knowns as the Chromosome and the Megaplasmid (Genin and Boucher, 2004). All RSSC strains grow best at warm temperatures around 28°C and are most prevalent in tropical and sub-tropical regions of the world and unable to survive temperatures below 4°C (Milling et al., 2009). However, a cold-infective strain called the "potato race" can infect potatoes in the temperate climatic zones in Europe and North America, leading to a disease called “brown rot” that causes the rotting of underground tubers (Peeters et al., 2013b). The strain is also known as Phylotype IIB/Sequevar1 (IIB-1) or historically Race 3 Biovar 2 (R3bv2) and can infect tomatoes, eggplant, geranium, weeds, and wild plants (Prior and Fegan, 2005). It is thought that phylotype IIB-1 originated in the cool highland regions of the Andes and spread from there to many potato-growing areas worldwide, presumably with human movement in the trade of infected seed potato tubers (Champoiseau et al., 2009). If seed tubers are infected with *R. solanacearum* and stored at cool temperatures, the infection can remain latent, and symptoms are not observed. In this form, the pathogen can be spread over large distances with national and international trade of seed potatoes. Consequently, if the seed tubers are traded to warmer climates where the pathogen meets favourable conditions, symptoms of brown rot can develop (Elphinstone, 2001). Brown rot disease is a huge issue for the potato industry and expensive monitoring of potato seed and contaminated areas are required after an infection has been detected. Bacterial wilt on potatoes alone has been estimated to affect 3.75 million acres in approximately 80 countries, with global damage costs estimated from the early 2000s exceeding 950 million US dollars per year (Yamada, 2016). Therefore, all of the RSSC strains are listed as recommended quarantine pathogens by the European plant protection organisation and are listed in the EU and UK Plant Health legislation as such (EPPO, 2018).

Phylotype IIB-1 became agriculturally relevant when it was first found in potatoes in Europe in the 1990s. It was accidentally introduced to North America and Europe with infected geranium cuttings, but it has not been established in North America (Janse et al., 2004; Kim et al., 2003). However, Phylotype IIB-1 is a problem not only for European potato farmers but also in many mountainous regions of Africa, South America, Australia, and Asia. Coalescent models estimate that from South America, the strain moved to Africa and Europe and America (Clarke et al., 2015). In the Mediterranean region, Phylotype IIB-1 was first detected in the 1920s in Egypt, Italy, and Spain (Clarke et al., 2015). The strain can currently be found in several countries in Europe and around the world in the temperate climatic zone (EPPO, 2018). Moreover, the pathogen's distribution is not confined to agricultural fields alone, as it has been established in freshwater ecosystems in association with the woody nightshade (*Solanum dulcamara*) in multiple European countries (Parkinson et al., 2013).

Insight from evolutionary genomics can tell us about *Ralstonia solanacearum's* constraints in the river water environment and its potential to evolve and adapt locally. *R. solanacearum* phylotype IIB-1 has spread across the old and new worlds and has remained surprisingly clonal (Clarke et al., 2015). Strains isolated 50 years apart show a mutation rate as low as 1.99×10^{8} base pairs per year (Clarke et al., 2015), which suggests that mutations are not playing a major role in the evolution and adaptation of this strain. Still, very little is known about the variation in the accessory genome of this strain. The accessory genome is the gene content of a species that can vary across strains and isolates and depending on ecological niche occupation (Brockhurst et al., 2019). The accessory genome is not constant and not shared by all isolates or strains of the species like the core genome. Together the accessory and the core genome make up the pangenome, which can be defined as all the genes in the species or strain of bacteria studied (Tettelin and Medini, 2020). The accessory genome of a species can vary in the absence of core genome variation. For instance, a study of the river water and *Solanum dulcamara*-associated population of *R. solanacearum* phylotype IIB-1 over two years in the Netherlands showed no variation in core virulence genes but differentiation at the accessory genome level, represented as insertion sequence differences (Stevens and van Elsas, 2010). By performing PCR-restriction fragment length polymorphism analysis of 7 selected genomic loci, they saw homogeneity across the strains with no single nucleotide polymorphisms (SNPs)

identified across the sample. In contrast, pulsed-field gel electrophoresis of restricted genomic DNA revealed the differential distribution ISRSo3 insertion sequence elements across four genetic groups, which differed in ISRSo3 copy number and tandem repeat differences in the gene which is predicted to be a hypothetical protein RRSL_04153 (strain UW551—(Gabriel et al., 2006)) (Stevens and van Elsas, 2010). This study suggests that the movement of insertion sequences play a vital role in the short-term diversification of phylotype IIB-1 isolates and can be observed in as little as two years before changes in the core genome are detectable.

In the UK, contaminated river water used for irrigation of potatoes has led to six out of the seven outbreaks on record since the bacterium was first detected in 1992 (Elphinstone et al., 1998b, 1998a; Tomlinson et al., 2009). Plant health inspections of the UK waterways have been carried out annually since the first outbreak in 1992 (Elphinstone et al., 1998a). Widespread river water and *S. dulcamara* sampling have been conducted during the surveys, which has led to detailed mapping of the distribution of this pathogen in the UK waterways. The six outbreak infections from contaminated irrigation water are linked to sewage runoff upstream of contaminated *S. dulcamara* plants (Elphinstone and Matthews-Berry, 2017). While the genetic diversity of the UK phylotype IIB-1 population has not been studied extensively at the whole genome level, Parkinson et al. (2013) used variable tandem repeat analysis to show that outbreaks are linked to *S. dulcamara* plants upsteam of the infected field site (Parkinson et al., 2013).

The bacteria can be detected in the river water in the summer months as they multiply during warm periods and leak out from the roots of *S. dulcamara*. The river water helps transmit the bacterial cells to new *S. dulcamara* hosts downstream or fields of potatoes and tomatoes if the water is used for irrigation (Elphinstone, 2001). *R. solanacearum* from the river water can be detected using a selective medium during summer if the water sample exceeds 2 million cells per cubic meter. During the winter, there is no detectable amount of *R. solanacearum* cells in the water (Elphinstone, 2001; Elphinstone and Matthews-Berry, 2017). An eradication program of contaminated *S. dulcamara* plants from 1998 to 2001 led to a significant decrease in the mean population size in the river water (Elphinstone, 2001). However, this procedure is difficult and unfeasible to be used for control. Therefore, the ultimate removal of the

bacterial population from the river requires a full understanding of *R. solanacearum's* phylotype IIB-1 genetics and biology. Studies of low genetic diversity pathogens, such as the phylotype IIB-1, can help elucidate fine-scale local adaption and help us detect the early stages of adaptation to the environment (Cellier et al., 2012).

Here, we investigated the genetic diversity of the population of *Ralstonia solanacearum* phylotype IIB-1 in time and space within an environmental reservoir: the river network. We used a 26-year long time series comprising of 170 *Ralstonia solanacearum* isolates, covering the full known distribution in the UK waterways. The samples were sourced by Fera ltd from river water surveys since the first recorded potato outbreak in 1992, consisting of either sampled water, host plant *S. dulcamara*, and samples from potato and tomato outbreaks. We hypothesised that an observable evolution signal could be observed in time (year of isolation) or space (geographic location of isolation), influencing the population structure of *R. solanacearum* phylotype IIB-1. We also hypothesised that time would be significantly correlated with the genetic distance from the original isolates, indicative of genetic diversification in the rivers, which could further provide a time estimate for the *R. solanacearum* phylotype IIB-1 origin in the UK. We found that the population has a very small amount of genetic variation based on whole genome analysis of the extensive collection and a well-supported phylogeny was not obtained. However, the regression of time (year of isolation) and root-to-tip distance was sufficient to obtain a time estimate for the UK population origin. Moreover, based on the accessory genome, we see early signs of the development of two pathogen subpopulations co-existing in time and space in the environmental reservoirs in the UK. Our results show that after a reduction in accessory genome size during the first recorded brown rot disease outbreak, there has been little genetic diversification in the *R. solanacearum* UK population indicative of punctuated equilibrium theory.

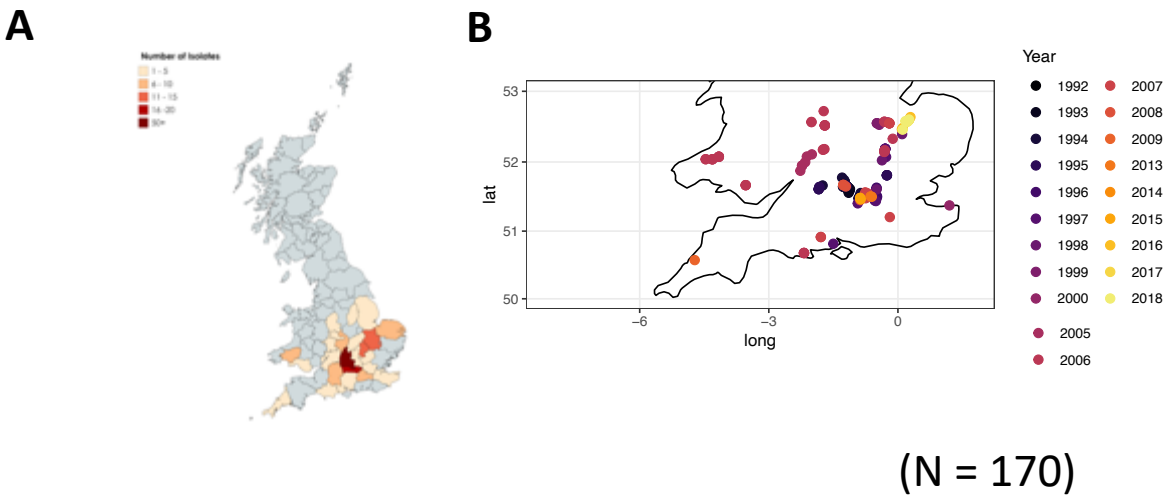
3.3. METHODS

3.3.1. Data

3.3.1.1. UK sampling and strain verification

The UK *Ralstonia solanacearum* river dataset spans 27 years of sampling since the first recorded outbreak in 1992 until 2018, covering 24 counties within England and was provided by Fera ltd. However, the sampling across time and space is uneven. This is because the river water surveys and sampling are conducted annually in September, but positive findings are only sometimes found. This can be due to detection limits and fluctuations in population sizes in the river water. Also, the testing happens in areas where they have previously detected *Ralstonia solanacearum*, and randomly in new areas to uncover expansions distribution. This has resulted in uneven distribution of samples across time and space and an overrepresentation of samples from Oxfordshire (where intensive sampling was conducted following the first finding in 1992), with 1/3 of the samples originating from there and from the year 2006 that was heavily sampled (Figure 11 & Figure 12).

The data used here are genome sequences of isolates belonging to the *Ralstonia solanacearum* species complex (RSSC), confirmed to belong to the RSSC complex by real-time PCR using a published protocol (Weller et al., 2000). The PCR protocol was designed to identify isolates within any of the four phylotypes of the RSSC. The isolates are part of a collection assembled at Fera Science Ltd. (York) and maintained within the National Collection of Plant Pathogenic Bacteria (NCPPB) as both research and reference isolates, preserved both at -80 °C in a cryoprotectant system (Protect) and freeze-dried cultures respectively. A large working cryostock library of the isolates and genome database of RSSC was created at The University of York. The metadata available for the isolates is based on records from the culture collections at Fera Science Ltd. 384 *Ralstonia solanacearum* species complex isolates were chosen from the Fera collection to represent the UK and a World collection: 176 isolates representing the UK *Ralstonia solanacearum* population and 208 representing the known worldwide diversity across the RSCC.

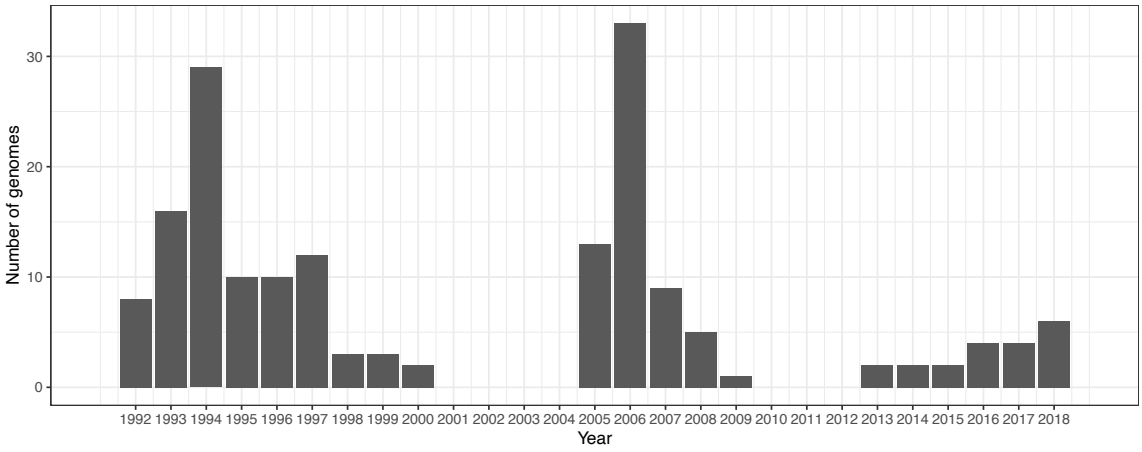


(N = 170)

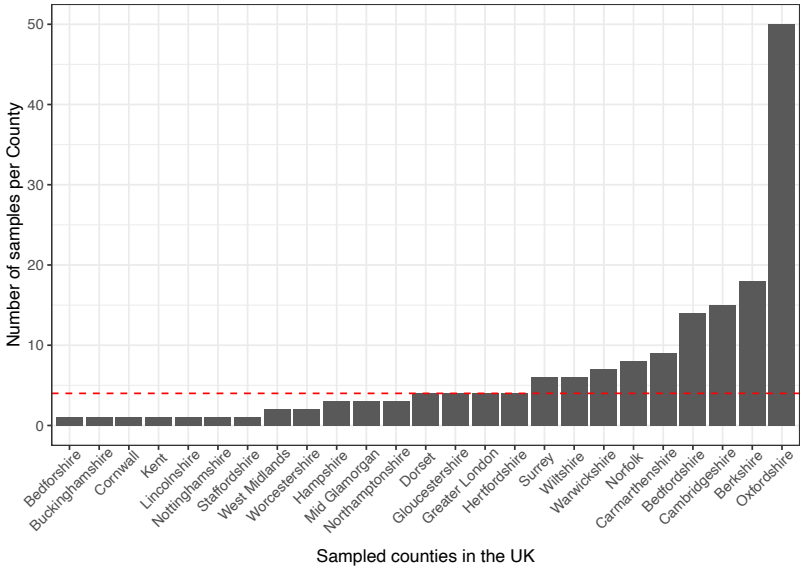
Figure 11. **Sampling map of *Ralstonia solanacearum* across the UK**

The maps on show location sampled across the UK: A) number of samples summarised by county; B) samples coloured by year of isolation and plotted based on geographical coordinates in the map.

A


Number of genomes
0
10
20
30
1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018
Year

B


Number of samples per County
0
10
20
30
40
50
Bedforshire
Buckinghamshire
Cornwall
Kent
Lincolnshire
Nottinghamshire
Staffordshire
West Midlands
Worcestershire
Hampshire
Mid Glamorgan
Northamptonshire
Dorset
Gloucestershire
Greater London
Hertfordshire
Surrey
Wiltshire
Warwickshire
Norfolk
Carmarthenshire
Bedfordshire
Cambridgeshire
Berkshire
Oxfordshire
Sampled counties in the UK

C

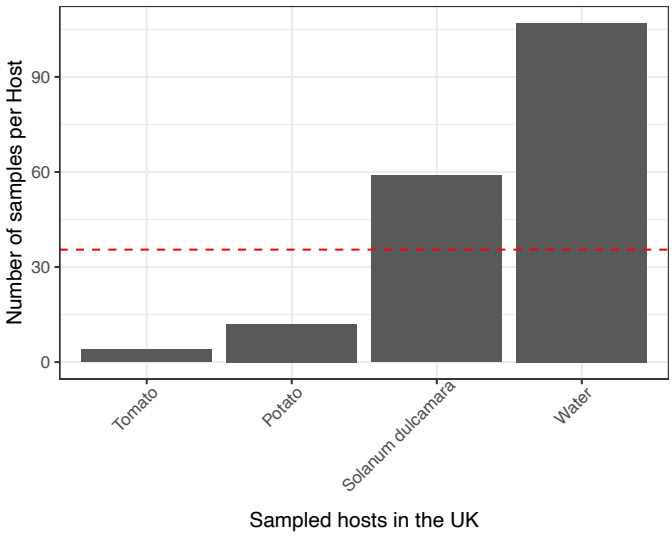

Number of samples per Host
0
30
60
90
Tomato
Potato
Solanum dulcamara
Water
Sampled hosts in the UK

Figure 12. **Uneven sampling of the UK *Ralstonia solanacearum* across time and space**
Figure shows plots of the number of samples: A) per year B) per county C) per host. Most of the samples are from Oxfordshire where the first outbreak in 1992 was detected. Missing years are due to lack of samples in the Fera Ltd. record for these years. The sampled hosts are predominantly associated with water and the riparian host *S. dulcamara* as the control methods of preventing farmers from irrigating crops with contaminated water has been highly effective.

3.3.1.2. Whole genome sequencing

The isolates were sequenced with Illumina MiSeq technology at the Earlham Institute UK. We received raw untrimmed paired FASTQ files for all 384 genomes. In addition, 24 of the isolates were re-sequenced with Nanopore MinIon technology at the University of York. The data received was base-called by the Bioscience Technology Facility at the University of York using Guppy, and raw FAST5 and FASTQ files were received. Following Martina Stoycheva quality checked the data and assembled the genomes. Illumina reads were filtered based on PHRED score with appropriate pairing, and the adapters were trimmed with TrimGalore. Kraken (Wood et al., 2019) was used to identify potential contaminant sequences in the reads. For the UK analysis presented in this study, six genomes from 176 UK genomes sequenced were deemed contaminated or low quality after quality control described below in sections for genome assembly and base calling. As a result, they were excluded from the downstream analysis resulting in 170 genomes in the dataset. All the raw FASTQ files from the Illumina and Nanopore runs are publicly available in the SRA archive and metadata under project number PRJNA823737.

3.3.2. Detecting genetic variants and construction of phylogeny

Single nucleotide polymorphisms (SNPs) and small insertions and deletions (indels) were called against strain YO199 (NCPPB 3854) - a reference genome originating from the first UK outbreak in 1992, which we created- and against an NCBI available high-quality reference genome UY031 (CFBP 8401; GCF_001299555). UY031 is the most complete and highest-

quality assembled genome of phylotype IIB-1 *Ralstonia solanacearum* from Uruguay (Guarischi-Sousa et al., 2016, p. 031). YO199 (NCPPB 3854) genome assembly is a high-quality hybrid assembly made from Nanopore MinION reads, and Illumina reads using Trycycler and polishes with Pilon. Snippy software (https://github.com/tseemann/snippy) was used to align all UK isolates against the YO199 genome. SNPs and small indels were called using minimum coverage of 10 and a minimum variant calling quality of 100. Core genome alignments were made only from the SNPs using the assistance script provided by snippy called snippy-core. The resulting core genome alignment had 171 sequences with 115 nucleotide sites, of which 42 were parsimony informative and 105 distinct site patterns. Gubbins (Croucher et al., 2015) with default parameters was used to clean the alignment from recombinant sites and construct a phylogeny with RAxML (Stamatakis, 2014). SNPs were also called against UY031, a reference strain from NCBI (GCF_001299555). Snippy with default parameters was used to generate BAM files. However, variant calling was done with the haplotype variant caller Freebayes (Garrison and Marth, 2012) on all isolates at once, as suggested by the developer to generate a whole genome VCF file which was annotated with SnpEff based on the gbff file for UY031. Freebayes (v1.3.2) was run with the following command: freebayes --min-coverage 5 -q 10 -m 60. The VCF generated contained 205, 693 small indels and single nucleotide polymorphisms. The VCF was then filtered for quality 30 based on the QUAL field using VCFtools (v0.1.16). Only reference calls were removed to get the calls for the population studied and not differences to the reference using VcfFilter (v0.2) with flag filterRefCalls. The filtering resulted in 250 SNPs and small indels. These were annotated using SnpEff (v 5.0) (Cingolani et al., 2012) and examined further manually using R, tidyverse and IGV (Integrative Genome Viewer) (Grolemund, 2017; R Core Team, 2017).

3.3.3. Genome assemblies

De-novo assemblies were made using Unicycler (v 0.4.7) (Wick et al., 2017a), and contigs were filtered based on GC content (>=0.66<0.67) and minimum length (5Mb). As a result, 357 draft genome assemblies were acquired. Mash (v2.3) was used to align contigs against a local RefSeq installation to identify potential contaminants in the assemblies (Katz et al., 2019). In addition, 24 of these genomes were re-made by a hybrid assembly of MinIon, and Illumina

reads using Unicycler. Furthermore, 48 assemblies for *Ralstonia solanacearum*, *Ralstonia pseudosolanacearum* and *Ralstonia syzygii* were downloaded from NCBI GenBank FTP on 10/06/2020 along with 1 *Ralstonia pickettii* 12J representative genome strain for the species and type strain for the *Ralstonia* genus. The 47 genomes were chosen if labelled as *Ralstonia solanacearum* (taxonomy number 490) and as complete assemblies in NCBI (See Appendix 2 for accessions). All assemblies were annotated with prokka (v1.14.6) (Seemann, 2014a). Also, the annotation of *R. solanacearum* type III effectors (T3Es) was added using the RalstoT3E pangenomic database (Sabbagh et al., 2019). We obtained a local installation of the T3E database and ran the blast searches against default parameters using the preconfigured Docker/Singularity container from the database publisher. We used default software configurations.

3.3.4. Pangenome

Panaroo (v1.2.9) pipeline (Tonkin-Hill et al., 2020) was used to create pangenome networks for all the genomes, including those sequenced here and those downloaded from NCBI. The coding sequences obtained from prokka were corrected for mis-annotation using a custom script provided in the panaroo accessory script GitHub page. Within the pipeline, the coding sequences were clustered by CD hit and alignments of the core genomes were created with ClustalO using the default setting - 90% of clusters shared by all isolates as a definition of the core genome. Maximum likelihood phylogenies were created from the core genome alignments with IQtree (v2.1.4) (Nguyen et al., 2015b). The GTR+G model was used, and for error correction, the consensus of two bootstrapping methods: 1000 UltraFast bootstrap (Hoang et al., 2018) and 1000 Alrt.

3.3.5. Insertion sequence detection

Insertion sequences were detected in the YO199 reference genome we created with ISEScan (Xie and Tang, 2017) using default parameters. To account for potential false positives, all putative ISs were blasted against the ISFinder database (https://isfinder.biotoul.fr/), with true positives determined if they had an E-value < e-04. ISs identified using ISEscan were used as queries to identify the insertion sites of IS elements in 176 UK isolates using ISMapper (v2.0.2)

(Hawkey et al., 2015) with default settings. Briefly, ISMapper identifies IS insertion sites by mapping isolate reads to previously identified reference ISs. It then parses out reads that overlap with the left and right IS flanking regions and maps them against the reference assembly, highlighting the specific positions of ISs in the query isolates. In line with a previous publication using ISMapper (Hawkey et al., 2020), insertion site precision was improved by running ISMapper using an IS-removed YO199 assembly. The genes flanking putative IS sites were determined using an annotated ancestral assembly generated using the stand-alone NCBI prokaryotic genome annotation pipeline (Tatusova et al., 2016).

3.3.6. Statistical analysis and computation

3.3.6.1. Tajima's D

To obtain the classic population genetics statistic for deviations of standard expectations of genetic diversity Tajima's D we used the software Popgenome (v2.7.5) (Pfeifer et al., 2014). The core genome alignments of the UK isolates were used to estimate Tajima's D and Pi with default settings.

3.3.6.2. Molecular clock fitting

In order to estimate time of origin for the UK *R. solanacearum* population, Bayesian inference of molecular clock using Markov Chain Monte Carlo (MCMC) performed with BactDating (v1.1) (Didelot et al., 2018). BactDating is ran on a core genome phylogeny and dates of isolation of the samples. We used the phylogeny generated from the Freebayes alignment and IQtree described above. We used default model settings. Tests for MCMC convergence were done with the R package coda (v0.19-4). All graphs were generated with R (R Core Team, 2017) in RStudio (R Studio Team, 2020) using tidyverse and ggplot2 (Wickham, 2009). Jaccard distances were calculated using the Vegan package (v2.6-2) (Dixon, 2003).

3.3.6.3. GWAS

To link phenoptypic traits with genetic differences, genome-wide association study (GWAS) was performed with pyseer (v1.3.10) (Lees et al., 2018) on genetic variants: VCF file from UY031 variant calling with Freebayes against all BAM files from snippy, unitig variation from unitig-counter (v1.1.0) (Jaillard et al., 2018), orthologous gene cluster variation obtained as gene presence/absence file from panaroo. The hclust algorithm was used to determine the hierarchical clustering of the mash pairwise genetic distances. The genetic diversity measures

of the genomes (*k*-mers, core mutations and gene presence/absence) were converted to pairwise distance using the mash algorithm for the *k*-mers and Jaccard distance for binary variables. Averages were used to linearise the distances. To turn the geographic distances between the geolocation coordinates available for the places of isolation for each bacterial genome (isolate) into a linear variable, we created a geographic distance measure from the geo-coordinates of each isolate in our collection. We took a mean for each isolate of all pairwise distances. The p-value distribution for the GWAS was checked using a QQ plot and we made sure there were no major shelving visible on the plots which can be a sign of unaccounted for population structure and that at least the smallest points fall on the theoretical distribution. Afterwards p-values were filtered based on a threshold created by the pyseer software automatically using Bonferroni correction. The scripts for filtering based on the p-values are available as an accessory script for pyseer. The University of York's local high-performance cluster "Viking" was used to run intensive computation jobs: assemblies, phylogenies, alignments, and GWAS.

3.4. RESULTS

3.4.1. Based on core-genome phylogeny, the UK *R. solanacearum* population is a highly clonal group belonging to the clonal strain phylotype IIB-1.

We constructed a species complex phylogenetic tree to place the UK *R. solanacearum* population within the *R. solanacearum* species complex taxonomic groupings. The phylogeny was based on the core genome of 357 draft short-read *de novo* assemblies, 24 hybrid long-read and short-read *de novo* assemblies. These assemblies were all from the world and UK RSSC sample set available at York University, plus 48 complete genome assemblies from NCBI representing the known diversity of the *Ralstonia solanacearum* species complex along with *Ralstonia pickettii* strain 12J as an outgroup for the *Ralstonia* genus (Figure 13). We found that all the UK isolates in our collection belong to a clonal branch within the *R. solanacearum* phylotype II, known as phylotype IIB sequevar 1 (phylotype IIB-1), the "potato race" or Race 2 Biovar 3 (R3bv2). There was no distinct clade within the phylotype IIB-1 clonal branch specific to the UK population, as all the phylotype IIB-1 isolates belonged to the same clonal

branch regardless of country and year of isolation. A worldwide phylotype IIB-1 clonal branch in the *Ralstonia solanacearum* species complex phylogeny indicates that it has been driven by a relatively recent population expansion that has not yet resulted in patterns of local adaptation between different regions. Thus, due to the small amount of genetic variation within the phylotype IIB-1 and the UK, we were unable to construct a meaningful phylogeny for the UK population with branches in the tree not receiving sufficient bootstrap support to be deemed non-randomly distributed (Figure 14). Therefore, to investigate the genetic variation within the UK phylotype IIB-1 we used classic population genetics allele diversity measures based on the core genome alignment of the 170 UK isolates. There was a total of 115 segregating sites in the core genome alignment. Based on the observed allelic diversity (Pi = 3.8), we calculated the Tajima's D value for the UK population (Taj D = -2.6). Negative Tajima's D values indicate that genetic variation is lower than expected, given the number of variable sites in the alignment. This is the expected signal for a clade with expanding population size, indicating that one dominant haplotype exists, and rare alleles are overrepresented. Therefore, this result is consistent with the hypothesis that *Ralstonia solanacearum* has recently spread to the UK and has not had enough time to accumulate changes in its core genome.

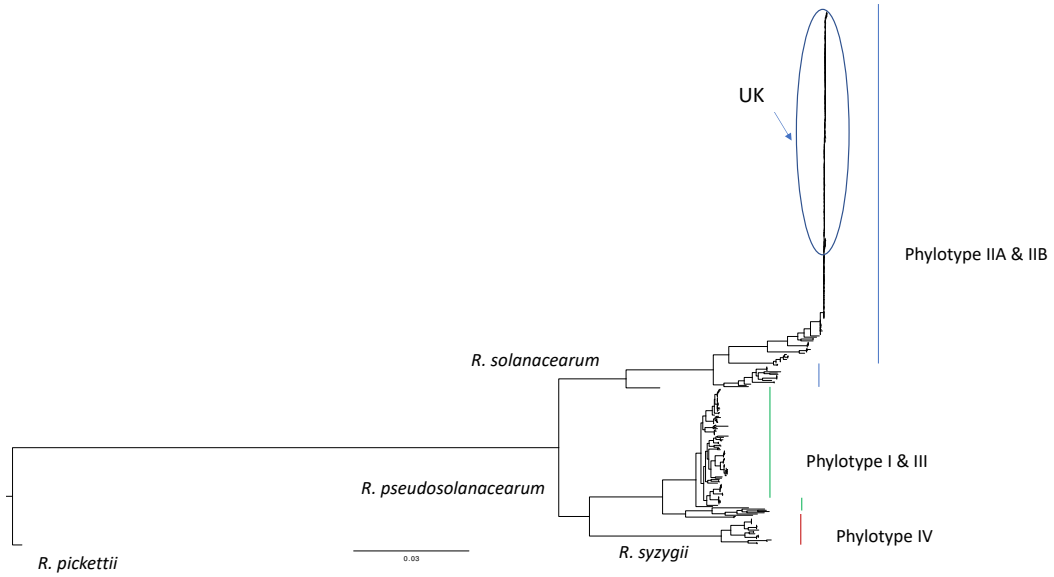
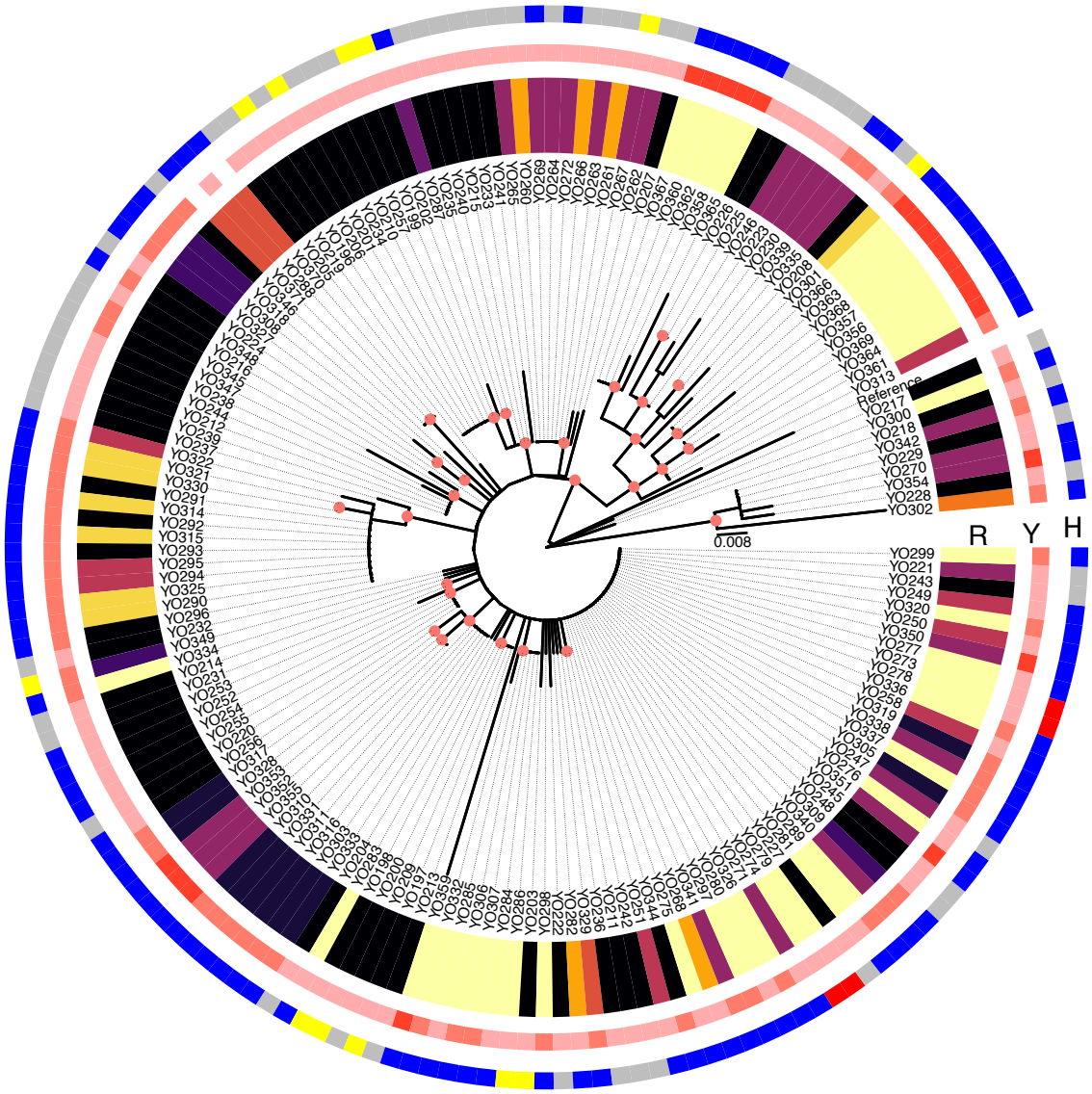


Figure 13. **Phylogeny highlighting the UK isolates belonging to clonal strain phylotype IIB-1 (race 3 biovar 2).**

Maximum Likelihood (ML) phylogeny based on 435 RSSC genomes: 357 draft short read data only assemblies + 24 hybrid long read and short read assemblies + 48 *Ralstonia solanacearum* species complex NCBI genomes and *Ralstonia pickettii* 12J used as an outgroup. All the species and phylotypes of the RSSC are labelled on their corresponding branches. The blue circle indicates all the genomes originating from the UK which all fall on the same clonal branch. The tree was generated with IQTree and GT4 model.



R=Region

- East Anglia
- East Midlands
- Essex & London
- Humberside
- Scotland
- Severn & West. Mid.
- South
- South East
- South West
- Wales
- West

Y=Year

- 2010
- 2000

H=Host

- Potato
- Solanum dulcamara
- Tomato
- Water

Figure 14. **Phylogeny of the UK *Ralstonia solanacearum* population**
Maximum likelihood GTR+Gamma phylogeny based on the core SNPs in the UK *Ralstonia solanacearum* population. The colour bars show metadata for isolation of the samples. The hosts the samples were isolated from, the region within the UK and the year of isolation. The coloured dots indicate branches with Bootstrap support from two bootstrap methods higher than 70 for SH-aLRT and higher than 90 for UFboot.

3.4.2. The UK *R. solanacearum* isolates have a large deletion in Megaplasmid compared to a South American isolate

To investigate the divergence of the UK population of *Ralstonia solanacearum* phylotype IIB-1 we wanted to compare it to a reference strain from South America where phylotype IIB-1 is thought to have originated. We ran the variant calling against a high-quality reference genome for phylotype IIB-1 isolated from infected potato tubers in Uruguay in 2003 called UY031. Interestingly, only a single core genome polymorphism in the *wapA* tRNA gene was identified as shared between all UK isolate genomes compared to UY031. In contrast, many genes were found to be missing in the Megaplasmid of the UK strain (Supplementary Figure 6) and all UK *R. solanacearum* genomes had a large 45 kilobase pair region missing in their Megaplasmid compared to the UY031 reference. This region is between bases 1,865,500 and 1,910,000 on the UY031 Megaplasmid and contained 183 genes (Supplementary Figure 6). Overall, the absence of many genes compared to an Uruguay isolate is striking compared to the large clonality observed in the core genome. This result indicates that variation in the gene content may play a larger role in the adaptation to local environments within phylotype IIB-1 strains compared to core genome polymorphisms. Therefore, the view of a highly clonal strain moving across the world may be biased if gene content is not considered as the lack of nearly two hundred genes is a significant difference between genomes.

3.4.3. *Ralstonia solanacearum* phylotype IIB-1 is a recent arrival in the UK

Due to the presence of a correlation between time and root-to-tip distance of the phylogeny of the UK phylotype IIB-1 *Ralstonia solanacearum,* we wanted further to investigate the time

signal in the core genome alignment. Therefore, we fitted a molecular clock in the phylogeny and used time as an explanatory variable in a genome-wide association study. First, we ran the BactDating software on the maximum likelihood phylogeny estimated constructed based on the core genome alignment of the 170 genomes to estimate the rate of evolution through time. The absence of a well-supported phylogeny meant the molecular clock estimates would not be highly accurate (Figure 14). Nevertheless, the analysis showed a correlation between time (year of isolation) and root-to-tip phylogenetic distance (Supplementary Figure 8). If the sampling of the parameters is correct, then we would expect an accurate dating. Even though we obtained a sufficient sampling of the MCMC parameters (Supplementary Figure 8 & Supplementary Table 3), the root branch placement was difficult for the model, probably due to most isolates having <5 SNPs and seven genomes with 0 SNPs compared to one of the original outbreak isolates from 1992 – isolate YO199, which we used as a reference for the core genome alignment (Supplementary Figure 9). Despite this, the root branch date was estimated to be in 1979 [1958-1988] and the substitutions per site (mu) in the genome were predicted to be 1.01 [0.68-1.51] per year (Table 1). Also, the year 1992, when the first outbreak in the UK was detected, is five years after the upper limit of the root date confidence interval 1987. The first outbreak was associated with contaminated river water, so an introduction into the UK 5 years before the outbreak is plausible (Elphinstone et al., 1998b). Overall, the dating analysis suggests a slowly diversifying population that originated shortly before the first recorded outbreak in the UK.

Mu [CIs]	**Sigma [CIs]**	**Alpha [CIs]**	**Root Date [CIs]**
1.01e+00 [6.75e-01;1.51e+00]	2.41e+00 [1.40e+00;4.03e+00]	8.13e+00 [6.35e+00;1.03e+01]	1978.66 [1958.44;1987.59]

Table 1. **Bactdating MCMC results.**

The MCMC is run for a total of 100,000 iterations, with the first 100 discarded as MCMC burnin and the remainder sampled every 100 iterations.

Mu = substitutions per site. Sigma = the average time it takes for two lineages to find a common ancestor. Alpha = coalescent time unit = Ne*g where Ne is the effective population size and g is the duration of a generation.

3.4.4. Pangenome variation in the UK phylotype IIB-1 is greater than the core genome variation

Pangenome analysis compares the presence and absence of all the genes in a set of whole-genome assemblies, which is sensitive to sequence quality errors that can inflate the estimates of absence of genes in some genomes. Therefore, two genomes with smaller total assembly lengths were excluded from the gene content analysis, leaving 168 out of the 170 *de novo* assembled Illumina genomes for this analysis. The pangenome of the UK population was estimated to consist of 4,725 genes, of which 4,569 could be considered as core genes, and 55 accessory genes (Table 2). Moreover, 100 genes were found to be variable at low frequency (<15%). Together, these findings suggest that there is more accessory genome variation between the strains compared to the small number of core genome variants (Supplementary Figure 9). This result indicates that the pangenome of phylotype IIB-1 in the UK is more dynamic than the core genome, with gene variation present at both low and intermediate frequencies.

Gene group	**Percentage of genomes sharing an orthologous group**	**Gene orthologous groups (N)**
Core genome	≥ 95%	4,569
Accessory genome	15% - 95%	55
Low-frequency genes	< 15%	100
Pangenome (all genes discovered)	100%	4,725

Table 2. **Breakdown of pangenome.**

The number of orthologous gene clusters (genes) shown in the different pangenome groupings. The proportion of genomes sharing an orthologous group shows criteria for belonging to a different gene group.

3.4.5. Changes in accessory genome content indicates gene loss after the first outbreak in 1992, resulting in two co-existing environmental populations

Since we observed much more variation in the *R. solanacearum* accessory than in the core genome we decided to use the accessory and core genome gene cluster variation to cluster the strains and compare the differences in their accessory genome. Therefore, we clustered the 168 genomes using hierarchical clustering based on gene presence-absence matrix from panaroo (Figure 15). By focusing on the seven genomes from the original outbreak in 1992 (YO196 - YO203), we found that the original seven genomes represented all the accessory genome diversity in the UK population (Figure 15). Moreover, isolates from the original outbreak were present in the three clusters formed by the hierarchical clustering (Populations 1-3; abbreviated from here on as Pop1-3; Figure 15). In other words, the observed accessory gene diversity was jointly represented by three original outbreak isolate groups that differed from each other. Interestingly, Pop1 was only present within the original outbreak samples and was not detected in later samples. In contrast, Pop2 and Pop3 persisted in time and could be identified through the whole sampling period (Figure 16).

Pop1 (grey) consisted of three genomes (YO196, YO197, YO199) and had the biggest number of accessory genes out of all the groups (Figure 15), which suggests a reduction in accessory genome size and diversity after the first outbreak. Pop1 had 29 unique genes, which were not present in either Pop2 or Pop3 (Supplementary Table 5). Moreover, these unique 29 genes were absent from all other later isolates (Pop2 and Pop3). Of these 29 unique genes, nine encoded the IS5 family transposase IS1021 and 16 other genes were located together on the genome next to a transposal element (group_1922 encoding the IS5 family transposase IS1021 (Figure 17). Moreover, the genes in the 16-gene loop were related to virulence or

antibiotic resistance (Supplementary Table 5). For example, the *vgb_2* gene is a homolog of the *vgb* gene (PubMed:11467949) in *S. aureus* and encodes the Virginiamycin B lyase, which inactivates the type B streptogramin antibiotics. Furthermore, two important virulence genes were detected in this cluster: 1) *epsE_2*, which is a Type II secretion system protein E, encoding protein EPS I polysaccharide export inner membrane protein EpsE in *R. solanacearum* and 2) *epsF_3,* a close homolog of EPS I polysaccharide export inner membrane protein EpsF (Q45412). Both proteins are probably involved in the polymerisation and export of exopolysaccharide EPS I, which functions as a virulence factor and plays a role in exporting EPS I or its intermediates across the membrane of RSSC cells. These results suggests that Pop1 strains encoded certain virulence factors, which were quickly lost from the UK *R. solanacearum* population after the initial outbreak.

Pop2 (beige) was represented by four original outbreak strains (YO198, YO200, YO201 and YO202) and comprised 40 isolates in total (Figure 15), while Pop3 (purple) was represented by one of the original outbreak genomes (YO203). Interestingly, Pop3 is missing 20 genes which were present in both Pop1 and Pop2 (Supplementary Table 6). These 20 genes were located together on the genome nearby to a transposase (group_1913 encoding the IS5 family transposase IS1021), suggesting they were lost together as a unit (Figure 18). These genes encoded multiple functions related to metabolism, such as maltose and trehalose metabolic pathways (Supplementary Table 6). This result suggests that the surviving Pop2 could be better at metabolising maltose and trehalose compared to Pop3, suggesting that observed loss of genes could related to metabolic adaptations.

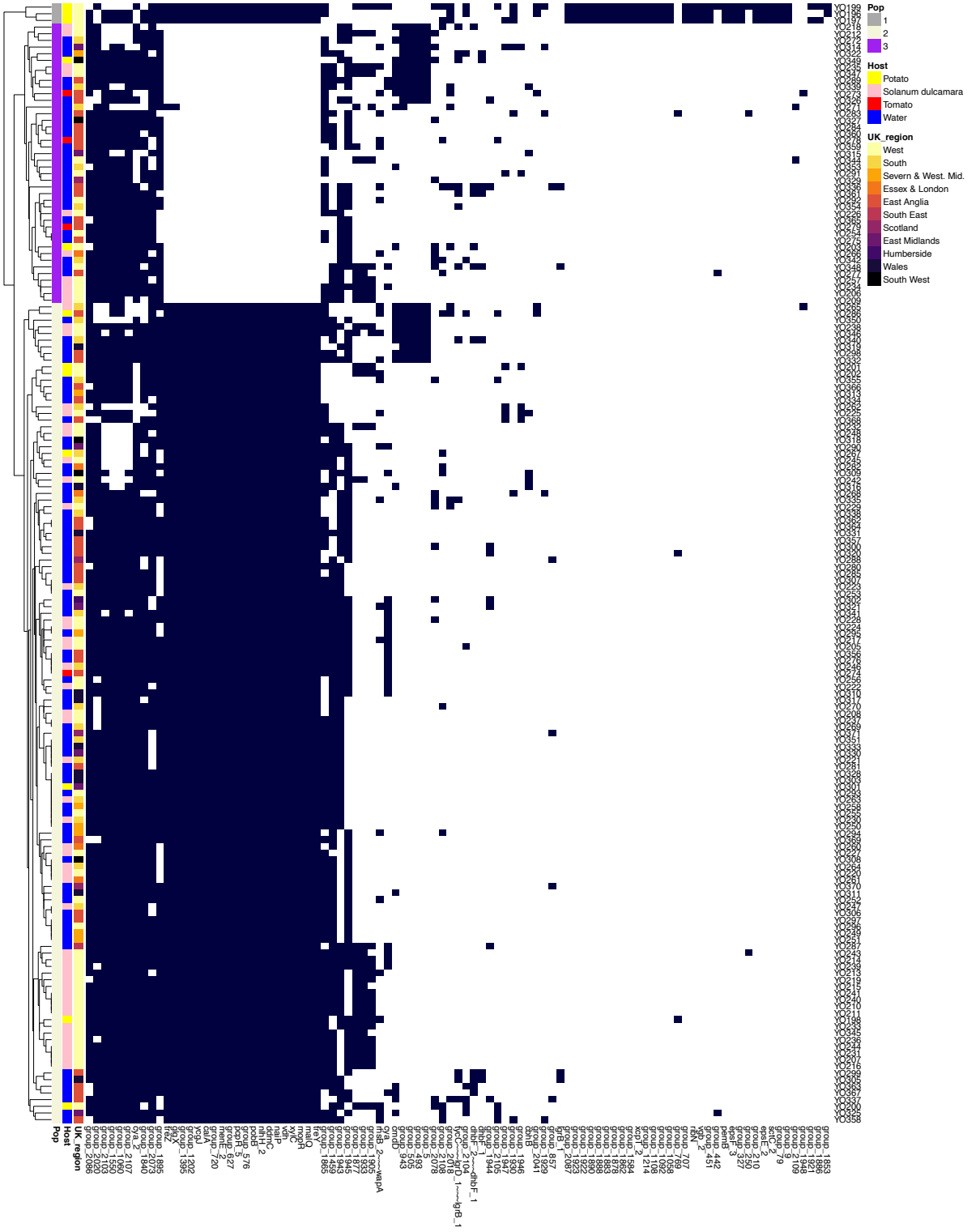

Pop
1
2
3
Host
Potato
Solanum dulcamara
Tomato
Water
UK_region
West
South
Severn & West. Mid.
Essex & London
East Anglia
South East
Scotland
East Midlands
Humberside
Wales
South West

Figure 15. **Accessory genome variation within the UK population of *Ralstonia solanacearum*.**

The binary matrix (presence and absence) of genes is shown as a heatmap where blue is the presence and white is the absence of genes. On the y-axis are all the bacterial isolates and on the x-axis are all the genes as named by panaroo after annotation correction. A gene here represents an orthologous cluster shared by the bacterial isolates studied and ~ represents the merge of sequence annotations by panaroo annotation correction. To show the true accessory genes in this plot we have also excluded genes shared at high frequency (>90%) and low frequency (<=1). A total of 95 variable genes are shown. Hierarchical clustering of the isolates (genomes) is applied to the heatmap.

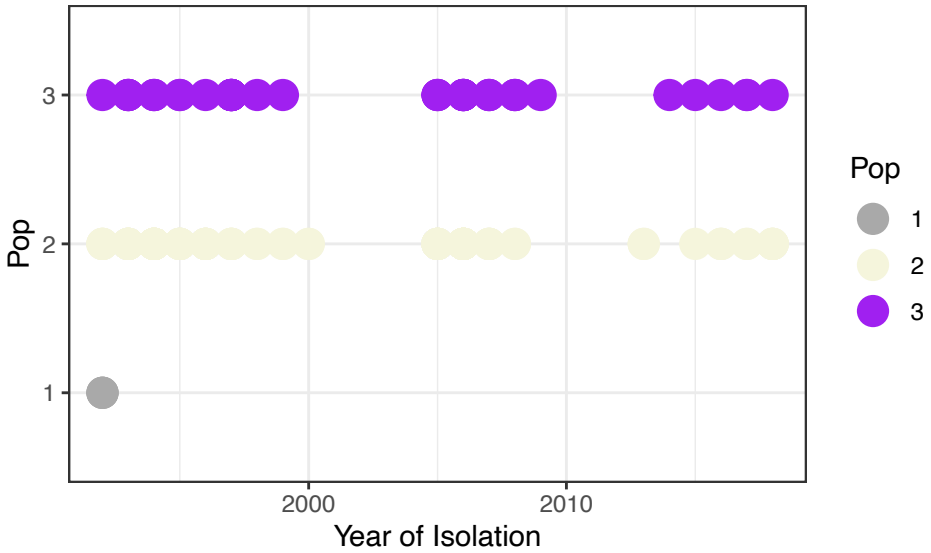


Figure 16. **Presence of Pop1-3 over time.**

The presence of each of the three genotypes Pop1-3 in the sampling across time in the UK. The x-axis represents time, and the y-axis shows each of the three genotypes. Pop 1 is only present in the first year of sampling 1992 and disappears after that.

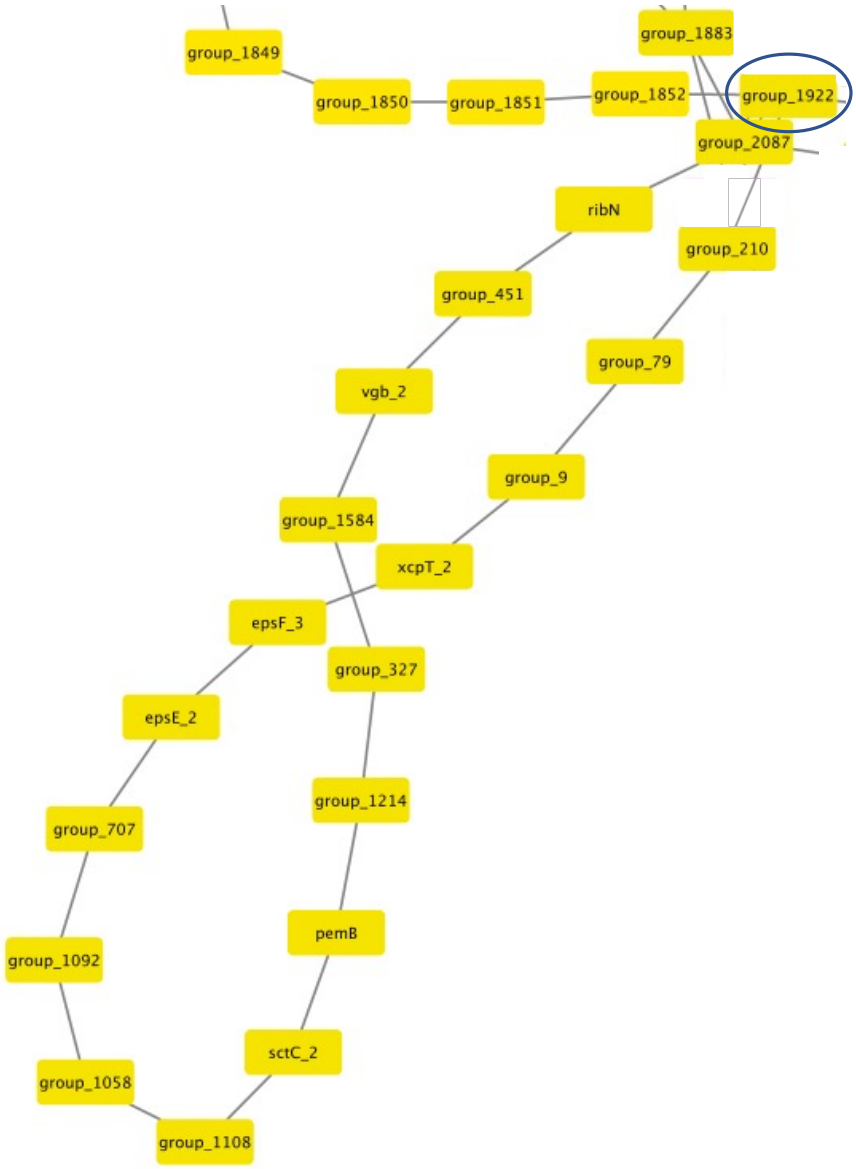


Figure 17. **Pangenome graph area of the genome unique to Pop1.**

The image shows the pangenome graph produced by panaroo visualised in Cytoscape. Only the area concerning the 29 genes unique to Pop1 is shown. The looped fragment represents the 16 genes specific to Pop1 including the vital for *Ralstonia solanacearum* exopolysaccharide (eps) genes. Circled in blue is nearby on the pangenome graph is the gene group_1922 encoding the IS5 family transposase IS1021.

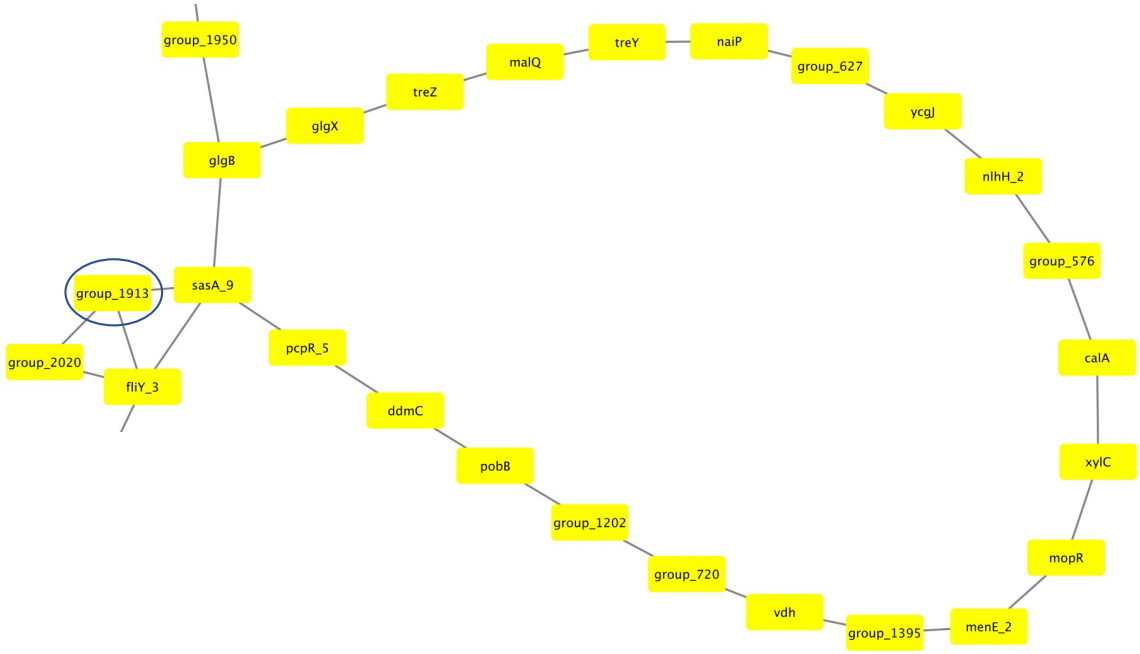


Figure 18. **Pangenome graph area of the genome missing in Pop3.**

The image shows the pangenome graph produced by panaroo visualised in Cytoscape. Only the area concerning the 20 genes missing in Pop3. The looped fragment represents the 20 genes specific to Pop2. Circled in blues is the gene_1913 encoding the IS5 family transposase IS1021.

3.4.6. GWAS identified only a few genetic variants associated with time

The genetic variation in the *Ralstonia solanacearum* phylotype IIB-1 population from the UK was tested for association with time in a genome-wide association study (GWAS). Time (year of isolation) was used as the phenotypic trait in the GWAS that was run using three different presence-absence genetic matrices: 1) shared unitigs (*k*-mers), 2) core genome SNPs called against UY031, and 3) accessory genes. We wanted to estimate whether variation in the sequences significantly increased in time and to be able to use accessory genome variation and core genome variation. The SNPs GWAS identified a variant on the 293,375 positions in the UY031 chromosome (Supplementary Table 4). This variant was detected in 23 genomes (YO246, YO208, YO230, YO335, YO367, YO313, YO361, YO223, YO360, YO339, YO225, YO226, YO365, YO356, YO368, YO363, YO369, YO357, YO301, YO358, YO362, YO364, YO366) with a

higher frequency during the last decade of sampling (post-2010). This locus was annotated as either *rhsC_3* gene homolog of putative deoxyribonuclease RhsC (NCBI-ProteinID: ALF86662) genome or as Betaproteobacteria toxic sRNA - RSUY_01040 - ncRNA T06155. The Unitig GWAS found another gene, which was significantly associated with time. This gene was annotated as a homolog of Tyrocidine synthase tycC (tycC_2 and tycC_3) in *Brevibacillus parabrevis* (Uniprot: TYCC_BREPA) (Supplementary Figure 11). In contrast, the gene presence GWAS identified four genes: group_2086, group_1933, group_1877, and group_1905 (Supplementary Table 4). The genes group_1933, group_1877 and group_1905 all encode an IS5 family transposase IS1021 (genes next to each other on the x-axis in Figure 15). This insertion sequence seems to be present or absent across the phylogeny regardless of where and when the samples were isolated, indicating that this transposase is actively moving across the UK phylotype IIB-1 population. The gene group_2086 encoded *lgrD*, which is the linear gramicidin synthase subunit D. This is a gene part of the four open reading frames, *lgrA*, *lgrB*, *lgrC*, and *lgrD* that all encode the linear gramicidin membrane channel-forming pentadecapeptide (Kessler et al., 2004). Gramicidin is an antibiotic molecule produced by another common soil bacteria *Bacillus brevis* during its sporulation phase (Hotchkiss and Dubos, 1940). Together, these results suggest the idea that IS elements might be important in generating genetic variation between the UK phylotype IIB-1 strains.

3.4.7. Insertion sequences show movement patterns indicative of location-specific population structure.

To gain further insight into accessory genome variation within the UK *Ralstonia solanacearum* population, we investigated the presence of insertion sequences (ISs) sites in the genomes, which are known to contribute to gene disruption. We found that ISs' insertion sites were location-specific based on the pairwise geographic distance of the sampled locations. We looked at the ISs' position of insertion within the genomes and compared it to the pairwise geographic distance. Nine out of 16 of ISs' positions in the chromosome and six out of 11 ISs' positions in the Megaplasmid were found predominantly in isolates that are situated within 25km of each other (Figure 19). In the chromosome, the genes disrupted included two hypothetical proteins and a response regulator transcription factor. Bacterial response regulators have been linked to various cellular processes ranging from basic metabolism and

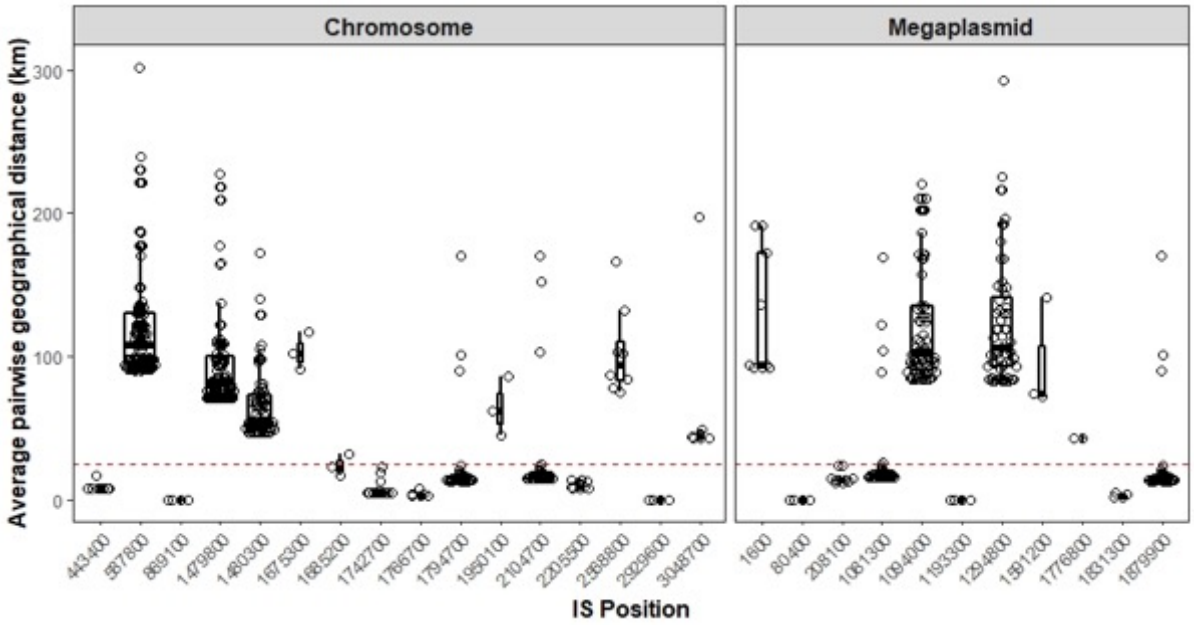


Figure 19. **Insertion sequences position of insertion in genomes closer together**
Isolates with specific IS positions are generally found within the same geographical location. The x-axis shows coordinates in the genome rounded to the closest 100bp. The y-axis shows the average pairwise geographical distance between isolates with each IS position. The dotted line shows 25km.

growth to biofilm formation (Gao et al., 2007). The disrupted genes included a hypothetical protein and an AraC family transcriptional regulator in the Megaplasmid. The major regulator HrpB, which regulates the type III secretion system, is an AraC family transcriptional regulator (Yoshimochi et al., 2009). In addition to disrupting genes, some ISs were intergenic and neighboured an ATP-dependent zinc metalloprotease FtsH and RlmE family RNA methyltransferase, phage-associated genes, and a glycoside hydrolase family six protein. Altogether, these results suggest that, on a location basis, the UK phylotype IIB-1 isolates may have undergone local adaptations through IS-mediated gene disruptions and re-activations.

3.5. DISCUSSION

In this report, we investigate a large set of whole genome assemblies based on a collection of 170 bacterial isolates of *Ralstonia solanacearum*. This collection is based on the annual environmental sampling of the river water in the UK performed by Fera Ltd. This sampling has resulted in a dataset covering all the detectable distributions of *Ralstonia solanacearum* in

the UK. There are samples from 24 different counties in the UK isolated across 26 years between 1992 and 2018. Most of the studied bacterial isolates are from river water and are thought to be associated with the asymptomatic secondary riparian host plant *Solanum dulcamara* (Elphinstone et al., 1998b). We found that all the isolates within our sample belong to the highly clonal phylotype IIB-1 strain of *Ralstonia solanacearum* (Figure 13). Although the investigated sample set is large, temporally, and spatially diverse, we found very little core genome variation. However, on the pangenome level, the population was relatively more diverse with 55 accessory genes present. The small amount of sequence variation in the core genome did not allow for the construction of a well-supported phylogeny. However, the significant temporal signal between time and root-to-tip distance meant the molecular clock fitting on the phylogeny was successful and we obtained an estimate for the origin of the population in the UK somewhere between 1959 and 1987. In addition, the accessory gene variation divided the UK population into three hierarchical clusters here called Pop1-3. Pop1 is only represented by 3 isolates from the first detected outbreak in the UK on potatoes in Oxfordshire whereas Pop2 and Pop3 are present from the first outbreak all the way to 2018. Moreover, Pop1 and Pop2 had lost multiple genes both in proximity to IS5 family transposase IS1021. In addition, the insertion sequence position of insertion was more commonly found to be the same if bacterial isolates were isolated within 25 km of each other indicating population structure could be detected in the movement of insertion sequences.

The link between European freshwater ecosystems and the long-term survival of *Ralstonia solanacearum* in the temperate climate environment needs to be better understood. For instance, whether the phylotype IIB-1 strain can survive in the soil or outside plant hosts for prolonged periods is unclear. Case field studies show that the phylotype IIB-1 strain survives in field soils for one to 4 months in regions where winter temperatures are around 4°C (Shamsuddin et al., 1978; Van Elsas et al., 2000). Interspecies bacterial competition and humidity are known to significantly affect the chance of *R. solanacearum* success in the soil environment. Previous studies have shown that the phylotype IIB-1 population from agricultural drainage water mixed with other microorganisms at 4°C goes extinct in 30 days (Van Elsas et al., 2001). Therefore, the combined stress of competition and low temperature is considered too great for phylotype IIB-1 to survive and thrive in the soil or freshwater for prolonged periods without a host (Granada and Sequeira, 1983; Van Elsas et al., 2000, p. 2).

Nevertheless, there is no systematic analysis of phylotype IIB-1 survival in temperate soils. The evolution of an environmental reservoir population of phylotype IIB-1 from river water in a long time series has yet to be investigated. In addition, the large host range of the species complex suggests that other wild plants apart from *Solanum dulcamara* can sustain infection and provide refuge.

In the UK, the population sizes of *Ralstonia solanacearum* in river water can be very large in the warm summer months, thus increasing the efficiency of selection and purging of deleterious mutations. Therefore, adaptation to new selection pressures could be observable in real time if samples are taken through a long enough period (Duchêne et al., 2016). The presence of *Ralstonia solanacearum* in the river water has been carefully monitored in the UK. This has allowed for outbreak containment. However, further investigation of RSSC strains' adaptation to the local environment and potential transmission routes to susceptible crops is needed.

We found that *Ralstonia solanacearum* phylotype IIB-1 from the UK is part of the highly clonal worldwide strain phylotype IIB-1 (Cellier et al., 2012). We saw an average of 3 SNPs per genome in the sample set studied here and a negative Tajima's D (Supplementary Figure 9). The lack of genetic variation is surprising considering the period of 26 years sampled here and the 24 counties across the UK. Due to the small number of SNPs, we could not construct a highly supported phylogeny as the small number of SNPs means a small number of parsimony informative sites in the core genome alignment (N=45) (Figure 14). However, the fitting of a molecular clock model is suitable when there is a statistically significant correlation between the year of isolation and the genetic distance from the most recent common ancestor also known as the root-to-tip regression (Nübel et al., 2010). Here, we observed significant albeit weak regression and we observed sufficient sampling of the Markov Chain Monte Carlo. Therefore, we believe our estimate of the population origin between 1959 and 1987 is the most accurate up-to-date estimate of the origin of the UK population of phylotype IIB-1 *Ralstonia solanacearum*. Moreover, from the model we estimated the mutation rate per year to be 1 (Table 1). Comparable rates of mutation to those found before worldwide and the small amount of total variation agree with a recent origin of the phylotype IIB-1 strain (Clarke et al., 2015). Clarke et al. (2015) found that 11 closely related phylotype IIB-1 (R3bv2) isolates

collected over 50 years did not accumulate more than seven mutations, and isolates collected in similar years were as different from each other as isolates collected 50 years apart. However, we found a 45 kb region missing in our samples compared to the UY031 South American isolate, which reflects the time phylotype IIB-1 isolates have had to evolve and differentiate in the different continents. This is in stark contrast to the single genetic variant difference between UY031 and YO199 identified in the rest of the genome.

Clarke et al. (2015) proposed that *R. solanacearum* bacteria may grow, and thus mutate, very slowly in soil and surface water populations, which often constitute the inoculum sources of outbreaks. Hence, most of the mutations that get fixed during fast pathogen replication during plant infection may constitute a sink rather than a source. The surface water concentrations of *R. solanacearum* must have been high enough to cause infections in the UK river water, as irrigation of plants with contaminated water has led to all five outbreaks of the disease in potatoes and two in tomatoes in the UK since the first recorded outbreak in 1992 (Elphinstone, 2001). It was shown that under optimum quarantine glasshouse conditions with single doses of aqueous suspensions of *R. solanacearum,* infection of potato only occurs when bacterial populations exceed 10^5 colony forming units (CFU) per plant and that wilting symptoms did not develop at the threshold level. Higher population sizes (> 10^6 CFU per plant) were required to infect 1-month-old *S. dulcamara* seedlings (Elphinstone et al., 1998a). Therefore, to be able to infect *Solanum dulcamara* the bacteria need to be present in high densities in the river. Thus, as suggested by Clarkson et al. (2015), slow growth and small population sizes seem unlikely reasons for the observed low genetic variation in the phylotype IIB-1 isolates.

One explanation could be that the bottlenecking of winter die-off in the water and growth limited to summertime with a leftover population as a refuge in the *S. dulcamara* combined with the already significant metabolic effort of surviving at temperate latitude is putting a large amount of pressure on the UK *Ralstonia solanacearum* making it hard for diversity to be generated for natural selection to act upon. It is, however, possible that the relationship is stochastic. Perhaps a very small number of founders is enough to cause an outbreak if it is in a very large volume of water. In this situation, a small number of plants could be initially infected by the water. The infection will spread further through the field through the soil and

during increasing generations of vegetatively propagated potato crops in a typical manner for *Ralstonia solanacearum*. Perhaps the lack of genetic diversity is due to population bottlenecks related to seasonal and environmental change. If the population is subject to a constantly changing environment and the bacteria experience huge fluctuations in selection and population sizes, perhaps caused by seasonal temperature changes and out-of-host versus inside-host environment. Then it is possible that as the autumn temperatures fall, the relatively high-density bacteria, after summer when they are sampled, seek refuge in the riparian host and therefore undergo fluctuating selection dynamics and bottlenecks.

Positive samples have not been detected in the colder months of the year (John Elphinstone personal communication). Similarly, phylotype IIB-1 could not be detected in irrigation water at 4°C in the Netherlands (Van Elsas et al., 2001). However, we note that we cannot conclude the selective pressure occurring in time due to the lack of repeated sampling of each geolocation present in the dataset over time. To the best of our knowledge, this sample set is the most extensive in time and space for the Phylotype IIB-1 strain of *Ralstonia solanacearum.* However, it does not provide enough replication to draw conclusions about the seasonal effects on population dynamics. Moreover, the sampling of the river water performed by Fera Ltd. is always done in September. Therefore, we are seeing only a late summer snapshot of the diversity present in the river water phylotype IIB-1 population. Moreover, the detection limit in the river water sampling is relatively high so presence of bacteria is possible below the detection limit of the protocol used by Fera Ltd.

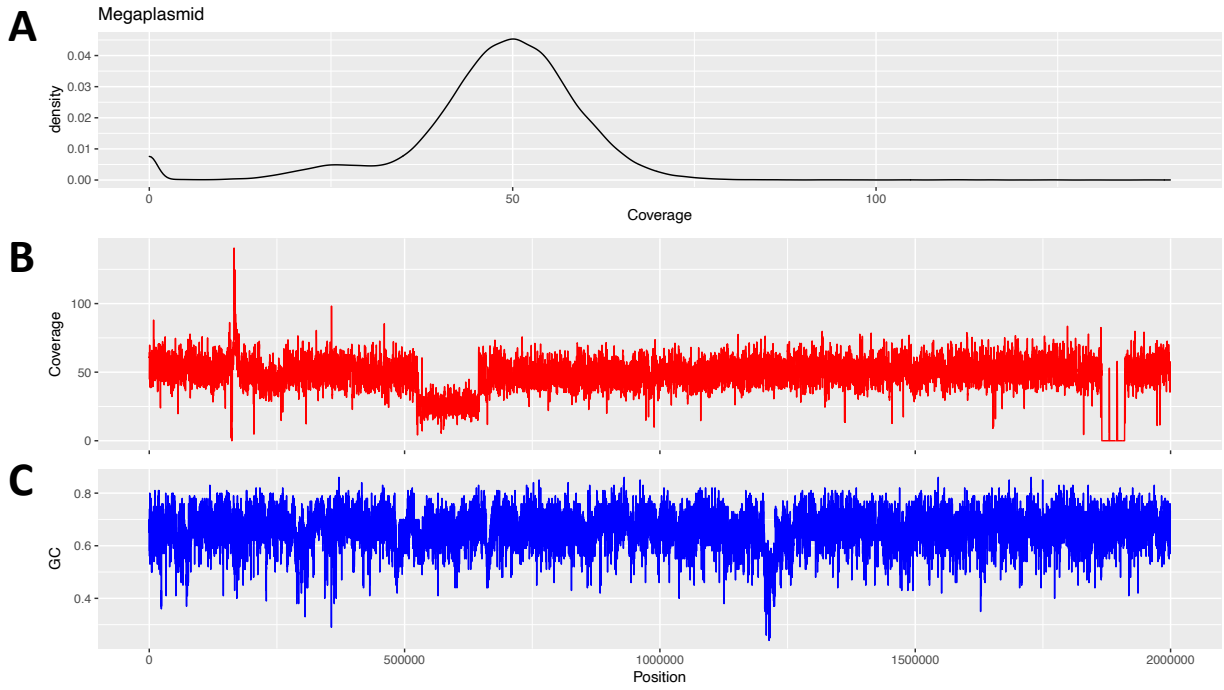
IS elements are showing movement with specific insertion sites regarding the location of isolation. Neutral genetic changes are expected to reflect some geographical sepration and the signature of geography found in the insertion sequence movement. Therefore, this type of analysis could detect neutral population structure apperin sooner tham traditional methods of detecting seqeunce variation such as variant calling analysis based on the core gene polymorphisms. Furthermore, this result aligns with the duplications and insertion sequence movement detected in the Netherlands' river water study of phylotype IIB-1 only over two years (Stevens and van Elsas, 2010). ISs disrupting transcriptional regulators may have a major effect on *Ralstonia solanacearum's* biology as the species is known to regulate major virulence pathways and disruptions in single genes can lead to avirulent phenotypes.

In the Netherlands, the ISRso3 was inserted in the phcA region. The phcA is the master regulator gene in *Ralstonia solanacearum* virulence and can lead to complete switching-off of the virulence network. Producing virulence factors is extremely costly and it is linked with a trade-off with growth. In *Ralstonia solanacearum* there is an experimentally proven trade-off between metabolic activity and growth (Peynard et al., 2016). Here, we show that the 2 populations existing in time (Pop2 and Pop3) differ in 20 genes associated with metabolic function and have lost genes associated with virulence present in Pop1. Perhaps after the initial outbreak, the strains needed to return to the river water and no longer need the exopolysaccharides for virulence so they can afford to lose them. Two genotypes with differing levels of metabolic capacity (Pop2 and Pop3) persisted in time. Therefore, insertion sequence movement provides the species with the ability to randomly generate fast-growing non-virulent clones that can be more competitive or pathogen resistant. Moreover, the lack of core genetic variation in phylotype IIB-1 is not due to the inability of the strain to evolve as evolution is observed in real-time when the strain is subjected to stress in laboratory conditions. Alderley et al. (2022) showed that phylotype IIB-1 isolate from the UK river water can rapidly evolve tolerance to antimicrobial plant allelochemicals. They showed that the tolerance was linked to the movement of insertion elements movement at into genes associated with stress responses, cell growth and competitiveness (Alderley et al., 2022). Thus, experimentally confirming the idea that the *Ralstonia solanacearum* phylotype IIB-1 strain evolves rapidly against stress after the stresses mobilise the movement into insertion sequences.

In conclusion, we found remarkably little core genome variation within the UK population of phylotype IIB-1 *Ralstonia solanacearum* with low mutation rate per year. However, variation in the accessory genome splits the population into three clusters related to gene loss associated with insertion sequence movement. We, therefore, conclude that there is an observable stronger effect of accessory genome loss compared to mutation accumulation within our sample *Ralstonia solanacearum* from the UK. The differences in accessory genes and insertion sequences are consistent with rapid accessory gene gain or loss and thus with the idea that the evolution of bacterial genomes is a highly dynamic process that involves extensive gain and loss of genes, with turnover rates comparable to, if not exceeding the rate of nucleotide substitution (Iranzo et al., 2019). Systems genetics approaches, such as genome-

scale transposon mutagenesis, would contribute to our understanding of the pathogenesis of this species. Overall, we have no means of estimating whether the presence or absence of accessory genes is adaptive from this data. However, we are seeing signs of some variation in the genomes in time, whether adaptive or neutral. Overall, the seven outbreaks in the UK show that prevention has failed on multiple occasions and control efforts need to be constantly improved. Genetic monitoring of pathogen populations could provide early warnings of the acquisition of antimicrobial resistance genes or virulence genes, such as Type 3 effectors recognised by plant immune systems (Straub et al., 2021a).

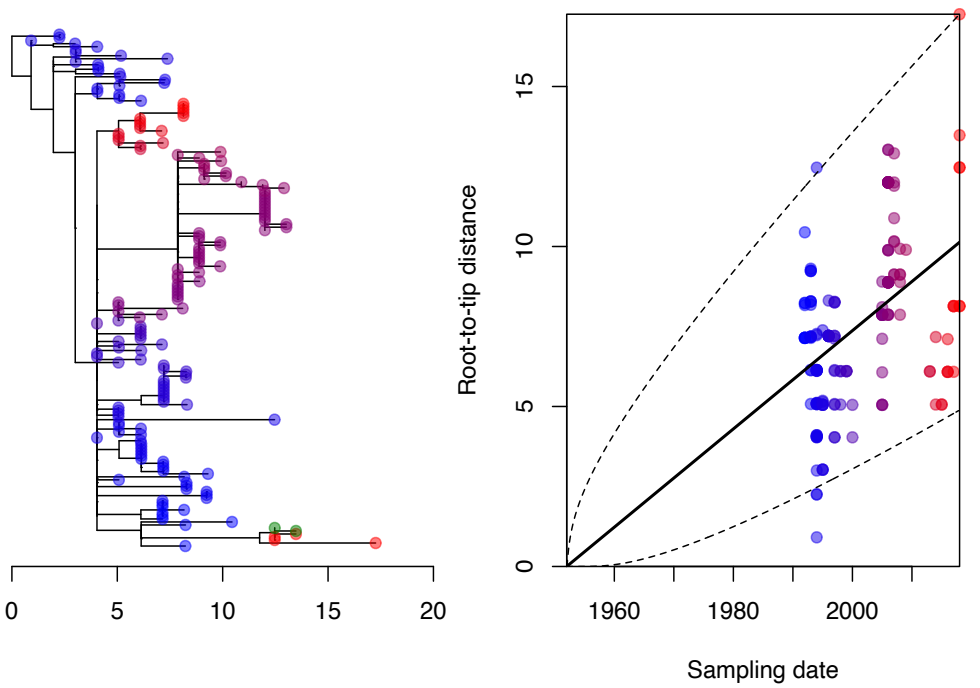
3.6. SUPPLEMENTARY



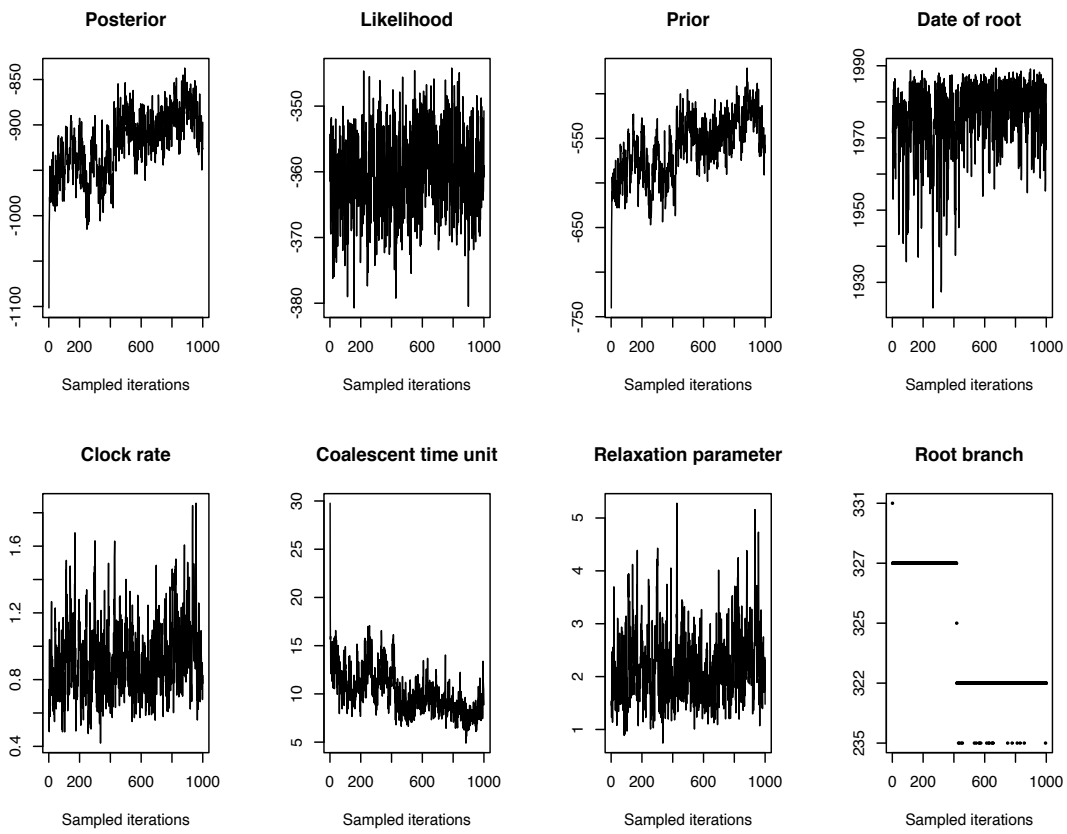
Supplementary Figure 6. **YO199 (NCPPB 3854) genome coverage plot against UY031**. Plot showing the Megaplasmid region missing from the UK isolate YO199 compared to UY031. a) Density plot of coverage across the genome. b) Coverage per 100 bp across the genome. c) GC per 100 bp across the genome (See supplementary for all genomes coverage plots.).

A

Rate=1.54e-01,MRCA=1952.00,R2=0.21,p<1.00e-04



B



Supplementary Figure 8. **Bactdating Quality Control**

A. Root-to-tip regression analysis. Simple test to estimate if there is significant correlation between time (year of isolation) and phylogenetic distance (distance from root to branch tip) in the phylogeny. B. Algorithm convergence plots. Root estimates and clock rate estimate convergence plots for the Markov chain Monte Carlo. 1000 iterations shown from the total of 100,000 iteration ran MCMC with BactDating.

Alignment (Sample size)	**Mu**	**Sigma**	**Alpha**
UK (N = 170)	102.3008	108.2189	129.1338

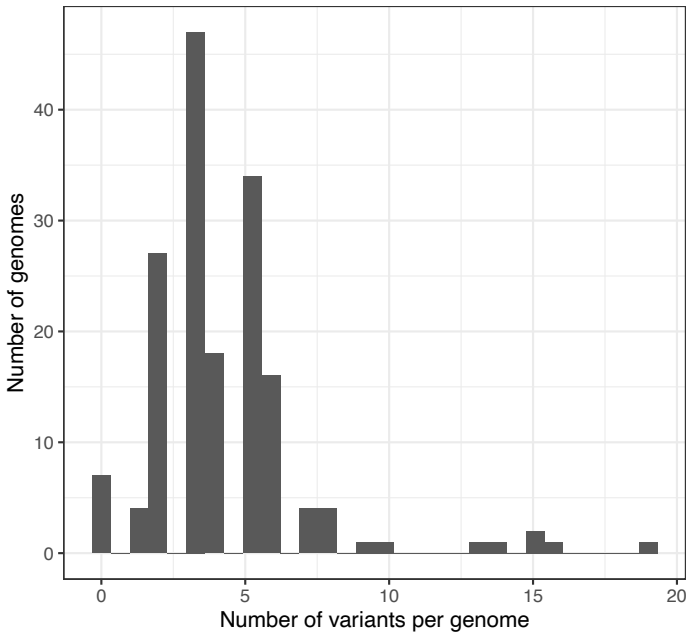
Supplementary Table 3. **Bactdating MCMC estimates tested with Coda**

Sampling of the parameters for Mu, Sigma and Alpha. Parameters need to reach 100 for the model to be valid.

Variant	Gene name	NCBI Protein ID	Allele frequency	P value (corrected)	Beta	Beta-std-err
CP012687_293375_A_G	rhsC_3	ALF86662	1.37E-01	2.50E-04	9.28E+00	1.59E+00
group_2086	lgrD		8.87E-01	7.85E-03	-7.98E+00	1.89E+00
group_1877	IS5 family transposase IS1021		2.32E-01	4.50E-07	-6.41E+00	1.32E+00
group_1933	IS5 family transposase IS1021		2.14E-01	2.16E-10	-7.33E+00	1.34E+00
group_1905	IS5 family transposase IS1021		1.96E-01	7.76E-09	-7.03E+00	1.40E+00

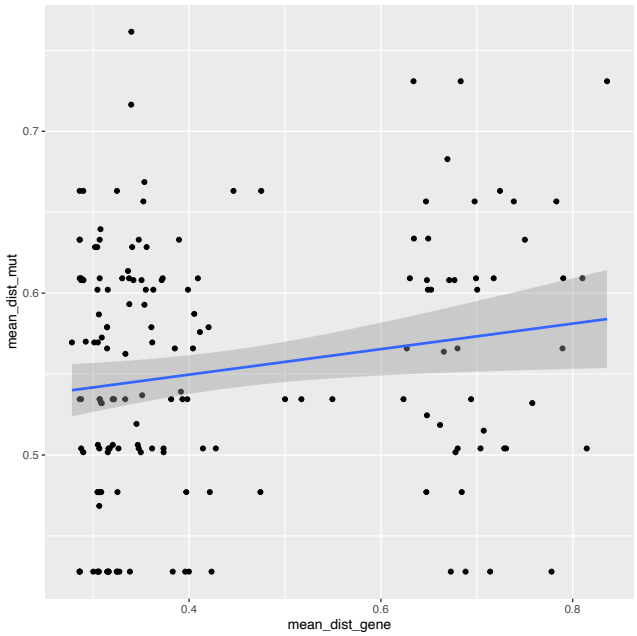
Supplementary Table 4. **Temporal GWAS.**

The table shows results for VCF and gene presence-absence (panaroo) results. Only results showed that pass p-value significance filtering. The variant column indicates either the location of the variant on the UY031 genome or the name of the gene as identified by panaroo.



Supplementary Figure 9. **Variants per genome against YO199**

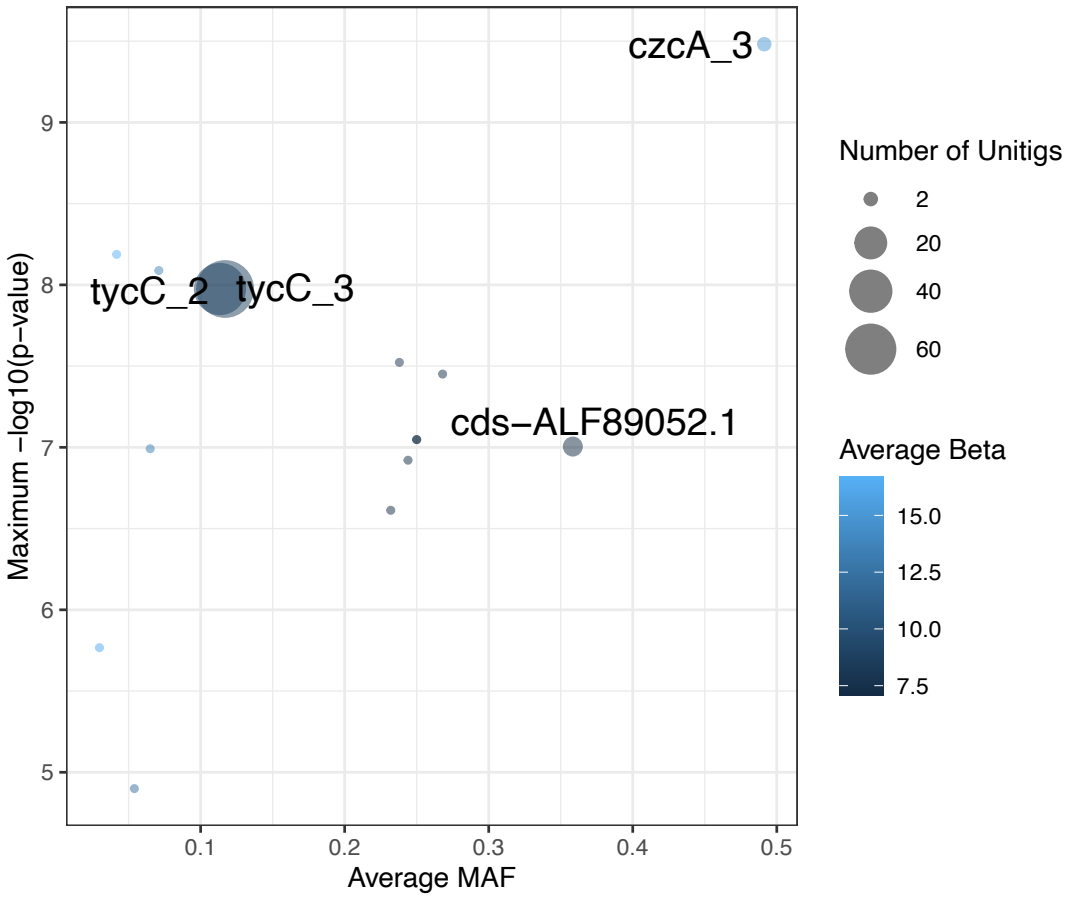
169 UK genomes variants were called against YO199 isolate from 1992 outbreak. The number of variants per genome are shown as a histogram of the count of the genomes that have that many variants.



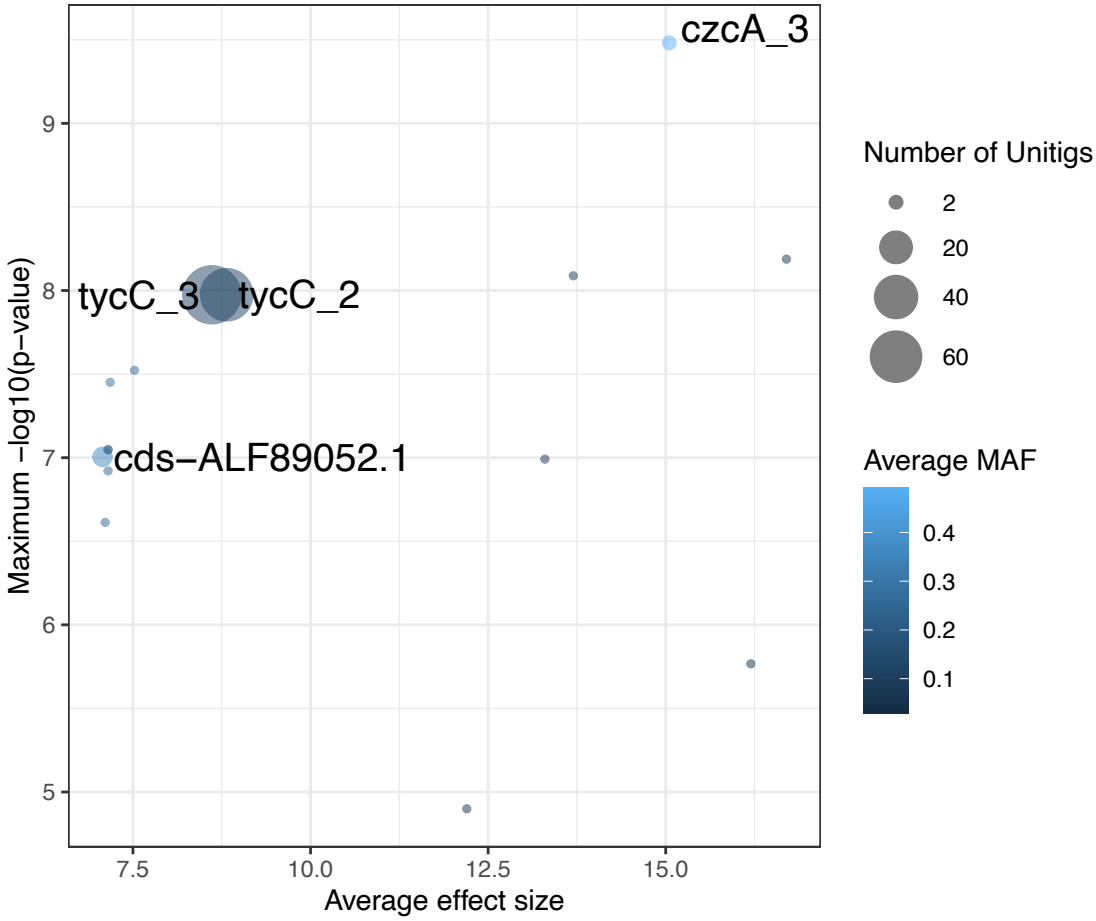
Supplementary Figure 10. Correlation between core genome distance and accessory genome distance

Jaccard distance based on presence absence used for both estimates. Mean distance for mutations based on core genome SNPs presence/absence and averaged over each pairwise relationship. Mean distance for genes based on the accessory genome (between 90% and 10% shared genes) presence/absence and averaged over each pairwise relationship.

A


czcA_3
tycC_2
tycC_3
cds−ALF89052.1
Maximum −log10(p−value)
Average MAF
9
8
7
6
5
0.1
0.2
0.3
0.4
0.5
Number of Unitigs
2
20
40
60
Average Beta
15.0
12.5
10.0
7.5

B


czcA_3
tycC_3
tycC_2
cds−ALF89052.1
Maximum −log10(p−value)
Average effect size
9
8
7
6
5
7.5
10.0
12.5
15.0
Number of Unitigs
2
20
40
60
Average MAF
0.4
0.3
0.2
0.1

Supplementary Figure 11. **Temporal GWAS with Unitigs**

Unitigs (k-mers) of variable sequence in the UK *Ralstonia solanacearum* phylotype IIB-1 population against time (year of isolation) GWAS results are shown in the figure A. Average effect size (Beta) (x-axis) against Log transformed corrected p-value for the k-mer. B. Minor allele frequency (MAF) (x-axis) against Log transformed corrected p-value for the k-mer.

Gene	**Annotation**
group_2087	hypothetical protein
group_1923	IS5 family transposase IS1021
group_1922	IS5 family transposase IS1021
group_1890	IS5 family transposase IS1021
group_1888	IS5 family transposase IS1021
group_1883	IS5 family transposase IS1021
group_1878	IS5 family transposase IS1021
group_1862	IS5 family transposase IS1021
group_1584	hypothetical protein
xcpT_2	Type II secretion system protein G
group_1214	hypothetical protein
group_1108	hypothetical protein
group_1092	hypothetical protein
group_1058	hypothetical protein
group_707	hypothetical protein
ribN	Riboflavin transporter
vgb_2	Virginiamycin B lyase
group_451	hypothetical protein
group_1921	IS5 family transposase IS1021
group_1886	IS5 family transposase IS1021

epsE_2	Type II secretion system protein E
sctC_2	Type 3 secretion system secretin
group_79	hypothetical protein
group_9	hypothetical protein
pemB	Pectinesterase B
epsF_3	Type II secretion system protein F
group_327	NADH oxidase

Supplementary Table 5. **Genes unique to Pop1**

Gene	Annotation
group_1913	IS5 family transposase IS1021
treZ	Malto-oligosyltrehalose trehalohydrolase
glgX	Glycogen operon protein GlgX
group_1395	hypothetical protein
group_1202	putative HTH-type transcriptional regulator
ycgJ	putative methyltransferase YcgJ
calA	Coniferyl-alcohol dehydrogenase
group_720	Hydroxycinnamoyl-CoA hydratase-lyase
menE_2	2-succinylbenzoate-CoA ligase
group_627	4-sulfomuconolactone hydrolase
pcpR_5	PCP degradation transcriptional activation protein
group_576	hypothetical protein
pobB	Phenoxybenzoate dioxygenase subunit beta
nlhH_2	Carboxylesterase NlhH
ddmC	Dicamba O-demethylase oxygenase component
naiP	Putative niacin/nicotinamide transporter NaiP
vdh	Vanillin dehydrogenase
xylC	Benzaldehyde dehydrogenase [NAD(+)]
mopR	Phenol regulator MopR
malQ	4-alpha-glucanotransferase
treY	Maltooligosyl trehalose synthase

Supplementary Table 6. **Genes missing from Pop3**

4. Multiple *Ralstonia solanacearum* lineages coexist in agricultural monocultures in China

4.1. ABSTRACT

Ralstonia solanacearum bacteria are globally important and highly diverse phytopathogens responsible for brown rot and wilt diseases on several plants. While *R. solanacearum* diversification can be explained by adaptations to different climates and plant hosts over time, it is poorly understood over shorter evolutionary time scales in agricultural systems. Here, we studied the phenotypic and genetic diversity of *R. solanacearum* within and between four tomato fields in China using a combination of genomics, phenotyping and physicochemical metadata collected at the level of individual plants. By phenotyping 1152 *R. solanacearum* isolates from the four fields, we show that *R. solanacearum* is highly variable showing a binomial trait distribution within each field, indicative of the presence of two coexisting ecotypes. Based on comparative genomics of a subset of 96 isolates (24 per field), we show that isolates split into two clades, which are further diversified into eight clonal groups that vary regarding their core (SNPs) and accessory genome content. Crucially, isolates from both clades coexist in 3 out of 4 fields, and more detailed analyses revealed that coexisting clonal groups differ regarding virulence traits and genes associated with type three effector proteins, quorum sensing, motility and iron-scavenging siderophores. In addition, microbial GWAS identified links between siderophore production and the IS5 family transposase IS1420, between growth and type III effector protein, and between *in planta* virulence and a RidA family protein. While physicochemical and biotic (bacterial community composition) properties clearly differed between the fields, they were not associated with the presence of specific clonal groups. Together, our results demonstrate that plant pathogenic *R. solanacearum* bacterium harbours high levels of diversity in agricultural ecosystems, which could potentially be attributed to character displacement between coexisting clonal groups.

4.2. INTRODUCTION

In agricultural ecosystems, target biocides, genetically homogenous resistant plant cultivars and crop rotations can be strong selection pressures leading to the evolution of virulence and resistance in plant pathogens. The evolution can be rapid if standing genetic diversity is present in the pathogen population. In contrast, the host population lacks the capacity to respond quickly, as agricultural breeding techniques usually lead to a lack of genetic diversity in commercial plant lines. In addition, the plant monocultures are planted densely, increasing the transmission efficiency of a pathogen across a field. In this environment, clones with a competitive advantage, such as high virulence or resistance to a control method, can quickly increase in frequency and sweep through the pathogen population. In theory, after such a clonal expansion event, the pathogen population should be rid of variation until either a new selection pressure leads to another sweep or the population survives for long enough to accumulate standing variation (Bargués-Ribera and Gokhale, 2020). However, in practice, plant pathogens often consist of several genotypes capable of infecting different or the same hosts but in different climatic conditions worldwide (Langner et al., 2021; Morris and Moury, 2019). The diversity observed within the pathogen species results from the differential long-term selection pressures acting on the pathogen. The long-term pressures could be highly heterogeneous, especially for pathogens capable of surviving in environmental reservoirs where conditions are less controlled compared to agricultural settings. In the short term, there will be a strong selection for traits contributing to within-host fitness like virulence, but over the long term, survival in the outside-host environment would select for traits allowing the pathogen to persist and compete within the environment until a susceptible host arrives. For opportunistic pathogens, it is a balance between the evolution of traits needed for host-pathogen interactions and traits required for the survival and competitive ability of the pathogen in the environment.

Different spatial scales with temporal crop rotations and seasonality make agricultural fields and soils heterogeneous or fluctuating environments. Therefore, selecting for local adaptation affects the fitness of pathogens due to differences in the deployment of resistance genes and seasonal temperatures (Croll and McDonald, 2017; Velásquez et al., 2018). Many

studies have investigated the global population diversity of different plant pathogens; however, it remains unclear to what extent pathogen variation is present at the local scale, such as within and between agricultural fields in the same country (Dietzgen et al., 2020; Sharma et al., 2021b). Despite the homogeneity of a plant monoculture, the soil within a field remains a heterogeneous environment providing multiple micro niches and pressures from viruses and competitors (Messiha et al., 2009). High-standing variation means that a pathogen population is more stable and can respond to environmental change. Consequently, plant resistance's longevity and antimicrobials' efficacy can become compromised (Lurwanu et al., 2021; Nelson et al., 2018; George W. Sundin and Wang, 2018). This threatens food security by making plant diseases harder to control, which in turn can result in considerable crop losses globally (Savary et al., 2019b). Therefore, identifying the genetic diversity of plant pathogens at the agricultural field scale can provide key insights into the mechanisms of their local adaptation and pave the way to designing effective strategies for controlling plant disease.

Ralstonia solanacearum is a prime example of an opportunistic plant pathogen characterised by high genetic diversity globally. It is able to infect multiple domesticated and wild plants, and can persist in the soil as a saprophyte making it an opportunistic plant pathogens, which has been adapted across different environments (Elphinstone, 2005; Genin, 2010; Genin and Denny, 2012b; Hayward, 1991b). Due to high genetic diversity, *R. solanacearum* is classified as a species complex (RSSC), which can be divided into four phylotypes based on geographical origin (M. Fegan and Prior, 2005), multiple biovars based on host range (Hayward, 1991b) and phenotypic characteristics (Buddenhagen, 1962). Currently, it is classified as three phylogenetic species (Safni et al., 2014b): *Ralstonia solanacearum*, *Ralstonia pseudosolanacearum* and *Ralstonia syzygii*. RSSC strains from phylotype I are widespread in China, present in most crop production regions and capable of infecting nearly a hundred local crop plants (Jiang et al., 2017). *R. solanacearum*-caused wilt in tomatoes can lead to yield loss from 35% to 90% under favourable high temperatures and high moisture conditions, such as in wet and warm tropical regions like the ones in a big part of China (Singh et al., 2015).

The genetic variation within the *R. solanacearum* species complex has been linked to several adaptations that can increase its fitness and survival under stressful and variable environments (Genin, 2010; Messiha et al., 2009; Scherf et al., 2010). For example, *r*esistant

tomato cultivars have been shown to become less resistant over time in the field, suggesting that the pathogen might be able to evolve to become more virulent (Singh et al., 2015). Several environmental conditions have been associated with *R. solanacearum* fitness (Álvarez et al., 2008; Caruso et al., 2005b; Jiang et al., 2021b), including growth (Kadam and Jagtap, 2018), tolerance to cold temperatures, ability to resist metal and salt stress (Um et al., 2013), nutrient availability (Gu et al., 2020a; Li et al., 2021; Yang et al., 2019, 2018), and antagonistic and facilitative bacterial interactions (Hu et al., 2017; Li et al., 2022, 2019; Wei et al., 2019, 2015b). Furthermore, sequevars have been shown to vary in their intraspecific competitiveness in cell cultures, rhizosphere, and within tomato xylem (Huerta et al., 2015). While the genetic and phenotypic diversity of *Ralstonia solanacearum* is recognised at the global (G. Cellier and Prior, 2010; Wicker et al., 2012b), country (Albuquerque et al., 2014; Ramírez et al., 2020; Xue et al., 2011) and regional level (Chesneau et al., 2017; Deberdt et al., 2014; Wicker et al., 2009), it is less well understood at the local scale within plants, fields and crop production areas (Grover et al., 2006). Specifically, while *R. solanacearum* diversity has been observed in natural populations (M. Fegan and Prior, 2005; Genin and Denny, 2012b), it is unclear which environmental factors drive this variation and if it can be associated with certain pathogen ecotypes in agricultural fields. Understanding small-scale geographic variation can help us understand the movement of *Ralstonia solanacearum* genotypes across fields and regions of one country and help us prevent future outbreaks not only in China but potentially within Europe, where *R. solanacearum* strains are quarantined pathogens (EPPO, 2018).

To compare the phenotypic and genetic diversity of *Ralstonia solanacearum* isolates within and between four tomato fields in China, we sampled 12 plants from each field located in four provinces in China (Figure 20 A): Nanjing (NJ) of Jiangsu province, Ningbo (NB) of Zhejiang province, Nanchang (NC) of Jiangxi province and Nanning (NN). We isolated 24 *Ralstonia solanacearum* clones per plant, resulting in a total of 1152 clones, which were all phenotyped regarding biofilm, maximum growth, and competitiveness. To quantify and compare genetic variation, 3 clones from 8 plants per field were sequenced and phenotyped more extensively (total of 96 clones). We also collected physicochemical and biotic (bacterial community composition and diversity) for each of these plants to investigate if genetic variation could be associated with certain local environmental conditions. We then investigated: i) the

biodiversity and physicochemical properties of the plants' rhizosphere, ii) the extent of genetic and phenotypic variation in *R. solanacearum* pathogen between the sampled fields, ii) the associations between genetic and phenotypic variation through genome-wide association study, iii) the diversity of virulence-associated genes in the *R. solanacearum* genotypes in the four fields. We found that pathogen diversity varied within and between fields in China. We found coexisting *R. solanacearum* genotypes within the same field that differ genetically and phenotypically, which can be explained by niche differentiation in ecologically important life-history traits and due to the wide resource utilisation spectrum of *R. solanacearum* strains. Together, these results suggest that *R. solanacearum* adapts to the agricultural environment through niche differentiation, leading to the coexistence of genetically and phenotypically distinct 'ecotypes' within individual fields.

4.3. METHODS

4.3.1. Tomato fields sites and collection of plant soil samples

Rhizosphere soils of tomatoes were sampled at four geographically disconnected fields in Nanjing (118°57' E, 32°03' N) of Jiangsu province, Ningbo (121°67' E, 29°91' N) of Zhejiang province, Nanchang (115°51'E, 28°41'N) of Jiangxi province and Nanning (108°21'E, 22°49'N) of Guangxi province between 26th of May and 13th of June in 2015. In each field, 12 tomato plants were randomly collected, resulting in a total of 48 rhizosphere samples (Figure 20 A). The excess soil was gently removed from the roots by shaking and the remaining soil attached to the surface of the roots was considered the rhizosphere soil (Wei et al., 2011). For microbial community composition comparisons, 5 g of rhizosphere soils were cryopreserved at -80 °C in 10 mL Falcon tubes containing 5 mL of 30% glycerol, while the rest of the rhizosphere soil samples were used to determine soil physicochemical (abiotic) properties as described later. From each of the 12 rhizosphere samples collected per field, 24 *R. solanacearum* single colony isolates were randomly selected using an *R. solanacearum*-specific semi-selective medium (SMSA) (Elphinstone et al., 1996) and cryopreserved on 96-well plates at -80 °C in 100 µL of nutrient medium (NB) with 15% glycerol.

4.3.2. Quantification of abiotic rhizosphere soil properties

Abiotic physicochemical properties included pH, soil moisture content (Moisture, %), electric conductivity (EC, us·cm^{-1}), soil organic matter (SOM, g·kg^{-1}), total nitrogen (TN, mg·kg^{-1}), total phosphorus (TP, mg·kg^{-1}), available nitrogen (AN, mg·kg^{-1}), available phosphorus (AP, mg·kg^{-1}) and available potassium (AK, mg·kg^{-1}). Soil physicochemical properties were mainly measured as described previously (Jiang et al., 2021b) following the Chinese standard for soil property determination: Soil pH (HJ 962-2018), EC (HJ 802-2016), SOM (NY/T 1121.6-2006), TN (LY/T 1228-2015), TP (GB/T 9837-1988), AN (LY/T 1229-1999). The difference in fresh and air-dried soil sample weight was used as a proxy for soil moisture for each rhizosphere sample. AP and AK were extracted with hydrochloric acid and ammonium fluoride and measured using the molybdenum blue method (Scrimgeour, 2007). The soil pH (Li et al., 2017) was measured in a soil-to-water 1: 5 suspension using a pH meter (PB-10, Sartorius, Germany). AP was extracted using hydrochloric acid and ammonium fluoride and determined using the molybdenum blue method (Scrimgeour, 2007). The total C and N were measured using a multi C/N analyser 3000 (Analytik Jena AG, Germany) as described previously (Scrimgeour, 2007).

4.3.3. Quantification of biotic rhizosphere soil properties

Bacterial community composition was calculated at the operational taxonomic unit (OTU) level within the 48 sampled plants. The total DNA was extracted from ~0.25 g of cryopreserved rhizosphere soil using PowerSoil DNA Isolation Kit (Mobio Laboratories, Carlsbad, CA, USA) following the manufacturer's protocol. DNA quality and concentration were checked using a NanoDrop 1000 spectrophotometer (Thermo Scientific, Waltham, MA, USA). Soil DNA was then subjected to 16S ribosomal RNA (rRNA) Illumina amplicon sequencing to determine the diversity and composition of bacterial communities at Shanghai Biozeron Biological Technology Co. Ltd. The V4 hypervariable region of the 16S rRNA gene was amplified with the primer pair 563F (5'-AYTGGGYDTAAAGVG-3') and 802R (5'-TACNVGGGTATCTAATCC-3'). All sequences were processed with QIIME (Caporaso et al., 2010). The OTU similarity cut-off was assigned at 97% identity level using USEARCH (Edgar, 2010). OTUs were assigned to corresponding bacterial taxa using the Ribosomal Database Project (RDP) database with the online version of the RDP classifier (Cole et al., 2014).

Microbial community composition was quantified as a dissimilarity index (Bray-Curtis) based on the average Bray-Curtis distance of each sample from each other at the OTU level. Bacterial community composition comparisons were done using the R vegan package (Dixon, 2003).

4.3.4. Phenotypic analysis

For the phenotypic analysis, all the 1152 *Ralstonia solanacearum* isolates collected from the 12 plants were used (4 field sites x 12 plants x 24 bacterial isolates; Figure 21 a). All isolates were characterised phenotypically for the following pathogen life-history traits: maximum growth rate (growth), biofilm formation (biofilm) and intra-species competition (competitiveness - against a type *R. solanacearum* strain isolated from Nanjing (Wei et al., 2011)). Each experiment was replicated three times unless stated otherwise. Furthermore, 96 isolates were further phenotyped for virulence traits siderophore production, exopolysaccharide production (EPS), *in planta* virulence (disease incidence) and swimming motility. Before virulence traits assays, the overnight liquid culture of *R. solanacearum* clones were washed and suspended in 0.85% NaCl (107 CFU/ml).

4.3.4.1. Maximum growth rate

Two µL of overnight *R. solanacearum* cultures of each isolate (revived as described earlier) were inoculated to 96-well microplate wells containing 198 µL of NA medium per well. Bacterial growth was assessed as changes in optical density (OD) at 600 nm measured every two hours for 24 h at 30°C using a SpectraMax M5 Plate reader (Molecular Devices, Sunnyvale, CA, USA). The maximum growth rates (at log phase) and biomasses were determined for each strain using the gcFitModel function to fit growth data in grofit (Kahm et al., 2010) package in R.

4.3.4.2. Biofilm formation

The biofilm formation was quantified using a modified crystal violet method (Yao and Allen, 2007). Briefly, two µL of overnight *R. solanacearum* cultures were inoculated to 96-well microplate wells containing 90 µL of NA medium. After incubation for 48 h at 30 °C without shaking, 200 µL of methanol was used to fix bacteria for 15 min. Microplates were then emptied and left to dry at room temperature before adding 25 µL of 1.0% crystal violet solution. After 10 min of staining, the unbound crystal violet was gently removed with a

pipette. The wells were then washed with distilled water, 70% ethanol and once more with distilled water. After the plates had been air-dried, the remaining crystal violet bound to adherent cells was re-solubilised with 150 µL 33% (v/v) glacial acetic acid. The biofilm was quantified as optical density (OD_{570}) using SpectraMax M5 spectrophotometer (Molecular Devices, Sunnyvale, CA, USA).

4.3.4.3. Intra-species competition (competitiveness)

The level of intra-species competitiveness of isolated pathogen strains was determined in direct co-culture competition assays with red fluorescent labelled *R. solanacearum* (mCherry-tagged Rs-RFP) type strain (Wei et al., 2011). To quantify differences, the growth of red fluorescent labelled QL-Rs1115-RFP (mCherry, excitation: 587 nm; emission: 610 nm) strain was measured in the absence and presence of *R. solanacearum* field isolates (in 50:50 starting densities, ~10^6 cells per mL for both strains) in 96-well microplates for 24 h at 30 °C. The intra-species competitiveness was measured as the difference in the growth of fluorescent QL-Rs1115-RFP strain in the absence (OD_{600a}) and presence (OD_{600p}) of *R. solanacearum* isolates using the following formula:

$$\text{Competitiveness} = \frac{OD_{600a} - OD_{600p}}{OD_{600a}} \times 100.$$

4.3.4.4. Siderophore production

Siderophore production was assayed using a modified version of the universal chemical assay developed by Schwyn and Neilands (Gu et al., 2020b; Schwyn and Neilands, 1987). Briefly, we used the liquid version of the CAS assay, where 100 ul cell-free supernatant (three biological replicates for all 2,150 soil isolates) or deionized water as a control reference were added to 100 ul CAS assay solution in a 96-well plate. After 2 h of static incubation at room temperature, the OD630 of the cell-free supernatants (A) and deionized water controls (Ar) was then measured using a plate reader (SpectraMax M5) at room temperature. Siderophores induce a colour change in the CAS medium, which lowers the OD630 measurements, and siderophore production can thus be quantified using the following formula: 1 − A ÷ Ar.

4.3.4.5. Disease incidence

The isolate of *R. solanacearum* clones was inoculated into the pot using a soil drenching method resulting in a final concentration of 5.0×10^6 CFU·g−1 of soil (Wei et al., 2011). The disease development was monitored daily and quantified by the proportion of wilted leaves per plant for the 0–4 scales disease index: 0 = no wilted leaf, 1 = <25% wilted leaves, 2 = 25-50% wilted leaves, 3 = 50-75% wilted leaves and 4 > 75% wilted leaves or dead plant (Schandry, 2017). The area under the curve progression of the disease dynamics curve (AUDPC) is used as a measure of the overall severity of wilt disease referred to as disease incidence. The disease dynamics curves were fitted individually for each plant using the gcFitModel function in the R grofit-package (Kahm et al., 2010).

4.3.4.6. Swimming motility and EPS

Swimming motilities were determined on CPG agar plates at 30°C containing 0.3% (w/v) agar, respectively. The zone of migration was measured in four directions after two days for the swimming motility (Raza et al., 2016) and the extracellular polysaccharides (EPS) production assay (Dubois et al., 1951). For the EPS liquid culture samples of *R. solanacearum* were inoculated (20 ul) in fresh CPG medium and after three days of growth at 30°C, extracellular polysaccharides were precipitated using ice-cold ethanol and quantified by the phenol-sulfuric acid method (Dubois et al., 1951).

4.3.5. Genome data processing

The genome sample set sequenced was from a subset of 8 plants. Three bacterial isolates were taken randomly from each plant, making up 24 isolates within each field. This is a total of 96 *Ralstonia solanacearum* genomes from tomatoes from 4 distant field sites in China, each within a different province: Nanjing (NJ), Ningbo (NB), Nanning (NN), Nanchang (NC).

4.3.5.1. Sequencing and assembly

Genomic DNA was exacted from overnight cultures of *R. solanacearum* with the Bacteria DNA extraction Kit (OMEGA) according to the manufacturer's instructions. Quality control was subsequently carried out on the purified DNA samples using TBS-380 fluorometer (Turner BioSystems Inc., Sunnyvale, CA). High qualified DNA samples (OD260/OD280 = 1.8~2.0, > 6

ug) were send to Shanghai Biozeron Biothchnology Co., Ltd. (Shanghai, China.) for genome sequencing with Illumina NovaSeq 6000. For each strain, at least 1 ug genomic DNA was used for sequencing library construction. Paired-end libraries with insert sizes of ~400 bp were prepared following Illumina's standard genomic DNA library preparation procedure. Purified genomic DNA is sheared into smaller fragments with a desired size by Covaris, and blunt ends are generated by using T4 DNA polymerase. After adding an 'A' base to the 3' end of the blunt phosphorylated DNA fragments, adapters are ligated to the ends of the DNA fragments. The desired fragments can be purified through gel-electrophoresis, then selectively enriched, and amplified by PCR. The index tag could be introduced into the adapter at the PCR stage as appropriate and we did a library quality test. After sequencing genomic data quality control and filtering was applied. Following we chose 95 genomes for the further analysis based on the total genome length of the assembly and contaminant sequence presence (1 isolate excluded in NJ field (NJ1823)). Contamination was checked with a *k*-mer alignment against a local copy of the RefSeq database using Mash provided as an accessory script in Panaroo (v1.2.4)(Tonkin-Hill et al., 2020). The genome assembly was performed using Unicycler (v0.4.9) with known contaminant sequences provided in the software (Wick et al., 2017b). Five of the assemblies still contained contaminants after this processing. As a result, an alternative method was applied to them where all reads were aligned to *R. solanacearum* GMI1000 [GCA_000009125.1] reference genome and only aligned reads were assembled with Unicycler. Draft annotations were then performed with prokka (v1.14.5) (Seemann, 2014b).

4.3.5.2. Constructing phylogeny and SNP calling analysis

Genetic distance methods were used to estimate the population structure amongst isolates and construct a phylogeny. First, single nucleotide polymorphisms (SNPs) and small indels were called against a high-quality *R. solanacearum* GMI1000 reference genome downloaded from NCBI [GCA_000009125.1] using snippy (v5.0). Then, we generated a core single nucleotide polymorphisms (SNPs) presence-absence matrix and core genome alignment of SNPs with the accessory script "snippy-core" of snippy and the variant calling files (VCFs) produced by snippy. Maximum likelihood phylogeny was then constructed using the IQtree GTR+G4 model and two different bootstrap methods (Alrt and Ultrafastboot) (Nguyen et al., 2015a). Recombination in the genome sequences was detected using 100 runs of prokaryotic recombination estimation software ClonalFrameML (Didelot and Wilson, 2015), and

recombination-aware phylogeny was constructed. Recombination calculations were conducted using the following formula for the ratio of the relative effect of recombination and mutation (r/m) for each simulation run:

$$\frac{r}{m} = \frac{R}{\theta} \times \delta \times \nu$$

$\frac{R}{\theta}$ - The ratio of recombination and mutation rates

δ - the mean length of imports

ν - the average distance of the imports

4.3.5.3. Pangenome analysis

To assess the overall gene count variation in the sample set, we performed a pangenome analysis on the genome assemblies using the Panaroo pipeline (Tonkin-Hill et al., 2020), which generated: Orthologous gene clusters, core genome alignment with ClustalO from 90% shared gene clusters between the genomes and gene cluster presence/absence matrix. The total size of the pangenome was 6575 genes, of which 2364 were classified as accessory genes (excluding genes >90% and singletons or doubletons) and 4211 as core genes (>90% shared; Supp. Fig. 2). When compared to the reference genome of GMI1000, 303 genes were missing from our pangenome dataset.

4.3.5.4. Type III effectors and prophage analysis

The 95 genomes were further analysed regarding known virulence-associated sequences: prophage elements and Type III effector proteins (T3Es) that are used to evade plant immunity (Peeters et al., 2013a). First, intact prophages were identified using Phaster (Arndt et al., 2016). Then, extracted prophage sequences' genetic distance was estimated using a *k*-mer approach provided with Mash and Mashtree software (Katz et al., 2019), and distances were compared using Euclidean clustering and Neighbour-joining. Identification of prophages was made based on queries against RefSeq Virus. In addition, prophage gene re-annotation was done using Viga (González-Tortuero et al., 2018), and Roary (Page et al., 2015) was used to determine the prophage pangenome. Second, the genomes were searched for Type III

effector (T3Es) sequences using a local installation of the Ralstonia T3E database (Sabbagh et al., 2019) with default stringent blast parameters.

4.3.5.5. GWAS analysis for linking phenotypes with underlying genetic variants

Genome-wide association study (GWAS) was performed on the seven phenotypic traits (list them here) to identify potential underlying SNPs and small indels, Unitigs and gene presence/absence from panaroo output using the pyseer microbial GWAS pipeline (Lees et al., 2018). First, all the phenotypes were centred using the base R scale function scale (centre=T, scale=F). Then, the SNPs and small indels were re-called with Freebayes using all the bam files obtained from snippy to obtain a genome-wide VCF. The VCF was filtered for sites with mean coverage between 100 – 300 bp, quality of 30, minor allele frequency (MAF) over 0.1 and below 0.90, only one alternative allele and no missing sites allowed leaving: 41,947 SNPs (~25,000 filtered out). The variants were annotated with SnpEff (Cingolani et al., 2012). The Unitigs were obtained using the unitig-counter software. Finally, the p-value threshold was obtained using a helper script within pyseer for each phenotypic trait analysed based on Bonferrorni correction. In addition, QQ plots of the p-values were made for all the phenotypes tested and visually observed for diviations of the points from the expected distribution and shelving which is a sign of population stricture. Due to the nature of the bacterial data high population structure is expected and not all of it is accounted for by the correction applied by pyseer. Therefore, we chose p-value plots that had small levels of shelving and at least the smallest values of p were falling on the theroretical line.

4.3.5.6. Post-processing and statistics

All post-processing, statistical analysis and graphics were constructed using R (v4.0.2) (R Core Team, 2017) with packages tidyverse, ggtree (Yu et al., 2018), vegan (Dixon, 2003), reshape2, ape, phytools. All plot visualisation was done with ggplot and ggtree packages (Yu et al., 2018). PERMANOVA was used to compare differences between the experimental groupings of principle component analysed data. The function adonis2 in the vegan package in R was used for the test (Dixon, 2003). R libraries factorextra and cluster were used to perform *k*-means clustering visualised in Figure 20 C and optimal cluster number (Supplementary Figure 13).

Estimates of nucleotide diversity (Pi) were done on the core genome alignment using the R library popgenome (Pfeifer et al., 2014).

4.4. RESULTS

4.4.1. Tomato plants' rhizosphere originating from four different fields have physicochemical and biodiversity differences explained by field site origin

First, we investigated the abiotic and biotic differences between the microenvironment of the tomato plants' rhizosphere *Ralstonia solanacearum* resides. We compared the rhizosphere of 12 plants from each of the four field sites studied, totalling 48 rhizosphere samples analysed (Figure 20 A). We compared the measurements for six abiotic factors: pH, AK, AP, total C, AN and Moisture, and biotic diversity using Bray-Curtis's index, including *Ralstonia solanacearum* (pathogen) density and total bacteria density. We see the clustering of the rhizosphere samples from each field site together in the principal component analysis based on the six abiotic soil properties measured (Figure 20 B, PERMANOVA: $F_{3,48} = 66$, $R^2 = 0.82$, $p = 0.001$). Also, we see field site-specific clustering on the multiple-dimension scaling plot based on the Bray-Curtis dissimilarity index (Figure 20 C). These results indicate that the biotic and abiotic environment is showing field-specific structure. The field sites are a great distance apart in four different provinces (Figure 20 A), so the presence of variable abiotic conditions is expected. However, conditions within each field also varied, and we did not observe tight clustering in either the MDS or PCA analysis. Therefore, the variance is not fully explained by the four-way clustering according to the field site.

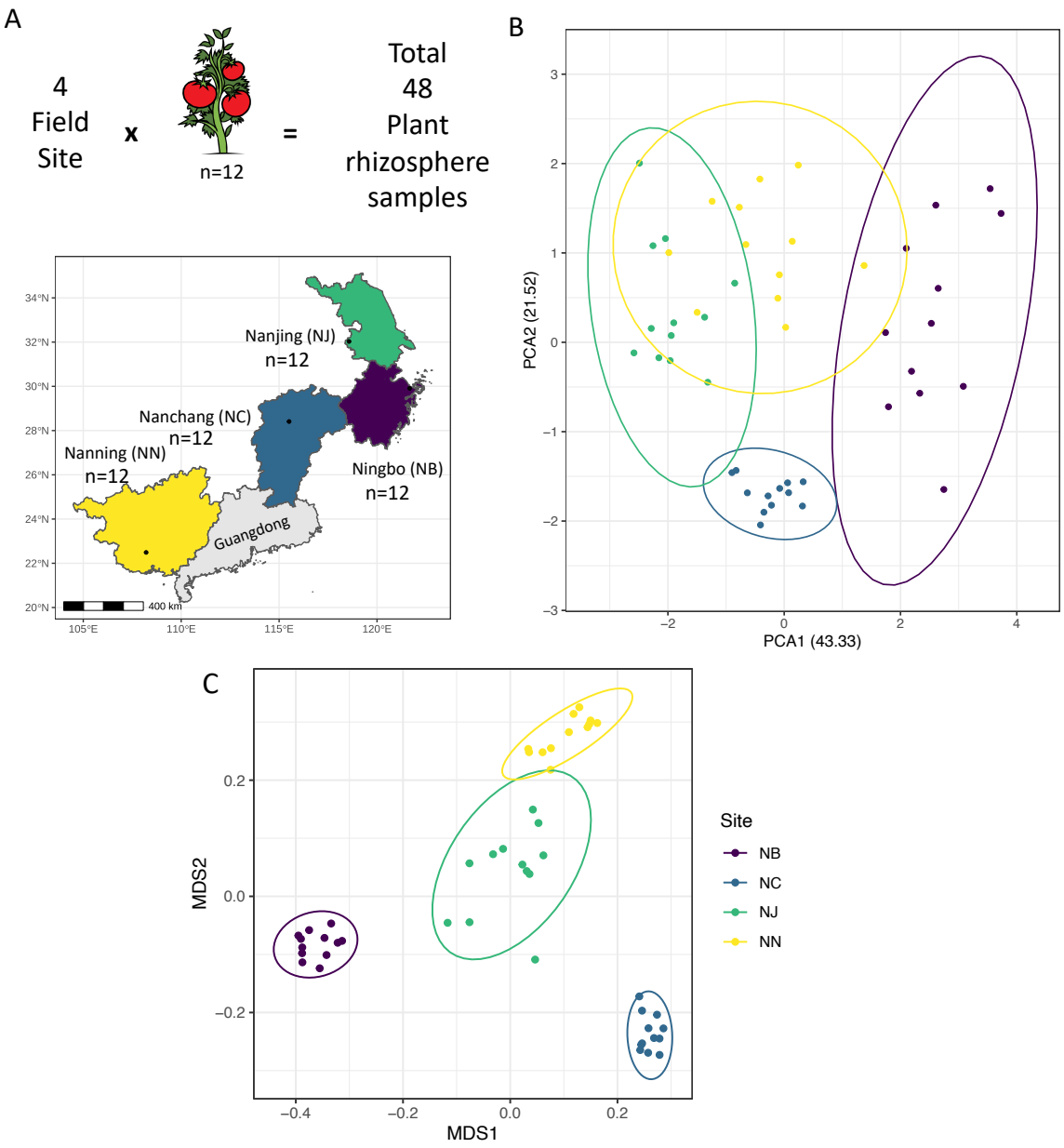


Figure 20. **Plant rhizosphere properties explained by field site**

Rhizosphere of 12 plants per fields site sampled and tested. A) Plant sampling design and Map showing plants sampled per field site; B) PCA analysis of 6 soil physicochemical properties: pH, AK, AP, total C, AN and Moisture; C) Soil biodiversity measurements with Bray-Curtis index. In all panels data points are coloured by field site.

4.4.2. Two phenotypic groups of *Ralstonia solanacearum* cooccur within each tomato field in China

To understand *R. solanacearum* variation within and between fields at the phenotypic level, we compared 1152 *R. solanacearum* regarding their growth, biofilm production and competitiveness (Figure 21 a & b). The PC analysis showed that there is large variation within and between the fields, with field identity explaining 21% of the total trait variation (Figure 21 **c**, PERMANOVA: $F_{3,1151}$ = 100, R^2 = 0.21, p = 0.001; Supplementary Figure 12). To understand this variation in more detail, we used *k*-means clustering to group our data phenotypically. The gap statistic value showed that the optimal number of clusters in the data is two (Supplementary Figure 13), indicating that our data distribution could be explained by two phenotypic groups (Figure 21 **D**). Moreover, isolates from both clusters were present in all four field sites (Supplementary Table 7).

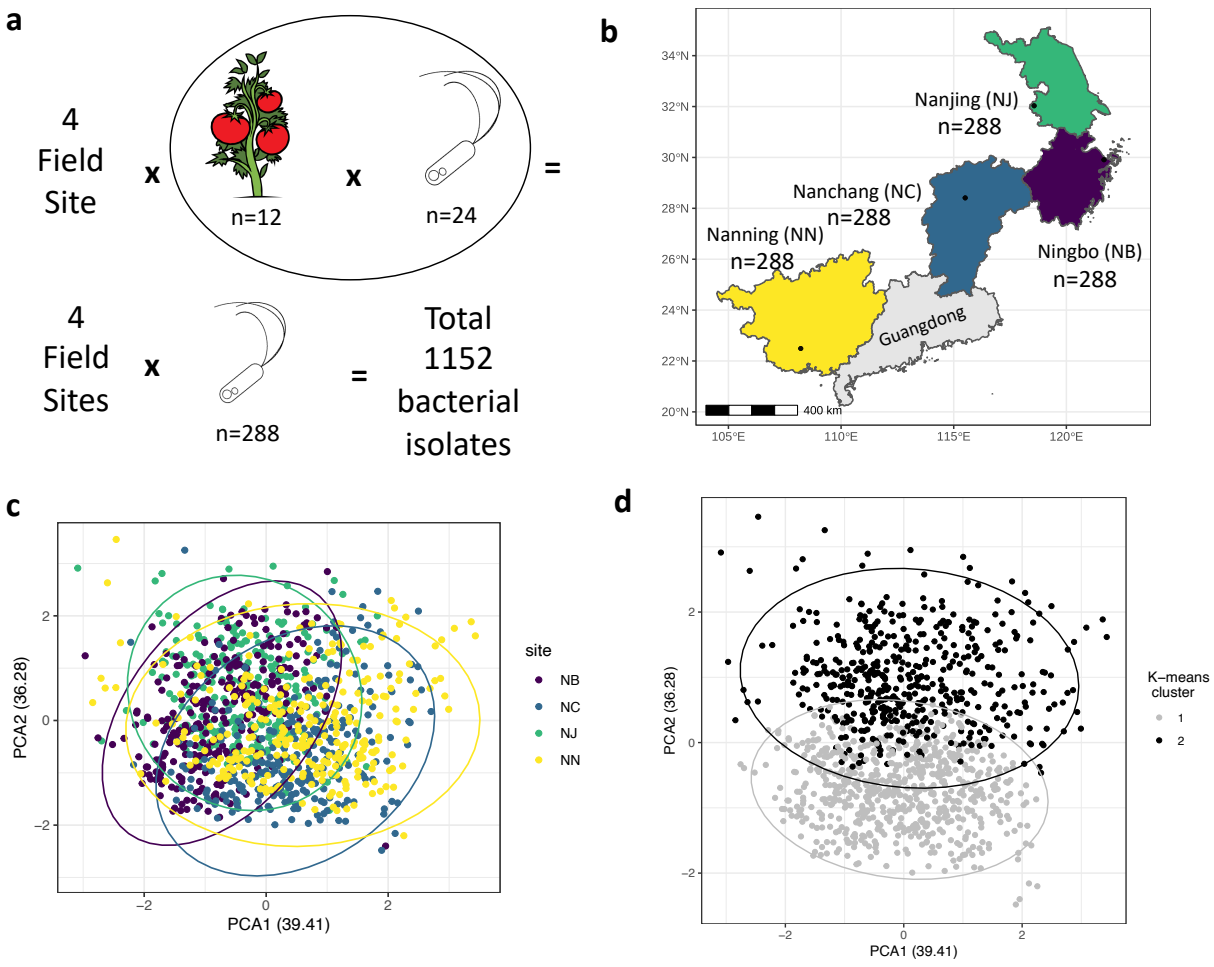


Figure 21. **Geography does not explain Ralstonia solanacearum phenotypic variation in tomato fields in China.**

The figure shows the phenotypic variation of 1152 Ralstonia solanacearum isolates from 4 Chinese fields based on three virulence-associated phenotypic traits: Competitiveness, Maximum growth rate and Biofilm formation. a) schematic of sampling design for the 1152 bacterial isolates; b) map of the field sites sampled c) PCA analysis of the three phenotypic traits coloured by sampling site; d) PCA analysis of the three phenotypic traits coloured by k-means cluster (Supplementary Figure 13 & Supplementary Table 7).

4.4.3. Two different *Ralstonia solanacearum* genotypes cooccur within tomato field sites in China

To examine the genetic diversity of *R. solanacearum* within the sampled fields we sequenced the whole genomes of 96 bacterial isolates (4 field sites x 8 plants x 3 clones; one genome excluded due to poor quality. First, we compared the 95 isolates with the *R. solanacearum* GMI1000 reference genome to compare how genetically distant our strains are from this well annotated reference genome. This showed that all the isolates are part of the *Ralstonia solanacearum* phylotype I, or *Ralstonia pseudosolanacearum* species (Safni et al., 2014b), even though the isolates were all distantly related to the GMI1000 reference, with the number of SNP differences ranging from 15,000-35,000 per isolate. Overall, the large number of SNPs observed within the sample set (48,300 variants) meant that a well-resolved phylogeny could be constructed based on the core genome SNPs. After recombinant sites were removed, a highly supported tree was constructed that split the isolates into 2 clades and a further 8 terminal branches (Supplementary Figure 19). The branches represented eight clonal lineages with little variation within each lineage (median of 1 segregating site and median of 0.2 nucleotide diversity index). Hence, from here on these branches are referred as 8 clonal lineages representative of distinct *R. solanacearum* genotypes.

To investigate the source of the sequence diversity within our genome sample set, we estimated the homologous recombination rates relative to mutation rates using the whole

genome alignment of the genomes. We used ClonalFrameML software as it can estimate recombination from outside the sample set of genomes studied. We saw a total number of 2728 predicted recombination events across the dataset. The ratio of the relative effect of recombination and mutation was predicted to be r/m = 1.87 [1.82-1.92], and the ratio of the frequency of recombination and mutation R/theta = 0.56 [0.54-0.57]. The average length of recombined fragments was estimated to be δ = 182 bp, and the average divergence between donor and recipient was ν = 0.019. Together, these results show that mutation happens twice as often as recombination. However, on average recombination contributed to longer sequence changes (δν = 3.4bp), causing almost twice the number of nucleotide substitutions compared to mutations. These results hence confirm the importance of recombination events for the diversification of *R. solanacearum* and show that the observed diversity was not a result of strictly clonal diversification.

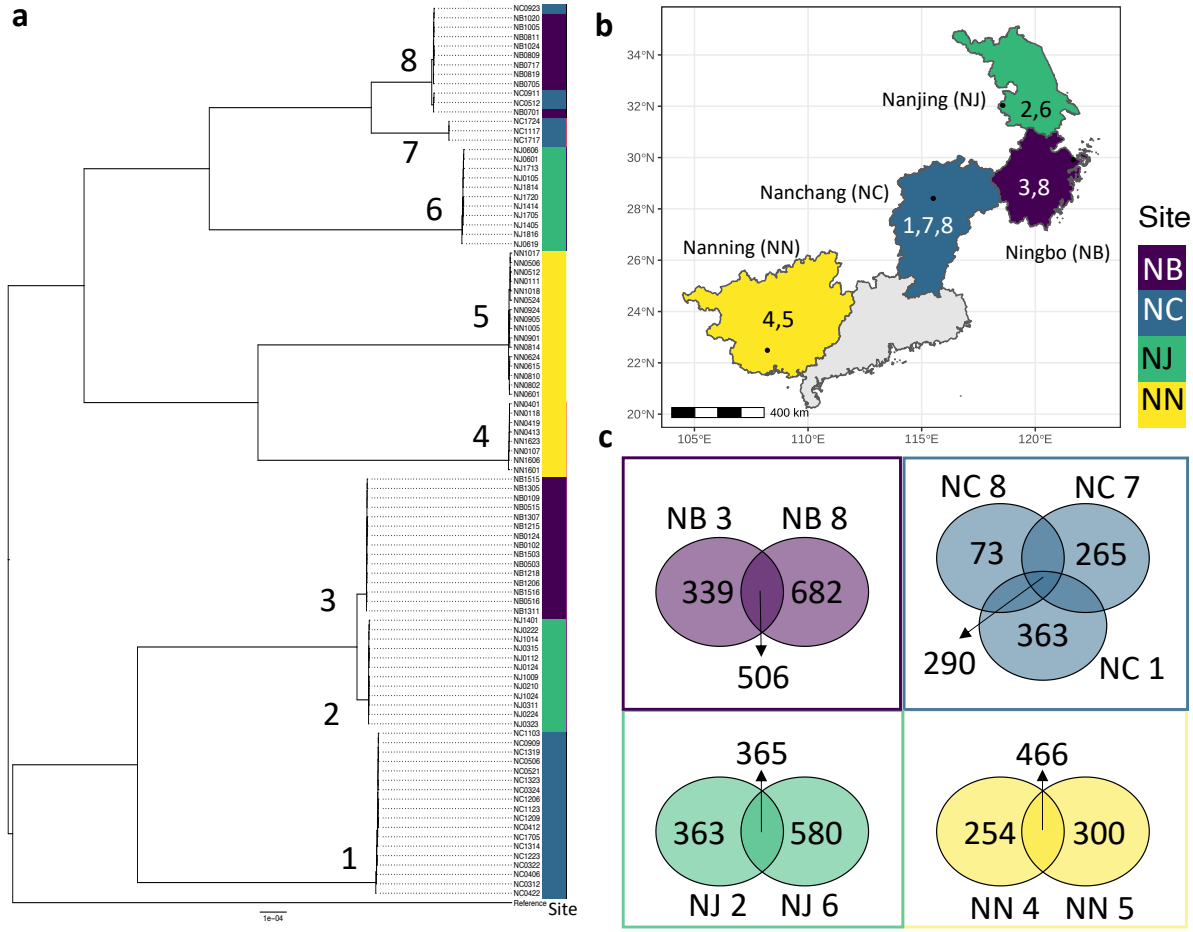

a
8
7
6
5
4
3
2
1
Site
b
Nanjing (NJ)
2,6
3,8
Nanchang (NC)
1,7,8
Nanning (NN)
Ningbo (NB)
4,5
400 km
34°N
32°N
30°N
28°N
26°N
24°N
22°N
20°N
105°E
110°E
115°E
120°E
Site
NB
NC
NJ
NN
c
NB 3
NB 8
339
682
506
NC 8
NC 7
73
265
363
290
NC 1
365
363
580
NJ 2
NJ 6
466
254
300
NN 4
NN 5

Figure 22. **Genetic divergence of *Ralstonia solanacearum* phylotype I sampled from 4 Chinese provinces.**

a) Maximum likelihood phylogeny constructed based on core SNPs compared to reference genome for R. solanacearum GMI1000. Clonal groupings are shown in red and metadata colour strips for the site of isolation (inner) and sequevar classification (outer) are added on the right of the isolate names. Recombinant sites were removed from the alignment of 3522 core genes using ClonalFrameML and phylogeny was constructed using IQTree (GTR+G4), midpoint rooted. (See Supplementary Figure 19 for bootstrap; b) The map is showing the geographic location of the four fields within the four Chinese provinces the bacterial samples originate from: Nanjing (NJ), Nanning (NN), Ningbo (NB) and Nanchang (NC)); c) Venn diagram of the 2346 accessory genes (orthologous groups of genes removed if present in <=2 strain and in >= 90% of strains). The name of each bacterial isolate is representative of the sampling. For example, NC0923 means an isolate from the NC field from plant number 9 and clone number 23.

Two clonal lineages co-occurred in each of the four fields apart from NC, where three lineages were observed (Figure 22 b). Each clonal lineage was specific to its location except for lineage eight, which was observed in both NC and NB. The co-occurring lineages originated from the two main clades with three out of four fields (Figure 22 a), and each of the clonal lineages contained accessory genes that were field-specific (Figure 22 c). The presence of the clonal group eight within both NC and NB fields suggests that this strain was potentially recently introduced to both fields potentially along with agricultural practices due to the large geographic distance between the two fields (Figure 22 b). We also compared the genetic variation in terms of nucleotide diversity (Pi) within the plant, field site, and clonal group levels (Supplementary Figure 14). It was estimated that each plant individual harbours low nucleotide diversity with an average of 1 clonal lineage per plant (median Pi per plant = 0.1). In contrast, nucleotide diversity observed per field was much higher due to two lineages co-occurring in each field (median Pi per field site = 4023). Together, these findings suggest that while different *R. solanacearum* genotypes co-occurred within each tomato field but presence of *R. solanacearum* at the plant level rhizosphere was limited to one genotype.

4.4.4. Coexisting lineages differ phenotypically regarding their virulence

The 95 genotyped strains were further investigated for phenotypic diversity by measuring traits indicative of *Ralstonia solanacearum* virulence: biofilm production, disease incidence, growth rate, competitive ability, swimming motility, exopolysaccharides (EPS) and siderophore production. To observe the differences between the coexisting genotypes, we compared the bacterial isolates' phenotypes split by field site and genotype (Figure 23). In the NB field, differences between the cooccurring genotypes 3 and 8 were significant ($p<0.001$) and the observed difference was especially affected by the biofilm formation trait (Figure 23 a-c). In the NJ field site, there was a significant difference between the coexisting clonal lineages 2 and 6 ($p<0.001$) (Figure 23 g-i). In addition, to compare the performance of the observed coexisting genotypes and provide an ecotype explanation for the coexistence we observed in each field site, we measured the growth of the 95 *Ralstonia solanacearum* isolates in 48 different substrate media broadly categorised as sugars, amino and organic acids, and other (Supplementary Figure 21). No significant per-field genotype differences were observed with the permutational multivariate analysis of variance. Therefore, the difference in the metabolic capacity of the strains was not linked to the coexisting genotypes per field (Supplementary Figure 21). We compared the metabolic profiles of the coexisting genotypes by creating a binary matrix of growth vs no growth from the growth rate data (Supplementary Figure 21 b).

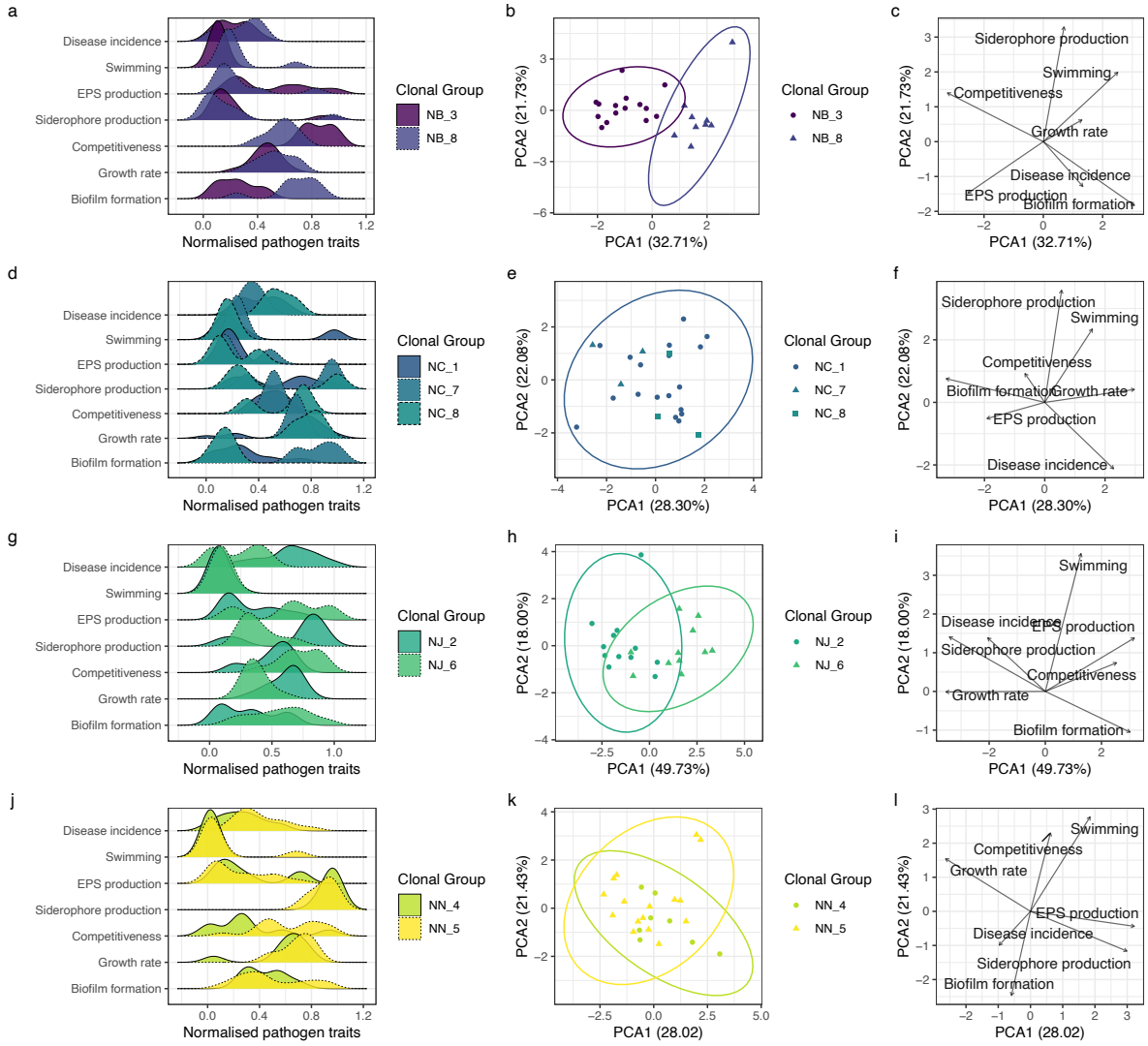


Figure 23. **Phenotypic variation in 95 *Ralstonia solanacearum* isolates from China**

PCA analysis of the phenotypic space of the 96 sequenced isolates. Panels a,d,g, j show frequency distribution of the 7 phenotypes measured: Disease incidence, Swimming motility, EPS production, Siderophore production, Competitiveness, Growth rate and Biofilm formation. Panels b,e, h,k show the PCA of the 7 phenotypes and panels c, f, i, l show the loadings for these phenotypes.

4.4.5. Cooccurring lineages have different virulence gene profiles

Substantial variation in the accessory genome was observed from pangenome analysis of the 95 isolates, and many genes were accessory between the genotypes found within each field (Figure 22 c & Supplementary Figure 16). Thus, we wanted to deploy a comparative genomics

approach to investigate the presence of virulence-associated genes and gene elements within the 8 *Ralstonia solanacearum* genotypes found here. We hypothesised that different virulence gene repertoire provides variable infection abilities to different *Ralstonia solanacearum* phylotype I genotypes. This provides an explanation for the coexistence of the different ecotypes within the same location. We investigate the variation in the core virulence genes' non-synonymous mutation profiles and the following genetic elements related to virulence: Type III Effectors (T3Es) from the Ralstonia T3E database & presence of prophages in the genomes. For this comparative analysis, we removed the clonal group 8 strains from the NC site as their genomes were smaller and missing many genes.

The mutation profile of the core virulence genes was investigated. The presence of missense mutations within core virulence pathways between the coexisting genotypes indicates there will be differences in the virulence mechanisms employed by them within the same field. Here we observed multiple missense mutations present within the clonal groups' virulence pathway genes and differing profiles of missense mutation depending on the clonal groups' phylogenetics. The differences are mainly associated with the three-clade ancestral split in the phylogeny before the further differentiation of the observed eight clonal groups. The three groups can be seen as a three-cluster split in a distance plot based on the presence-absence of mutated virulence genes (Figure 24 b). Interestingly when the mutated virulence pathways are compared between the genotypes cooccurring in each field, we can see that three out of the four fields studied have major differences in the mutation profile of cooccurring genotypes (Figure 24 a). These differences indicate a link between virulence gene profile and co-occurrence of *Ralstonia solanacearum* genotypes within an agricultural field.

Furthermore, the Ralstonia T3E database was searched to identify differences in major virulence genes that can contribute to different modes of infection deployed by the eight genotypes. In total, 72 accessory T3E genes were found in our sample set of 95 bacterial isolates along with GMI1000 (Figure 24 C). The whole genome distance data suggests that clonal groups 1-3 and clonal groups 4-8 have a distant common ancestor, and this two-way split is reflected in the presence of 3 effector genes: RipJ, RipAX1, and Hyp6 (Supplementary Figure 17). Multicopy of RipJ were observed in the upper half of the phylogeny and have previously been recorded in strains from phylotype I from China and South Korea, indicating

it could be associated with southeast Asian strains from phylotype I of *Ralstonia pseudosolanacearum* (Pandey et al., 2021). In addition, most of the strains belonging to phylotype I from around the world in the Ralstonia T3E database have a pseudogene for RipAX1; however, the same gene is always present in clonal groups 6-8 and perhaps indicating a very important and under-investigated role of the gene within this lineage. Hyp6 or Hypothetical 6 is a more recently discovered effector and not very well documented, but it appears to be very variable both between the eight clonal groups compared here and compared to GMI1000. Moreover, we detected six hypothetical T3Es (2, 6, 7, 8, 14, 17), all of which are absent from reference genome GMI1000 but have been detected in other strains from phylotype I from around the world and in other phylotypes (Sabbagh et al., 2019). Hypothetical 17 is the only clonal type specific (unique for that clonal type) T3E gene detected here for clonal group 7. Overall, the agreement of phylogeny and T3E effector clustering suggests there is a distinct profile of T3Es associated with each genotype studied, and multiple T3E profiles exist per field (Supplementary Figure 17).

In addition, prophage's presence within the genome sets was investigated. There is a known link between insertion or prophage sequences gene disturbance and virulence in *Ralstonia solanacearum* strains (Gonçalves et al., 2020b). The clonal group-specific pattern was also observed in the prophage analysis, with prophage profiles differing between cooccurring isolates, suggesting that field cooccurrence of genotypes may be due to differential virulence and gene content (Figure 24 D). Notably, we found the prophage RSS1, shown previously to enhance virulence in tomatoes (Addy et al., 2012, p. 1). RSS1 was in clonal group 2 but not in clonal group 6 suggesting the cooccurrence of these two genotypes in the same field may be due to virulence differences.

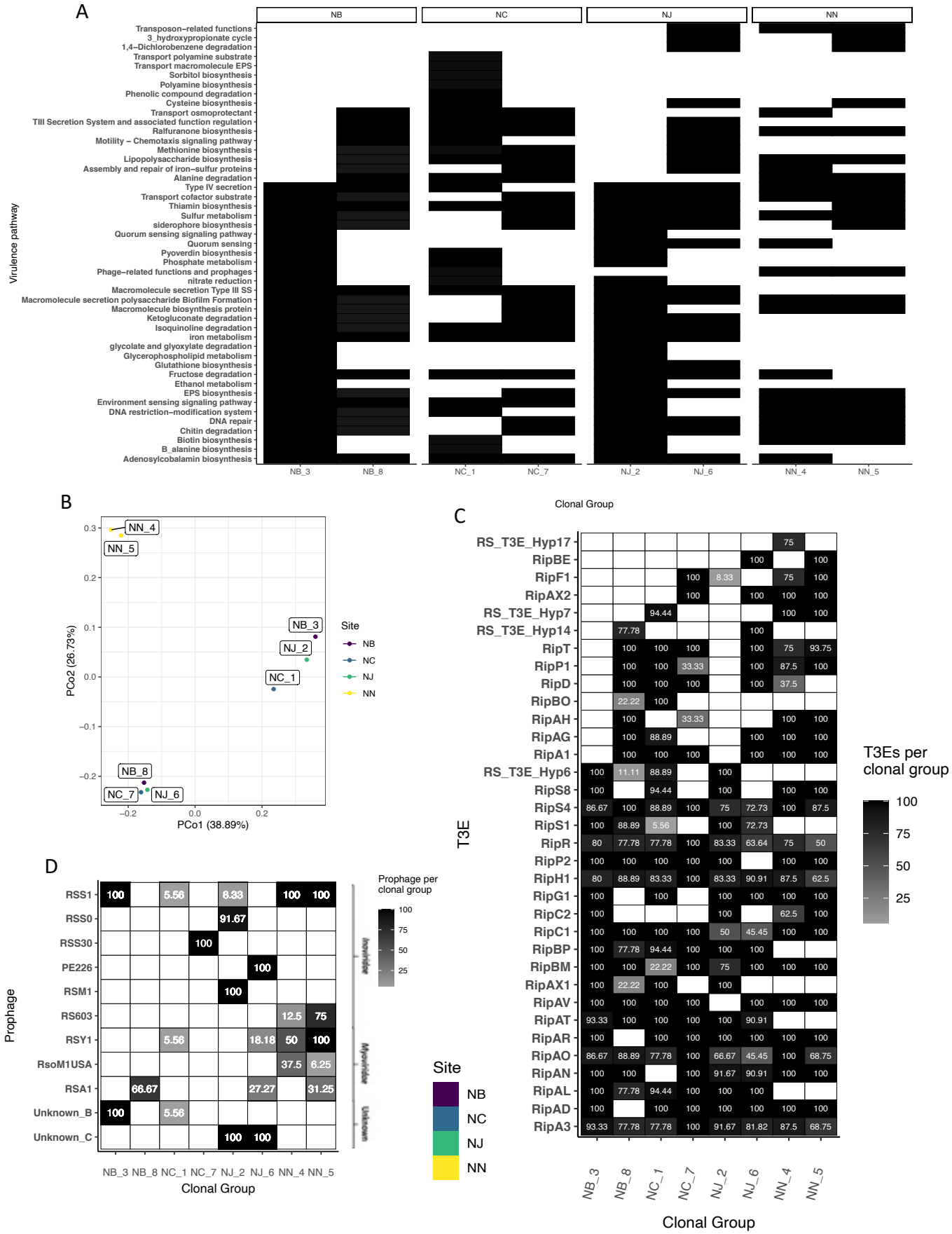

A
NB
NC
NJ
NN
Virulence pathway
Transposon−related functions
3_hydroxypropionate cycle
1,4−Dichlorobenzene degradation
Transport polyamine substrate
Transport macromolecule EPS
Sorbitol biosynthesis
Polyamine biosynthesis
Phenolic compound degradation
Cysteine biosynthesis
Transport osmoprotectant
TIII Secretion System and associated function regulation
Ralfuranone biosynthesis
Motility − Chemotaxis signaling pathway
Methionine biosynthesis
Lipopolysaccharide biosynthesis
Assembly and repair of iron−sulfur proteins
Alanine degradation
Type IV secretion
Transport cofactor substrate
Thiamin biosynthesis
Sulfur metabolism
siderophore biosynthesis
Quorum sensing signaling pathway
Quorum sensing
Pyoverdin biosynthesis
Phosphate metabolism
Phage−related functions and prophages
nitrate reduction
Macromolecule secretion Type III SS
Macromolecule secretion polysaccharide Biofilm Formation
Macromolecule biosynthesis protein
Ketogluconate degradation
Isoquinoline degradation
iron metabolism
glycolate and glyoxylate degradation
Glycerophospholipid metabolism
Glutathione biosynthesis
Fructose degradation
Ethanol metabolism
EPS biosynthesis
Environment sensing signaling pathway
DNA restriction−modification system
DNA repair
Chitin degradation
Biotin biosynthesis
B_alanine biosynthesis
Adenosylcobalamin biosynthesis
NB_3
NB_8
NC_1
NC_7
NJ_2
NJ_6
NN_4
NN_5
B
PCo2 (26.73%)
PCo1 (38.89%)
Site
NB
NC
NJ
NN
C
Clonal Group
T3E
RS_T3E_Hyp17
RipBE
RipF1
RipAX2
RS_T3E_Hyp7
RS_T3E_Hyp14
RipT
RipP1
RipD
RipBO
RipAH
RipAG
RipA1
RS_T3E_Hyp6
RipS8
RipS4
RipS1
RipR
RipP2
RipH1
RipG1
RipC2
RipC1
RipBP
RipBM
RipAX1
RipAV
RipAT
RipAR
RipAO
RipAN
RipAL
RipAD
RipA3
T3Es per clonal group
D
Prophage
RSS1
RSS0
RSS30
PE226
RSM1
RS603
RSY1
RsoM1USA
RSA1
Unknown_B
Unknown_C
Prophage per clonal group
Inoviridae
Myoviridae
Unknown
Site
NB
NC
NJ
NN
Clonal Group

Figure 24. **Virulence gene profiles**

a) Non-synonymous mutations in virulence pathways Core non-synonymous mutations presence in genes is shown in this plot and mutations are summarised over pathway and clonal group. If a mutation is present in >80% of isolates in a clonal group in each pathway it is shown in black and it is shown in white if that pathway is mutated in <80% of the isolates in a clonal group. Therefore, virulence pathways that are mutated in most of the isolates in a clonal group are represented with black colour.

b) Non-synonymous mutations in virulence genes PCoA shows that mutation profiles align with expected clustering based on phylogenetic relationship between strains.

c) Presence of Type III effector proteins Presence absence matrix for the T3Es was constructed as multicopy or presence was accepted as 1 and pseudogene or absence as 0. The data shown is proportion of isolates per clonal group that have a certain T3E gene on the y-axis.

d) Presence of prophages the y-axis shows clonal groupings (1-8) colour coded by origin of samples within the Chinese province, and the x axis shows the prophage cluster and phage family (top) and the prophage species (bottom) – for prophage cluster determination (

Supplementary Figure 18). Cells are shaded based on the proportion of isolates that contain

4.4.6. GWAS identified a transposase and a type III effector linked to siderophore production and maximum growth phenotypes in *Ralstonia solanacearum*

Microbial genome-wide association study (GWAS) approach was used to link the genetic and virulence-linked phenotypic diversity across the whole dataset of 95 isolates (Supplementary Table 8). We used three different genetic components for the GWAS: 1) shared k-mers (unitigs) of sequence; 2) shared genes from the pangenome and 3) genetic variants from variant calling analysis against GMI1000. The unitig GWAS identified a few genes associated with growth and swimming motility. In case of growth, association with the gene RSp0842, annotated as a putative Type III effector protein, was found. Instead, swimming motility was associated with three genes: Rsc3183 (probable hemagglutinin related protein) and two phage-related proteins (Rsc1924, Rsc1923). The SNP GWAS identified associations also for

swimming motility and competition. For the competition, a missense variant in the RSp1282 gene, which encoded for phosphoenolpyruvate - a protein phosphotransferase involved in binding magnesium. For swimming motility, a synonymous variant in the RSc0102 gene, which was putatively annotated as involved in calcium binding was found. The gene presence-absence GWAS identified the highest number of significant associations with disease incidence (virulence in planta), competition, swimming motility, exopolysaccharide production (EPS) and siderophore production. However, annotations were only present for disease incidence and siderophore production. For disease incidence, the gene group_4779 was found, which encoded for a RidA family protein. For siderophore production, the gene group_5250, which encoded an IS5 family transposase IS1420 was identified. Together, these results highlight potential role of these genes for *R. solanacearum* virulence evolution in tomato fields in China.

4.5. DISCUSSION

Our results show that *R. solanacearum* isolates from four geographically disconnected tomato fields in China form three genetically distinct clades, which are further divided into eight clonal lineages. Interestingly, in three out of four fields, isolates from two clonal lineages coexist, representing diversity from two distinct clades. In one of the four fields, we observe its own clade that split into two clonal groups. Coexisting lineages had unique virulence gene profiles and show phenotypic divergence regarding their virulence traits in two out of four fields. Using GWAS analysis, we further linked phenotypic variation with a transposable element and a type III effector protein, which are both known to contribute to virulence in *R. solanacearum*. Together these results show that *R. solanacearum* infections in China are not clonal but vary considerably within agricultural tomato monocultures. Specifically, the coexisting lineages within fields differed in their virulence, potentially resulting in two ecological strategies with low niche overlap.

We found that the abiotic and biotic environment of the tomato plants' rhizosphere was strongly associated with the field site the tomato plants originate from (Figure 20 b&c). Structure within biotic and abiotic measurements based on location was expected because the sampled fields were far apart and due to general heterogeneity between sampling

environments. Thus, we also expected to see the field site structure in the phenotypic and genotypic data of the sampled *Ralstonia solanacearum* isolates. However, we saw that the total phenotypic variation (based on 1152 clones) was explained by a two-way split in the phenotypic space with isolates belonging to both clusters present in all fields. Furthermore, the genotypic data showed even more heterogeneity in the dataset with 8 different genotypes present in a subset of 95 sequenced clones. We show that two genotypes in each field site cooccur and the genotypes differ in their core genome and accessory gene content within each field site. Many of the gene differences between the 8 genotypes can be accounted for due to an earlier phylogenetic split splitting the 95 genomes into two major clades: clade I (genotype 1-3) and clade II (4-8). However, this earlier genetic split does not align with the binomial split observed in the phenotypic data. Previously, in a tomato field study in China, it was shown that three pathotypes of *R. solanacearum* coexisted (Zheng et al., 2014).

This study shows that recombination rates were higher than mutation rates based on the phylogenetic recombination inference performed. Therefore, horizontal gene transfer is vital in the diversification process of *R. solanacearum* strains within China and has largely contributed to the variation observed. Typical levels of homologous recombination were observed here, with recombination contributing on average three times more diversity than mutation (Price and Arkin, 2015). However, the presence of recombination means that the clonal clades should be able to merge or get closer together as recombination leads to a purge of genetic variation and brings lineages closer together (Sheppard et al., 2018). On the contrary, we observe clearly separate genotype lineages with large genetic distances between them but virtually no variation within the clonal groups. This phenomenon fits well with the stable ecotype model hypothesis which can explain the existence of clonal groups in the environment (Bobay and Raymann, 2019). In this model, an ecological species or strain adapted to a specific niche can expand when favourable conditions arise and cause a selective sweep. In opportunistic bacterial species such as *R. solanacearum*, cells can survive in an environmental host reservoir or in the soil and avoid competition with a new expanding strain (Elphinstone et al., 1998b; Schönfeld et al., 2003). Then when conditions change the population surviving in a refuge can expand again leading to coexistence due to differential temporal dynamics. Thus, differential pressures over time can explain the observation of

differing dynamics between strains with different virulence strategies. Unfortunately, here we investigate a sample of bacterial genomes in a single point in time which makes it hard to conclude what the long-term evolutionary dynamics of *R. solanacearum* in the Chinese agricultural environment. However, due to the long population history *R. solanacearum* has had in China it can be assumed that there are multiple genotypes that can be found in the natural environment and agriculture (Hanson et al., 2012; Sun et al., 2017; Wicker et al., 2012b). Potentially these genotypes persist due to variable niche preferences and can survive within their own niche which may expand and contract as the environment varies over time.

We show that the eight *Ralstonia solanacearum* clonal lineages studied here have multiple gene differences based on their pangenome. Crucially, these differences included SNP variation and presence-absence of important virulence genes such as the Type III effector proteins and virulence-associated prophages. It is not clear to what extent the pangenome is expendable even within important systems such as the type III effectors where a lot of homology and redundancy can be observed in *Ralstonia solanacearum* strains (Sabbagh et al., 2019). Overall, the significance of the variability in pangenome is an ongoing discussion and it is not yet known to what extent accessory genes contribute to an individual's fitness (McInerney et al., 2017). However, there is mounting evidence that accessory genes can be associated with microbial growth and survival in different environments (Kent et al., 2016). We know that genes like type III effectors are hugely important for plant pathogens' ability to infect multiple hosts and we see a huge variation in the multi-host bacteria *Ralstonia solanacearum* so disregarding this variation as simply neutral would not be sensible. Moreover, we know that mobile genetic elements can contribute to virulence and previous studies have identified a link between prophage presence and virulence in *Ralstonia solanacearum* virulence in tomato (Addy et al., 2012; Yamada, 2013). The differing accessory genes between the two major lineages of the phylogeny could be signs of adaptation of the cooccurring clonal lineages to the local environment.

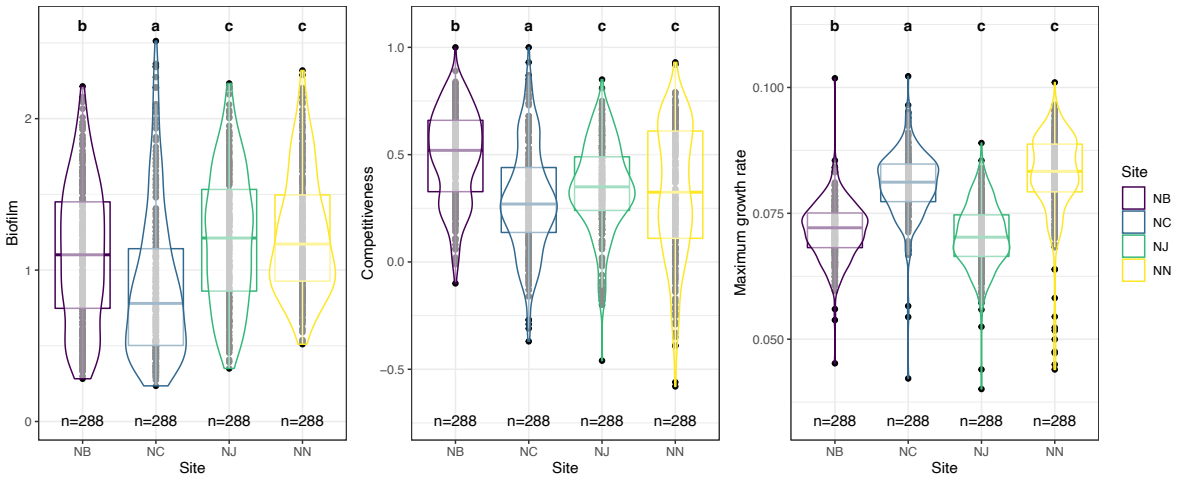
Our GWAS analysis also identified a few links between genetic variants and important virulence phenotypes. Swimming motility is essential for the early stages of the infection cycle in *R. solanacearum* strains but flagella are also a common receptor for the phage adsorption (Abedon, 2009; Corral et al., 2020). Here, we identified two phage-related proteins linked

with swimming motility. Furthermore, we linked disease incidence (virulence *in planta*) to a RidA family protein. It has been shown that this family of proteins are highly conserved across bacteria, archaea and eukaryotes and can provide protection against metabolic damage by reactive intermediates (Irons et al., 2020). Reactive oxygen species (ROS) are a big part of the defence response in plants against pathogen invasions (Waszczak et al., 2018). Therefore, a protein that aids bacteria to defend themselves against ROS could be beneficial and increase their infection success and fitness. In addition, we identified an association between IS5 family transposase (IS1420) and siderophore production. Transposable elements are known to contribute to *R. solanacearum* genome plasticity and 'phenotypic conversion' from virulent to avirulent strains (Jeong and Timmis, 2000). In a previous pangenome study, 20 IS families were found to be widespread across the *R. solanacearum* strains, and among them, IS5 and IS3 were the most abundant. Therefore, our finding on the association between IS5 family transposase IS1420 with siderophore production aligns with previous knowledge of *R. solanacearum* virulence being affected by the movement of transposable elements, and to our knowledge, IS1420 has not been linked to *R. solanacearum* virulence traits previously. Finally, we also identified a link between *R. solanacearum* maximum growth rate and the of a putative Type III effector protein. Trade-offs between growth and virulence have been experimentally quantified for *R. solanacearum* phylotype I strain previously (Peynard et al., 2016). Here, we found this further supports our idea of trade-offs and compensatory mutations leading to the evolution of different phenotypes of *Ralstonia solanacearum* cooccurring with each field.

Strikingly, we found that multiple *R. solanacearum* clonal lineages co-occurred in three out of four fields even though the infections remained clonal at the plant level. One crucial genetic and the phenotypic difference between coexisting lineages was found in virulence traits and genes at both core and accessory genome levels. This suggests that selection by the local environment was not strong enough to lead to selective sweeps and dominance of the relatively more virulent clonal lineage despite the relatively homogeneous tomato monocultures. This can potentially be explained by competition between the bacterial genotypes for resources at various stages of the infection, which might result in lower niche overlap or differential plant infection success at different phases of plant growth and development. Alternatively, it is possible that spatial heterogeneity and variation

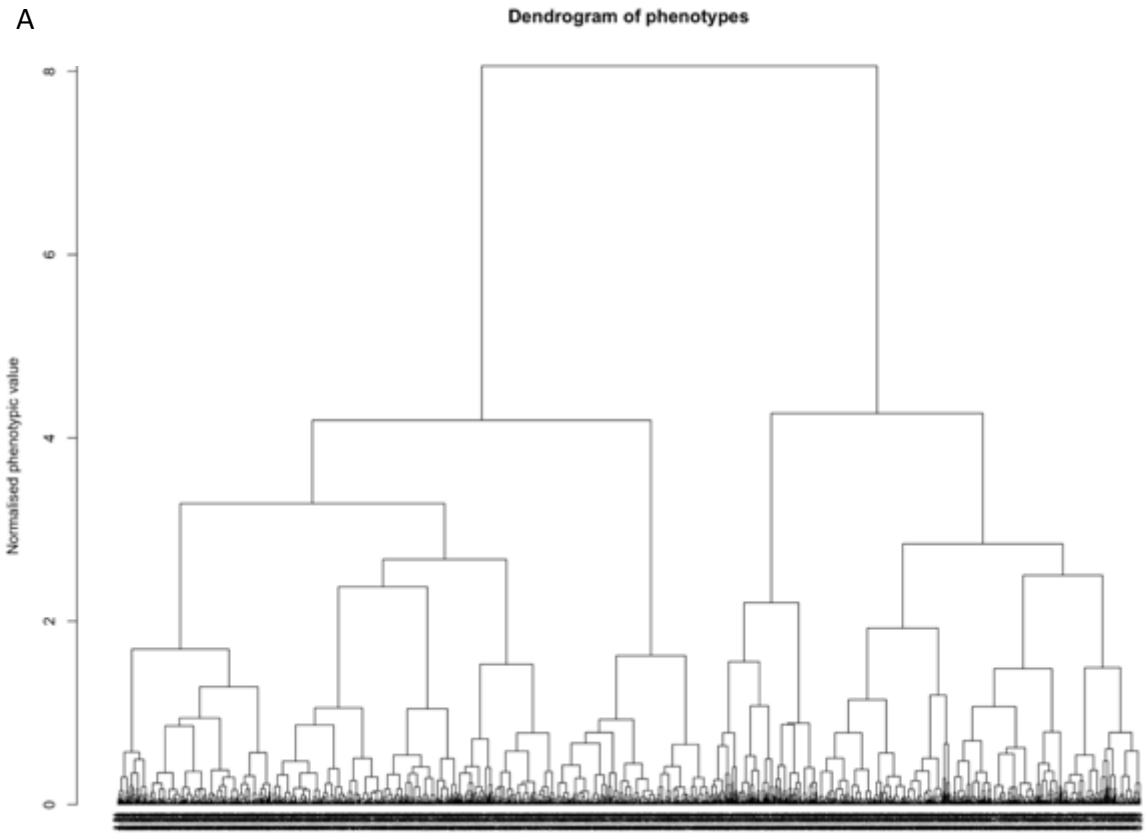
microenvironmental conditions favoured different coexisting lineages at different locations of the field. However, more data covering a wider spatial sample distribution is required to confirm this hypothesis in the future. Also, competitive selection could be imposed by the plant presenting a barrier which only a single genotype can overcome at a time. Alternatively, the infection may not be the bottleneck but rather competition between the strains at the early stages of infections which was not observed due to the time point in the bacterial lifecycle sampled here. The weight of these components' contribution to the success of infection should be assessed with further field and lab experiments focusing on these questions.

4.6. SUPPLEMENTARY

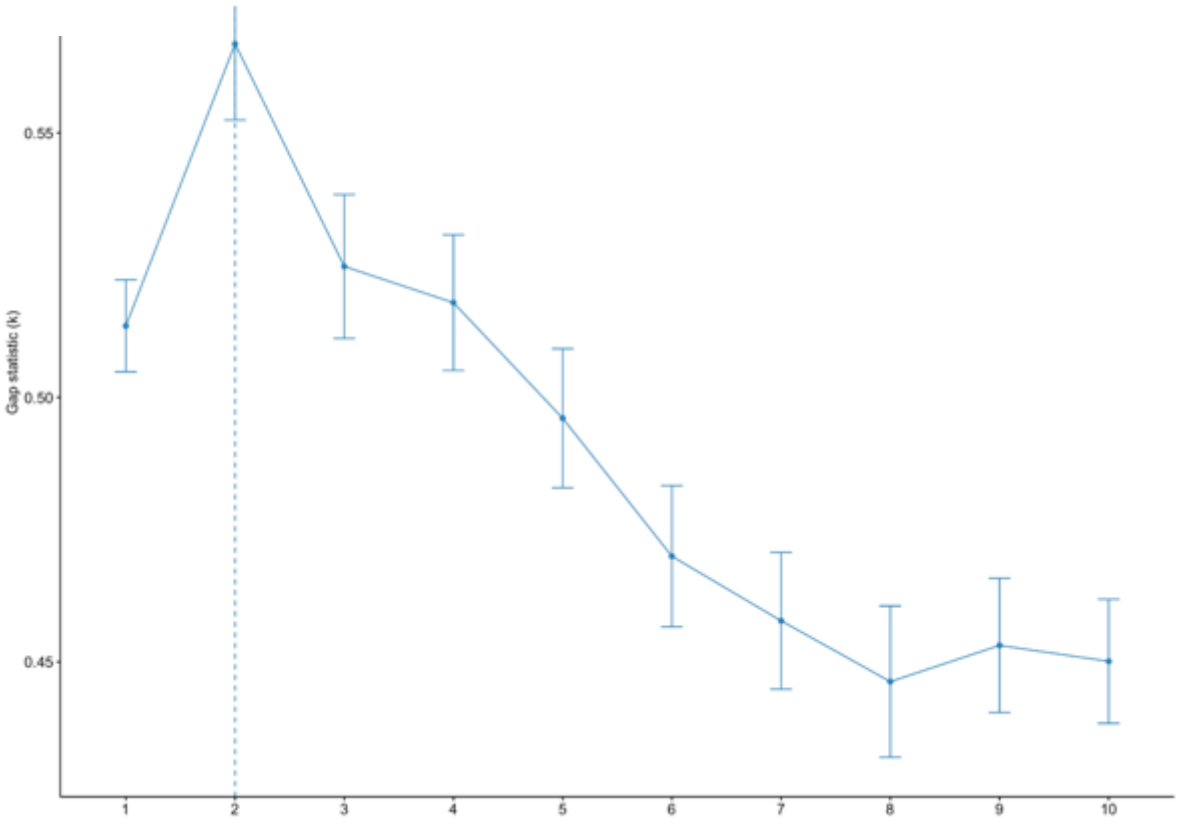


Supplementary Figure 12. **Phenotypic comparisons of 1152 bacterial isolates per field site** Biofilm, competitiveness, and maximum growth rate were compared using a linear model and sidak correction. Significant groups are shown with different letters.

A


Dendrogram of phenotypes
Normalised phenotypic value
0
2
4
6
8

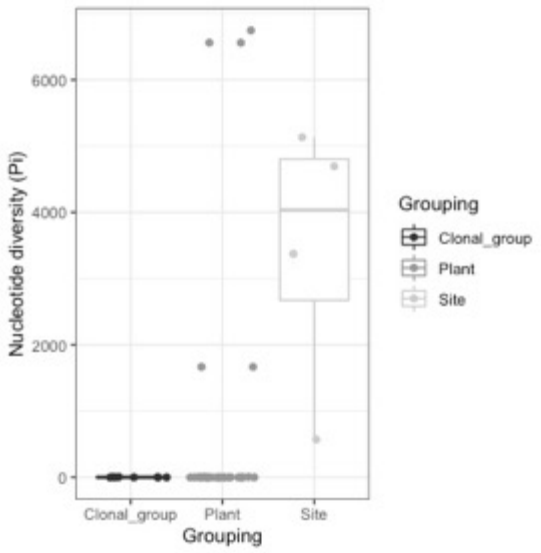
B


Gap statistic (k)
0.45
0.50
0.55
1
2
3
4
5
6
7
8
9
10

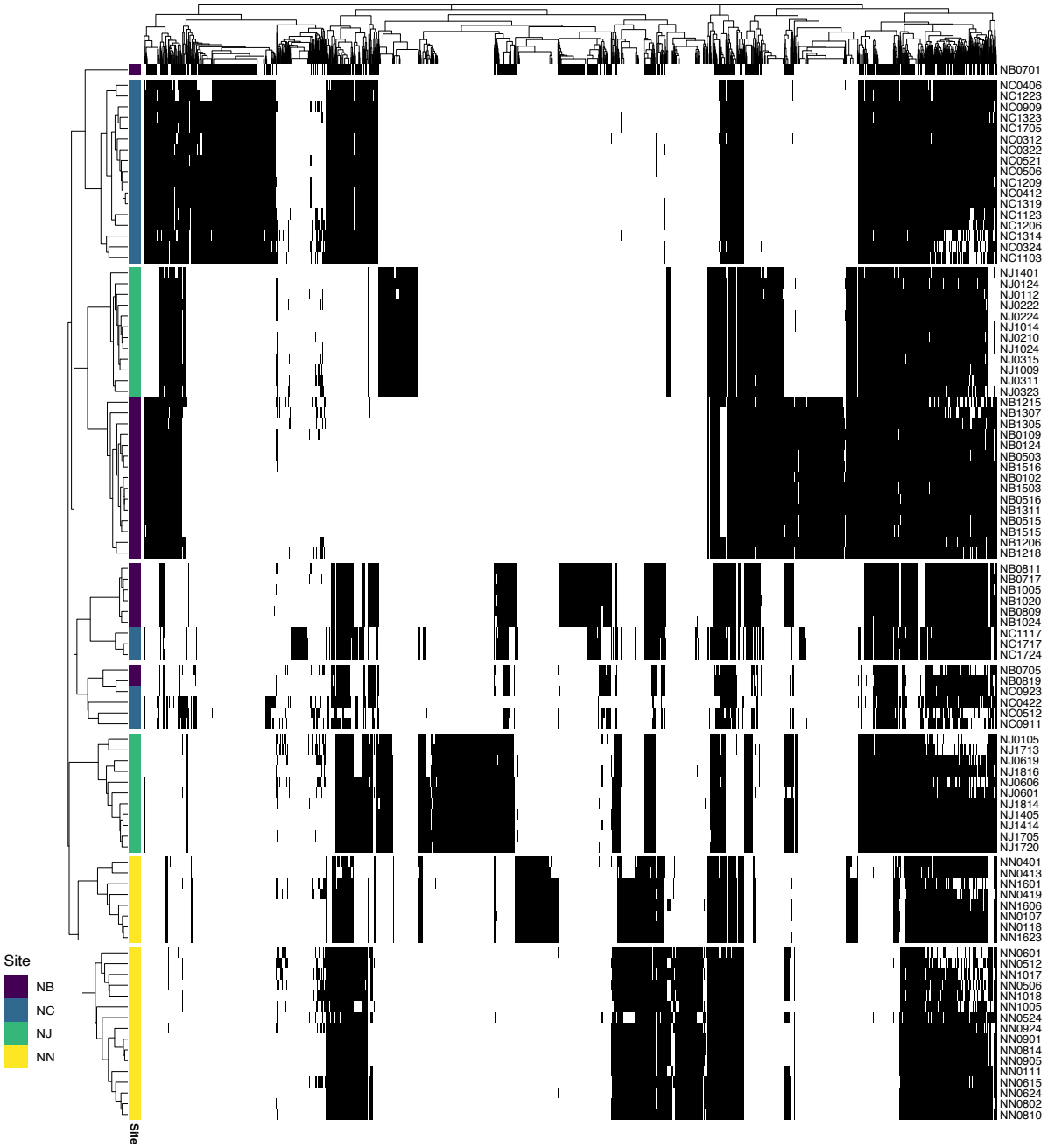
Supplementary Figure 13. **Phenotypic comparisons of 1152 bacterial isolates hierarchical clustering and gap statistic** Biofilm, competitiveness, and maximum growth rate measurements were hierarchically clustered (A) and optimal number of clusters was found using a gap statistic (B).

Site	Cluster 1	Cluster 2
NB	152	136
NC	222	66
NJ	130	158
NN	162	126

Supplementary Table 7. **Two-way split based on k-means clustering of the 1152 isolates based on 3 phenotypic traits.** The table shows how many isolates from each field site fall within each phenotypic cluster.

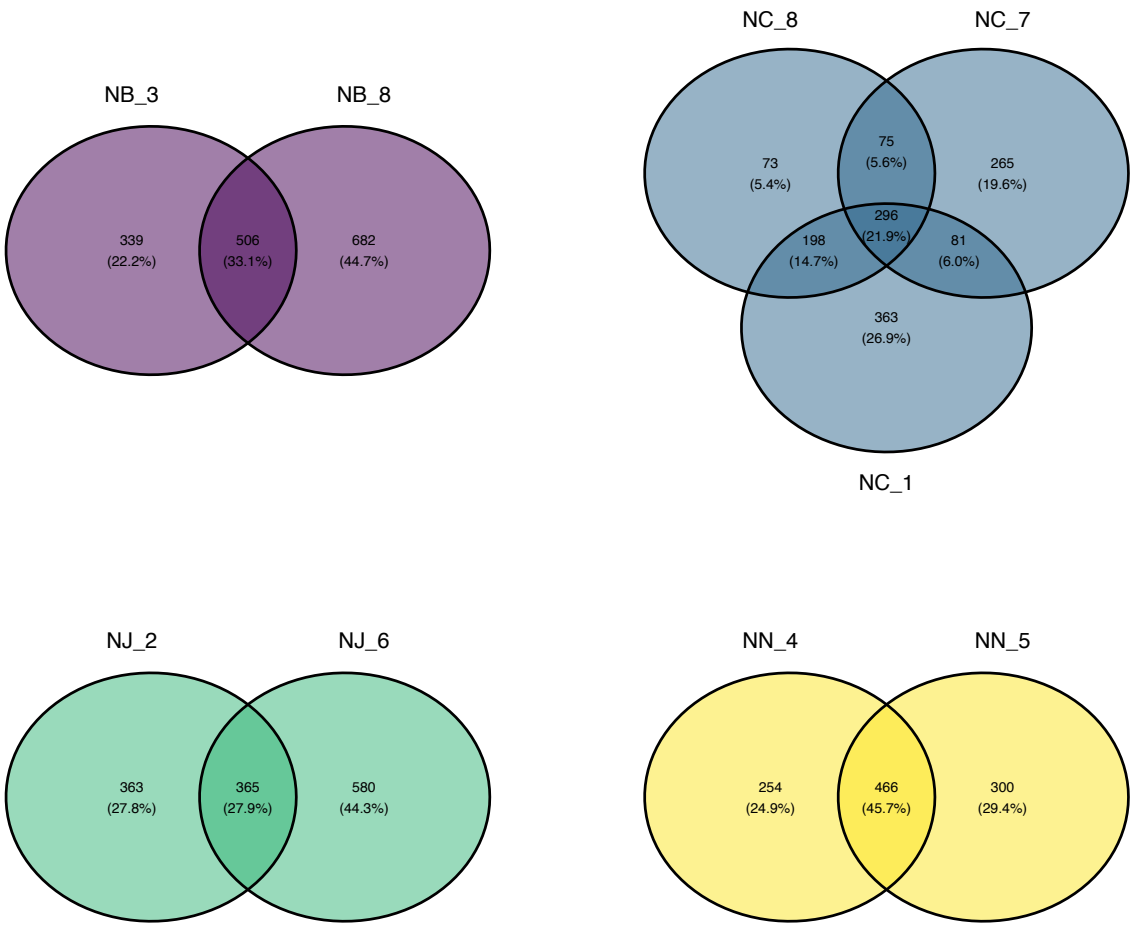


Supplementary Figure 14. **Nucleotide diversity Pi**



Supplementary Figure 15. **Accessory genes in the 95 Chinese isolates**

Gene presence/absence matrix **showing** hierarchically clustered genes. The singletons and core (95% of isolates) genes are excluded from the graph before clustering in order to show the accessory genome variation. White is absent and black is presence of a gene cluster within a genome. Colour strip on the left shows field site for the origin of the samples. Hclust dunction from base R used to cluster the matrix.



Supplementary Figure 16. **Venn diagram of accessory genes.**

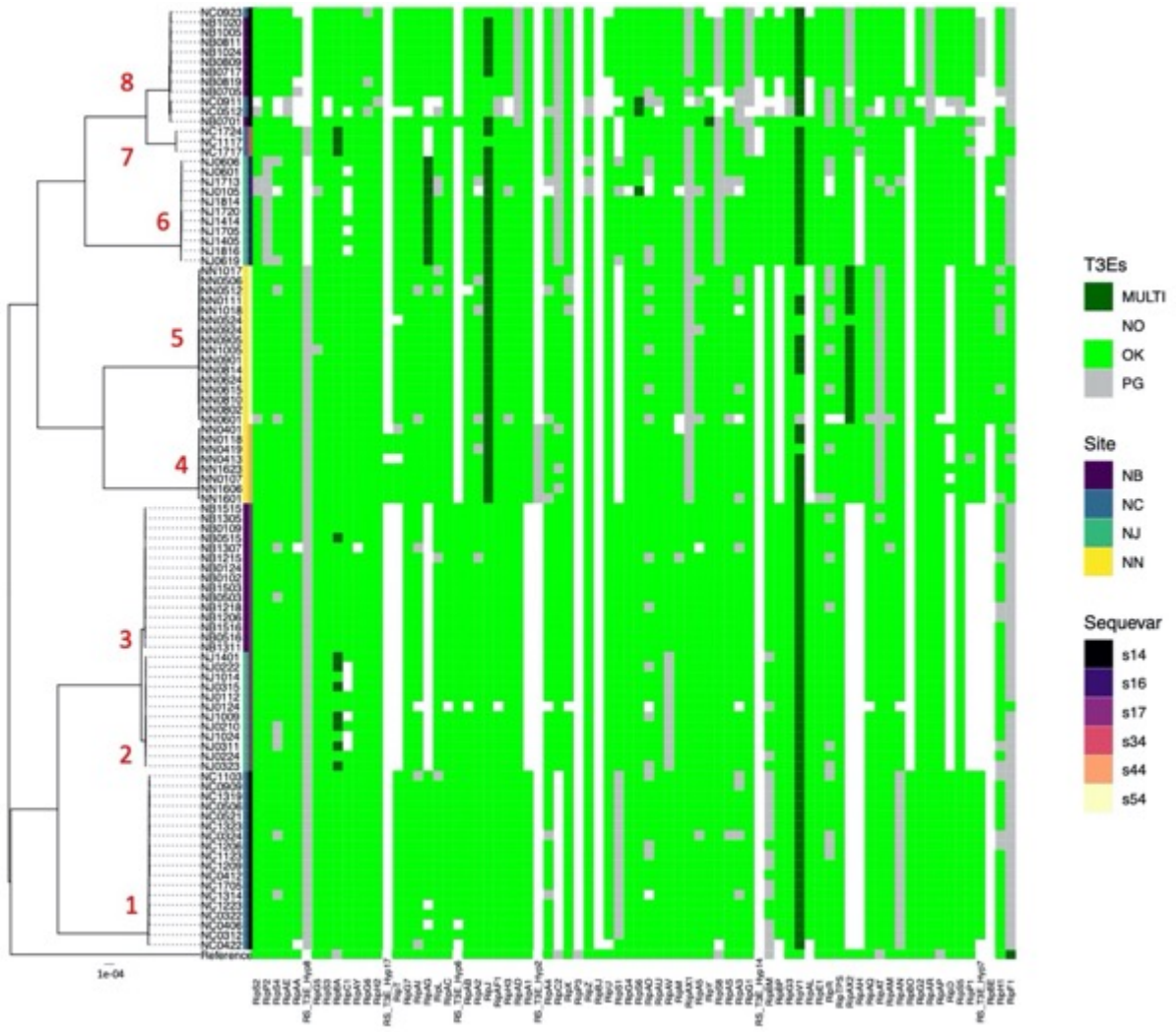
Accessory genes shared between clonal lineages divided by field site are shown.

	Phenotype	**Significant Gene**	**York name**	**Function**	**Uniprot link**	**Heritability (h2)**
Unitig	Growth rate	RSp0842		T3E associated	https://www.uniprot.org/uniprot/Q8XRI9	0.40
	Swimming	Rsc3183		Probable hemagglutinin related	https://www.uniprot.org/uniprot/Q8XUK5	0.25

		Rsc1924, Rsc1923		Phage related	https://www.uniprot.org/uniprot/Q8XY39 https://www.uniprot.org/uniprot/Q8XY40	
Gene						
	Disease incidence		Group_4652, Group_4779	Hypothetical, RidA family protein		0.87
	Competition		Group_4712, group_2873	Hypothetical, Hypothetical		0.71
	Swimming		Group_5353	Hypothetical		0.25
	EPS		Group_5225	Hypothetical		0.37
	Siderophore		Group_5250	IS5 family transposase IS1420		0.66
SNP	Competition	RSp1282	AL646053_1629852_G_A	Missense variant in RSp1282 - Phosphoenolpyruvate-protein phosphotrans-ferase – binding magnesium	https://www.uniprot.org/uniprot/Q8XQE5	0.71

	Swimming	RSc0102	AL646052_116265_G_A	Synonymous variant in RSc0102 – putative calcium binding	https://www.uniprot.org/uniprot/Q8Y378	0.25

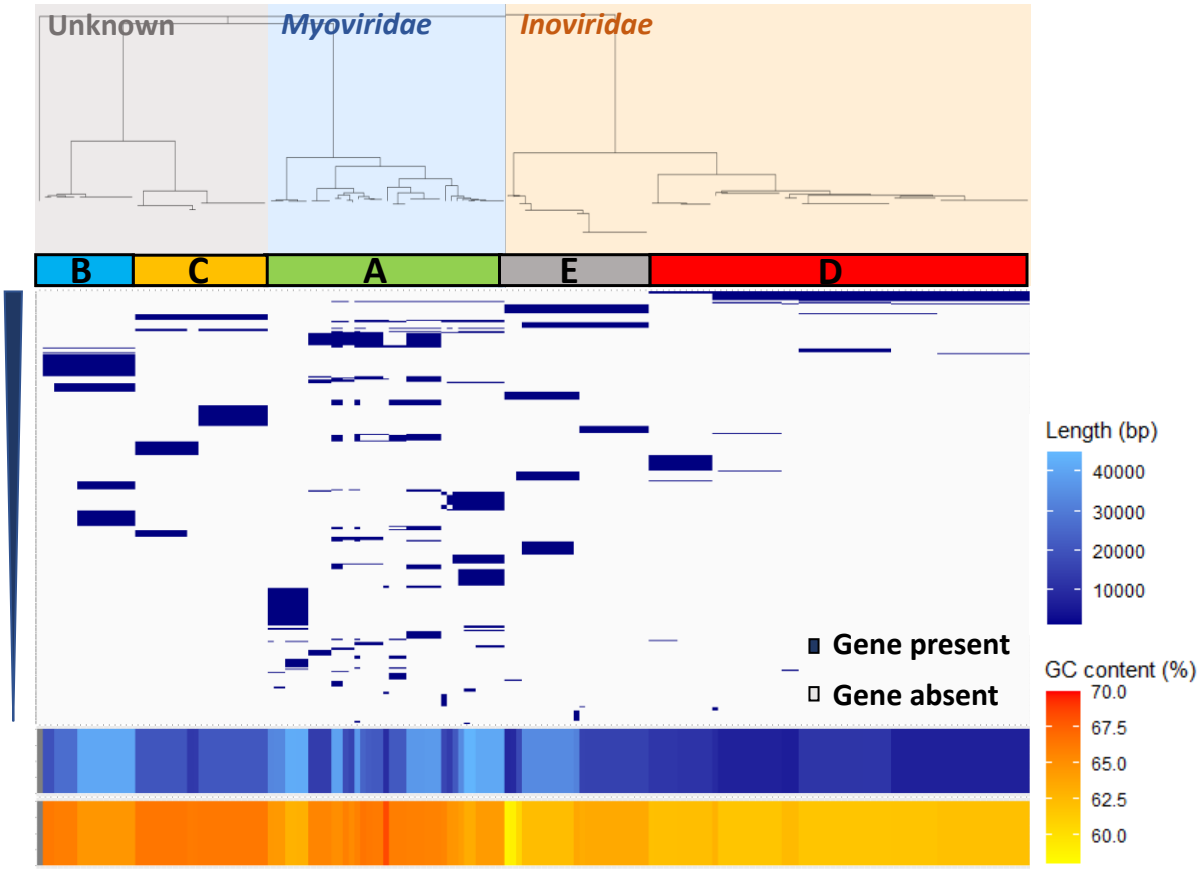
Supplementary Table 8. **GWAS results**. Only significant hits of the GWAS results are shown after Bonferroni correction.



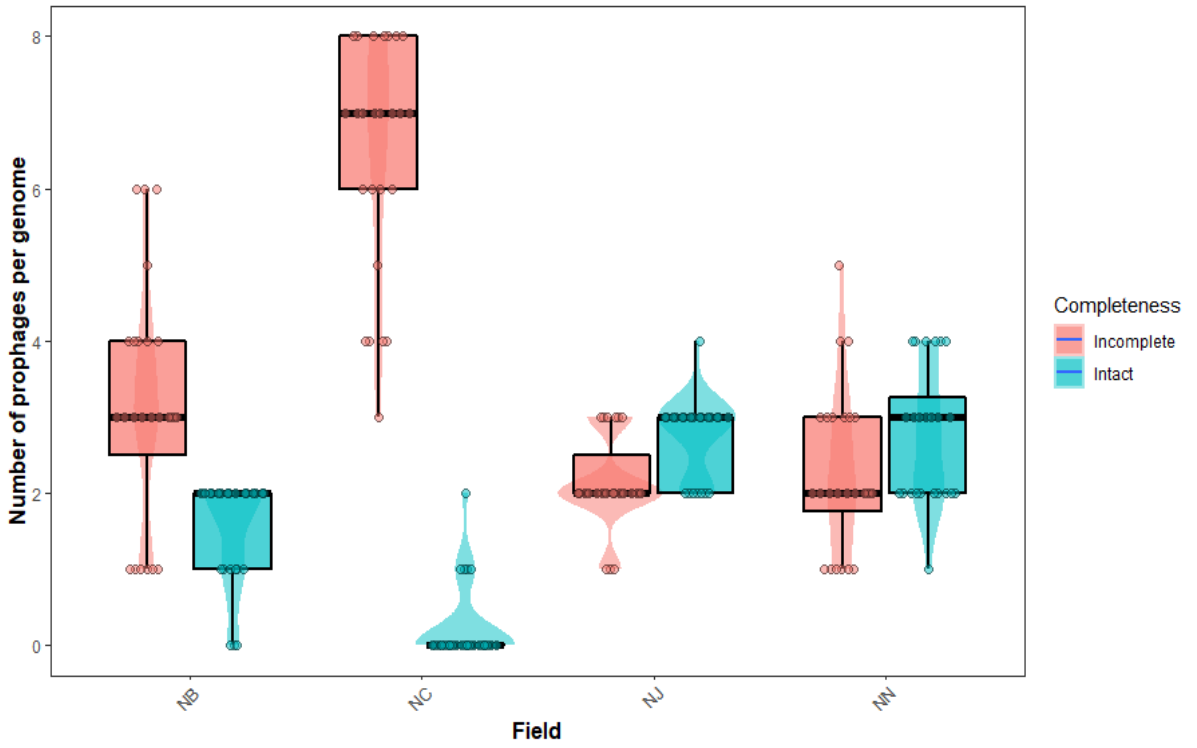
Supplementary Figure 17. **Categorical table showing presence-absence of T3Es compared to GMI1000.** The core SNPs compared to reference strain GMI1000 phylogeny is represented on the left and coloured bars for the field site and sequevar of the *Ralstonia solanacearum* isolate. The presence-absence pattern of Type III effector protein genes is aligned to each

isolate of the phylogeny. The x-axis shows 76 T3Es and their presence is indicated in 4 categories as defined by Peeters *et al.* (2013): Multiple copies of the gene (MULTI); absence of the gene (NO); the presence of the gene (OK) or pseudo gene (PG). Maximum likelihood core gene phylogeny presented on the left labelled with location of isolation and lineages.

A

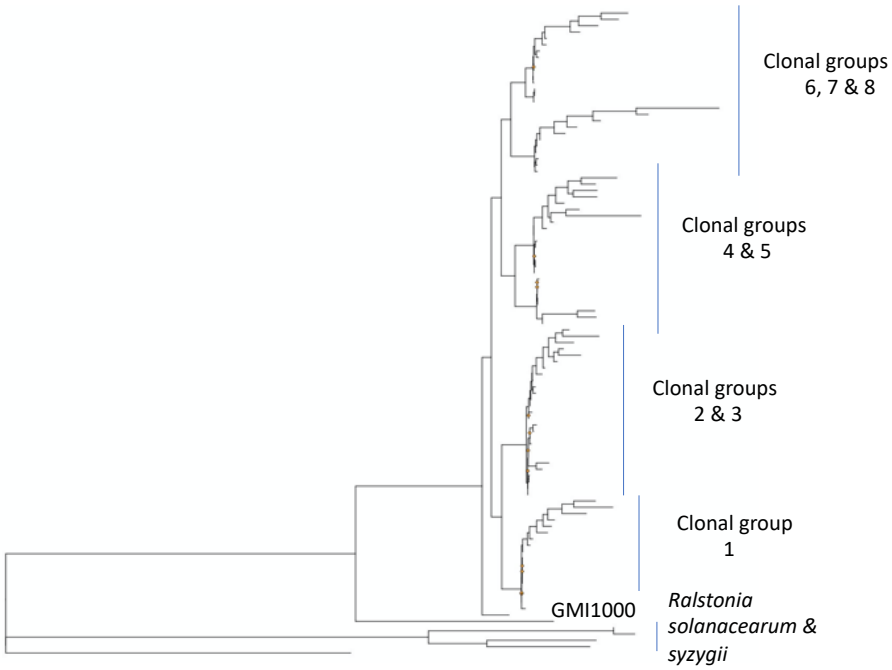

Unknown
Myoviridae
Inoviridae
B
C
A
E
D
Gene present
Gene absent
Length (bp)
40000
30000
20000
10000
GC content (%)
70.0
67.5
65.0
62.5
60.0

B


Number of prophages per genome
0
2
4
6
8
NB
NC
NJ
NN
Field
Completeness
Incomplete
Intact

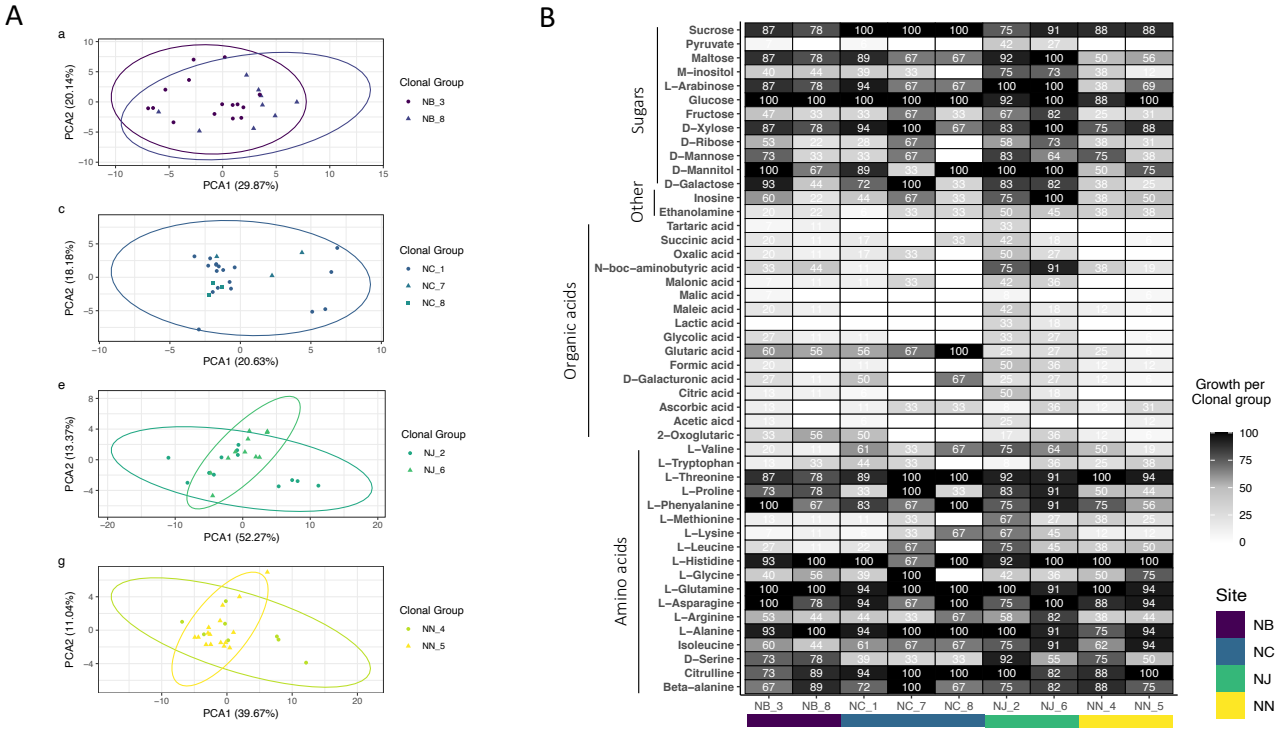
Supplementary Figure 18. **Prophage diversity in Chinese isolates**

A) Top – Neighbour-joining tree of intact prophages constructed using Mash distances, annotated with prophage clusters, and coloured based on phage family. **Middle** – heatmap showing the presence (blue) and absence (grey) of genes within prophages, with decreasing gene abundance further down the plot. **Bottom** – length and GC content distribution of prophages. **B)** Prophage prevalence per field from four Chinese provinces. Boxplot underlaid with a violin plot of the prevalence of incomplete and intact prophages per genome (y axis) in the four Chinese provinces (x axis). Completeness is labelled on the right.



Supplementary Figure 19. **Phylogeny of field samples from 4 Chinese provinces**

Maximum likelihood tree showing branches with LOW bootstrap support consensus from two bootstrap methods (alrt <70 & boot <95) with yellow diamonds. Only very few peripheral branches show low support.



Supplementary Figure 20. **Metabolic capacity of the 96 isolates**

a) PCA analysis of growth in carbons, amino and organic acids substrates plotted. Area under the curve of 24-hour growth curves shown.

b) Growth rate of the isolates in different substrates shown as a binary matrix heatmap. Percentage shows number of isolates per clonal group capable of growing in each substrate. The shade of white to grey also shows this percentage.

5. General Discussion

In a changing climate, food security is increasingly threatened by plant pathogens. The transfer of agricultural pathogens with human practices and trade has been an ongoing issue in the global economy. Still, the change in the environment provides increasingly favourable conditions for bacterial pathogens creating novel opportunities. Genomic evolution research can help us understand the selection drivers for virulence and pathogen survival within agricultural settings and environmental reservoirs. This can help us design control strategies with evolution in mind. The overall aim of this PhD project was to investigate the genetic heterogeneity of the diverse plant pathogenic bacteria *Ralstonia solanacearum* species complex at three geographical levels: global, one country and a crop field. The research presented used comparative genomics techniques to investigate the variation in mutations and gene content in the pangenome of the species complex at the three levels. This chapter provides an overview and synthesis of the results of this thesis. It discusses them in the context of the three central aims and their significance and contribution to the broader knowledge of pathogen evolution.

A central theoretical concept in this thesis has been that pangenomes vary in bacterial species (Tettelin et al., 2005). It was of research interest that bacterial pathogens with the ability to occupying large number of niches have multiple strains that differ in their accessory gene content (Freschi et al., 2019; Rasko et al., 2008; Xin et al., 2018). Such species often have large genome sizes, multiple strains, and an extremely large number of possible niches that they can occupy. *Ralstonia solanacearum* species complex is a diverse plant pathogen that can survive inside and outside host plants (Hayward, 1964). Therefore, we believed that investigating *Ralstonia solanacearum* species complex isolates from different countries, and from agricultural and environmental settings can reveal interesting patterns about the maintenance of a large pangenome in this species.

In the first results chapter, we attempted to show the significance of the variability in the pangenome *Ralstonia solanacearum* species complex (RSSC). We showed that important virulence genes, transposases and abiotic defence genes can cooccur in the accessory genome. We believe this result highlights and adds additional weight to an important recent finding that transposons contribute to the genome plasticity of RSSC (Gonçalves et al., 2020b).

Knowledge that insertion sequences are abundant in bacterial genomes is not novel but the contribution of insertion sequence generated genomic variation and its contribution to evolution and adaptability in certain niches is not clear (Siguier et al., 2006). A recent paper showed that IS-mediated mutations are abundant in the *E. coli* genome evolution. They showed that IS elements accounted for ~35% of the mutations observed over 50,000 generations of *E. coli*. They concluded that variation generated by IS movement can both promote and constrain evolvability as there were both detrimental and beneficial effects observed in the population (Consuegra et al., 2021). We had no means of estimating the size of the effect transposases have on the accessory gene movement in the RSSC pangenome. However, we believe the presence of genes cooccurring with transposases in 10% of the cooccurring components within the accessory genome shows a link between insertion sequences and accessory gene movement across RSSC strains. In other words, part of the gene level variability in *Ralstonia solanacearum* is due to mobile genetic movement.

The link between insertion sequences and accessory gene variation in RSSC was further shown in the second results chapter by the investigation of the genome changes within the UK river water population of *Ralstonia solanacearum* phylotype IIB-1 over 27 years. The phylotype IIB-1 strain is highly clonal around the world but can evolve in response to stress in the lab where mechanistically adaptation was linked to insertion sequence movement (Alderley et al., 2022; Clarke et al., 2015). We showed that the mutation rate in the UK *Ralstonia solanacearum* is as low as 1 nucleotide change per year, but we observed multiple gene differences between the bacterial isolates. The gene differences were mainly transposons or genes in proximity to transposons. Therefore, we believe the genome variation generated by insertion sequences is generating larger size sequence variation compared to genome duplication errors.

In addition, in the third results chapter we showed that genetic heterogeneity in RSSC is present not only worldwide but within individual crop fields. In theory, tomato plantations are homogenous environments, but we believe the ability of RSSC strains to survive in the soil and in the plants provides enough environmental variability and available niches for multiple genotypes to occur in a relatively small area. Our data does not allow for direct estimation of competition for resources within the plant or the soil. However, we show that the cooccurring *R. solanacearum* within three of the four fields studied have multiple gene differences.

Furthermore, the variation in siderophore production phenotype was linked to presence of a transposase in a genome-wide association study. Siderophore production is linked to the master regulator virulence network of RSSC with the PhcA gene on top (Bhatt and Denny, 2004).

Overall, this thesis has shown that accessory gene variation is present in *R. solanacearum* in the field, within the environmental reservoir of the freshwater ecosystem and within the wide global pangenome. The common result in the three levels of variation investigated has been that transposons constitute a part of the accessory genome and are often found in proximity to important virulence and survival genes. Therefore, this thesis provides evidence that genetic diversity in a non-model plant pathogen is present in the pangenome population. We suggest that insertion sequence movement can provide a means of detecting genetic variation at an earlier stage of diversification compared to the traditional approach of core genome variation detection through single nucleotide polymorphism analysis. Therefore, future studies of adaptation to environment and fields in plant pathogens should consider insertion sequence movement as well as mutation to achieve a more holistic picture of the adaptability of a pathogen. Moreover, the evidence of large level of variation of the accessory genome provided here shows that *Ralstonia solanacearum* species complex future research should focus on investigating the role of the accessory genes and their potential for spread across different Ralstonia and other soil microbe species. Overall, the role and function of most of the accessory genes found here was unknown or hypothetical. Thus, the role of these genes especially in virulence and spread of this and other soil borne pathogens should be thoroughly investigated. *Ralstonia solanacearum* research is growing due to interest in the species both its specific pathology and generally as a model organism for plant bacterial infection. Therefore, studying the accessory genome of *Ralstonia solanacearum* can provide general insight about accessory genome function in plant pathogens.

6. APPENDIX 1

York Number	Species York	Phylotype York	NCPPB	Host	Year	Country	County	UK_region
YO001	Ralstonia pseudosolanacearum	III	NCPPB1018	Potato	1961	Angola		
YO002	Ralstonia solanacearum	II	NCPPB1483	Potato	1963	Australia		
YO003	Ralstonia pseudosolanacearum	I	NCPPB2245	Stylosanthes humulis	1969	Australia		
YO004	Ralstonia solanacearum	II	NCPPB3980	Potato	1997	Australia		
YO005	Ralstonia pseudosolanacearum	I	NCPPB3992	Tobacco		Australia		
YO006	Ralstonia pseudosolanacearum	I	NCPPB4001	Ginger		Australia		
YO007	Ralstonia pseudosolanacearum	I		Potato		Australia		
YO008	Ralstonia solanacearum	II		Potato	2012	Bangladesh		
YO009	Ralstonia pseudosolanacearum	I		Potato	2012	Bangladesh		
YO010	Ralstonia solanacearum	II		Potato	2012	Bangladesh		
YO011	Ralstonia solanacearum	II		Potato	2014	Bangladesh		
YO012	Ralstonia solanacearum	II		Potato	1993	Belgium		

YO013	Ralstonia solanacearum	II	NCPPB0613	Potato	1958	Brazil		
YO014	Ralstonia solanacearum	II		Tobacco	1968	Brazil		
YO015	Ralstonia solanacearum	II	NCPPB3649	Banana	1979	Brazil		
YO016	Ralstonia solanacearum	II	NCPPB3650	Banana	1979	Brazil		
YO017	Ralstonia solanacearum	II	NCPPB3868	Potato	1991	Brazil		
YO018	Ralstonia pseudosolanacearum	I	NCPPB3864	Chili	1991	Brazil		
YO019	Ralstonia solanacearum	II	NCPPB3862	Chili	1991	Brazil		
YO020	Ralstonia pseudosolanacearum	I	NCPPB3863	Tomato	1991	Brazil		
YO021	Ralstonia pseudosolanacearum	I	NCPPB3866	Potato	1993	Brazil		
YO022	Ralstonia solanacearum	II	NCPPB3982	Potato		Chile		
YO023	Ralstonia pseudosolanacearum	I	NCPPB4006	Olive		China		
YO024	Ralstonia pseudosolanacearum	I	NCPPB4007	Mulberry		China		
YO025	Ralstonia pseudosolanacearum	I	NCPPB	Mulberry		China		

			4012					
YO026	Ralstonia solanacearum	II	NCPPB4011, NCPPB3850	Mulberry		China		
YO027	Ralstonia pseudosolanacearum	I	NCPPB3994	Olive		China		
YO028	Ralstonia pseudosolanacearum	I	NCPPB3998	Ginger		China		
YO029	Ralstonia pseudosolanacearum	I	NCPPB4003	Ginger		China		
YO030	Ralstonia pseudosolanacearum	I	NCPPB4008	Peanut		China		
YO031	Ralstonia solanacearum	II	NCPPB282	Potato	1950	Colombia		
YO032	Ralstonia solanacearum	II	NCPPB3594	Heliconia caribaea	1960	Colombia		
YO033	Ralstonia solanacearum	II		Tobacco	1966	Colombia		
YO034	Ralstonia pseudosolanacearum	I	NCPPB2154	Heliconia sp.	1958	Costa Rica		
YO035	Ralstonia solanacearum	II	NCPPB787	Banana	1959	Costa Rica		
YO036	Ralstonia solanacearum	II	NCPPB0788	Banana	1959	Costa Rica		
YO037	Ralstonia pseudosolanacearum	I	NCPPB	Nightshade	1959	Costa Rica		

			0790					
YO038	Ralstonia pseudosolanacearum	I	NCPPB0791	False daisy	1959	Costa Rica		
YO039	Ralstonia solanacearum	II		Potato	1972	Costa Rica		
YO040	Ralstonia pseudosolanacearum	I	NCPPB4004	Ginger		Costa Rica		
YO041	Ralstonia solanacearum	II	NCPPB3977	M. perfoliatum		Costa Rica		
YO042	Ralstonia pseudosolanacearum	I	NCPPB3993	Pepper		Costa Rica		
YO043	Ralstonia solanacearum	II		Potato		Costa Rica		
YO044	Ralstonia solanacearum	II	NCPPB643	Potato	1959	Cyprus		
YO045	Ralstonia solanacearum	II	NCPPB1584	Potato	1963	Cyprus		
YO046	Ralstonia solanacearum	II	NCPPB909	Potato	1961	Egypt		
YO047	Ralstonia solanacearum	II	NCPPB1115	Potato	1961	Egypt		
YO048	Ralstonia solanacearum	II	NCPPB1824	Potato	1966	Egypt		
YO049				Potato	1991	Egypt		
YO050	Ralstonia solanacearum	II		Potato	1991	Egypt		

YO051	Ralstonia solanacearum	II		Potato	1991	Egypt		
YO052	Ralstonia solanacearum	II		Potato	1991	Egypt		
YO053	Ralstonia solanacearum	II		Potato	1994	Egypt		
YO054	Ralstonia solanacearum	II		Potato	1995	Egypt		
YO055				Potato	1995	Egypt		
YO056				Potato	1995	Egypt		
YO057				Potato	1995	Egypt		
YO058				Potato	1995	Egypt		
YO059				Potato	1995	Egypt		
YO060	Ralstonia solanacearum	II		Potato	1995	Egypt		
YO061				Potato	1995	Egypt		
YO062	Ralstonia solanacearum	II		Potato	1996	Egypt		
YO063	Ralstonia solanacearum	II		Potato	1997	Egypt		
YO064	Ralstonia solanacearum	II		Potato	1998	Egypt		
YO065	Ralstonia solanacearum	II		Potato	1998	Egypt		
YO066	Ralstonia solanacearum	II	NCPPB4153	Potato	1998	Egypt		
YO067				Potato	1998	Egypt		
YO068	Ralstonia solanacearum	II		Water	2002	Egypt		

YO069	Ralstonia solanacearum	II		Water	2002	Egypt		
YO070	Ralstonia solanacearum	II		Potato	2002	Egypt		
YO071	Ralstonia solanacearum	II		Soil	2002	Egypt		
YO072	Ralstonia solanacearum	II			2002	Egypt		
YO073	Ralstonia solanacearum	II			2002	Egypt		
YO074	Ralstonia pseudosolanacearum	I	NCPPB1500	Potato	1963	Fiji		
YO075	Ralstonia pseudosolanacearum	I	NCPPB1702	Potato	1964	Fiji		
YO076	Ralstonia pseudosolanacearum	III	NCPPB1029	Pelargonium capitatum	1961	Reunion		
YO077	Ralstonia solanacearum	II	NCPPB2200	Tomato	1966	Guadeloupe		
YO078			NCPPB4157	Potato	1995	France		
YO079	Ralstonia pseudosolanacearum	I		Tomato	2011	France		
YO080	Ralstonia pseudosolanacearum	I	NCPPB2204	Tomato	1968	Guyana		
YO081	Ralstonia pseudosolanacearum	I		Tomato		Guyana	French Guyana	
YO082	Ralstonia pseudosolanacearum	III	NCPPB3181	Nightshade	1978	Gambia		

YO083	Ralstonia pseudosolanacearum	I		Tomato	2011	Georgia		
YO084	Ralstonia pseudosolanacearum	I		Tomato	2011	Georgia		
YO085	Ralstonia solanacearum	II	NCPPB4161	Potato	1996	Germany		
YO086	Ralstonia solanacearum	II	NCPPB1789	Potato	1965	Greece		
YO087	Ralstonia solanacearum	II	NCPPB2015	Potato	1967	Greece		
YO088	Ralstonia solanacearum	II	NCPPB3205	Banana	1979	Guyana		
YO089	Ralstonia solanacearum	II	NCPPB0789	Banana	1959	Honduras		
YO090	Ralstonia solanacearum	II		Potato	2012	Hungary		
YO091	Ralstonia solanacearum	II		Potato	2012	Hungary		
YO092	Ralstonia solanacearum	II		Potato	2012	Hungary		
YO093	Ralstonia solanacearum	II		Potato	2012	Hungary		
YO094	Ralstonia solanacearum	II		Potato	2012	Hungary		
YO095	Ralstonia solanacearum	II		Potato	2012	Hungary		
YO096	Ralstonia solanacearum	II	NCPPB1331	Potato	1962	India		

YO097	Ralstonia solanacearum	II	NCPPB1333	Potato	1962	India		
YO098	Ralstonia solanacearum	II	NCPPB3214	Banana	1980	India		
YO099	Ralstonia syzygii	IV	NCPPB3219	Clove	1980	Indonesia		
YO100			NCPPB3792	Clove	1985	Indonesia		
YO101	Ralstonia pseudosolanacearum	I	NCPPB3793	Potato	1985	Indonesia		
YO102	Ralstonia syzygii	IV	NCPPB3725	Banana	1987	Indonesia		
YO103				Banana	1987	Indonesia		
YO104	Ralstonia syzygii	IV	NCPPB3726	Banana	1987	Indonesia		
YO105				Banana	1987	Indonesia		
YO106	Ralstonia syzygii	IV	NCPPB3727	Banana	1987	Indonesia		
YO107			NCPPB3728	Banana	1987	Indonesia		
YO108	Ralstonia solanacearum	II		Clove	1987	Indonesia		
YO109	Ralstonia solanacearum	II		Clove	1987	Indonesia		
YO110	Ralstonia solanacearum	II		Syzygium agneum	1987	Indonesia		

YO111	Ralstonia solanacearum	II	NCPPB3794	Clove	1987	Indonesia		
YO112	Ralstonia solanacearum	II		Water	2007	Ireland		
YO113	Ralstonia solanacearum	II		Water	2007	Ireland		
YO114	Ralstonia solanacearum	II		Potato	2007	Ireland		
YO115	Ralstonia solanacearum	II			2007	Ireland		
YO116	Ralstonia solanacearum	II		Tomato	2007	Ireland		
YO117	Ralstonia solanacearum	II		Tomato	2007	Ireland		
YO118			NCPPB0928	Sugarcane	1956	Jamaica		
YO119	Ralstonia syzygii	IV	NCPPB3445	Clove	1983	Indonesia		
YO120	Ralstonia solanacearum	II	NCPPB173	Potato	1945	Kenya		
YO121	Ralstonia solanacearum	II	NCPPB1028	Potato	1961	Kenya		
YO122	Ralstonia pseudosolanacearum	I	NCPPB1045	Eggplant	1961	Kenya		
YO123	Ralstonia solanacearum	II	NCPPB1049	Tomato	1961	Kenya		
YO124	Ralstonia pseudosolanacearum	I	NCPPB4215	Water	2001	Kenya		

YO125	Ralstonia solanacearum	II	NCPPB4211	Pelargonium hortorum	2001	Kenya		
YO126	Ralstonia solanacearum	II	NCPPB4212	Pelargonium hortorum	2001	Kenya		
YO127	Ralstonia solanacearum	II	NCPPB4213	Water	2001	Kenya		
YO128	Ralstonia pseudosolanacearum	III	NCPPB4214	Soil	2001	Kenya		
YO129	Ralstonia solanacearum	II	NCPPB1489	Potato	1963	Madeira Islands		
YO130	Ralstonia pseudosolanacearum	I	NCPPB0792	Teak	1960	Malaysia		
YO131	Ralstonia pseudosolanacearum	I	NCPPB1052	Ginger	1961	Malaysia		
YO132	Ralstonia solanacearum	II	NCPPB1614	Potato	1964	Malaysia		
YO133	Ralstonia pseudosolanacearum	I	NCPPB3190	Tomato	1978	Malaysia		
YO134	Ralstonia solanacearum	II	NCPPB2199	Eggplant	1965	Martinique		
YO135	Ralstonia pseudosolanacearum	I	NCPPB253	Pine tree	1949	Mauritius		
YO136	Ralstonia pseudosolanacearum	I	NCPPB0500	Broad bean	1956	Mauritius		
YO137	Ralstonia pseudosolanacearum	I	NCPPB	Cabbage	1956	Mauritius		

			0501					
YO138	Ralstonia pseudosolanacearum	I	NCPPB0503	Dahlia sp.	1956	Mauritius		
YO139	Ralstonia pseudosolanacearum	I	NCPPB1621	Potato	1960	Mauritius		
YO140	Ralstonia pseudosolanacearum	I	NCPPB1484	Strelitzia reginae	1963	Mauritius		
YO141	Ralstonia pseudosolanacearum	I	NCPPB1485	Common bean	1963	Mauritius		
YO142	Ralstonia solanacearum	II	NCPPB3974	Tomato		Mexico		
YO143	Ralstonia solanacearum	II	NCPPB3238	Potato	1982	Netherlands		
YO144	Ralstonia solanacearum	II	NCPPB4156	Potato	1995	Netherlands		
YO145	Ralstonia pseudosolanacearum	III	NCPPB1703	Potato	1965	Nigeria		
YO146	Ralstonia solanacearum	II	NCPPB2088	Potato	1968	Nigeria		
YO147	Ralstonia solanacearum	II		Potato	2011	Pakistan		
YO148	Ralstonia solanacearum	II		Potato	2011	Pakistan		
YO149	Ralstonia pseudosolanacearum	I	NCPPB1123	Tomato	1961			

YO150	Ralstonia pseudosolanacearum	I	NCPPB1140	Tomato	1961			
YO151	Ralstonia pseudosolanacearum	I	NCPPB2937	Potato	1975			
YO152	Ralstonia solanacearum	II	NCPPB3985	Eggplant	1987	Peru		
YO153	Ralstonia solanacearum	II	NCPPB3990	Potato	1989	Peru		
YO154	Ralstonia pseudosolanacearum	I		Tomato		Peru		
YO155	Ralstonia pseudosolanacearum	I		Potato		Peru		
YO156	Ralstonia pseudosolanacearum	I	NCPPB3996	Tomato		Peru		
YO157	Ralstonia solanacearum	II	NCPPB2315	Banana		Peru		
YO158	Ralstonia solanacearum	II	NCPPB3986	Potato		Peru		
YO159	Ralstonia solanacearum	II	NCPPB3970	Banana	1992	Philippines		
YO160	Ralstonia solanacearum	II	NCPPB3971	Banana		Philippines		
YO161	Ralstonia pseudosolanacearum	I	NCPPB4005	Ginger		Philippines		
YO162	Ralstonia solanacearum	II			2014	Poland		

YO163	Ralstonia pseudosolanacearum	I		Rose	2016	Poland		
YO164	Ralstonia pseudosolanacearum	I		Rose	2016	Poland		
YO165	Ralstonia solanacearum	II	NCPPB1019	Tomato	1960	Portugal		
YO166	Ralstonia solanacearum	II		Potato	1995	Portugal		
YO167	Ralstonia solanacearum	II	NCPPB4158	Potato	1995	Portugal		
YO168	Ralstonia solanacearum	II	NCPPB1225	Tomato	1958	Puerto Rico		
YO169	Ralstonia solanacearum	II	NCPPB1226	Tomato	1958	Puerto Rico		
YO170	Ralstonia solanacearum	II		Potato	2012	Serbia		
YO171	Ralstonia solanacearum	II		Potato	2012	Serbia		
YO172	Ralstonia solanacearum	II		Potato	2012	Serbia		
YO173	Ralstonia solanacearum	II		Potato	2012	Serbia		
YO174	Ralstonia solanacearum	II			2013	Serbia		
YO175	Ralstonia solanacearum	II			2013	Serbia		
YO176	Ralstonia solanacearum	II			2013	Serbia		
YO177	Ralstonia solanacearum	II			2013	Serbia		

YO178	Ralstonia pseudosolanacearum	I	NCPPB1763	Tomato	1965	Seychelles		
YO179				Potato	1980	Slovenia		
YO180	Ralstonia solanacearum	II	NCPPB4160	Potato	1996	Spain		
YO181	Ralstonia solanacearum	II	NCPPB1323	Potato	1962	Sri Lanka		
YO182	Ralstonia pseudosolanacearum	I	NCPPB3217	Tumeric	1980	Sri Lanka		
YO183	Ralstonia solanacearum	II	NCPPB2505	Potato	1972	Sweden		
YO184	Ralstonia solanacearum	II	NCPPB2797	Solanum dulcamara	1974	Sweden		
YO185	Ralstonia solanacearum	II	NCPPB2796	Solanum dulcamara	1975	Sweden		
YO186	Ralstonia pickettii		NCPPB1695	Sugarcane	1965	Tanzania		
YO187	Ralstonia pseudosolanacearum	I	NCPPB4000	Ginger		Thailand		
YO188	Ralstonia solanacearum	II		Tomato	1957	Trinidad		
YO189	Ralstonia solanacearum	II	NCPPB0446	Banana	1957	Trinidad		
YO190	Ralstonia solanacearum	II	NCPPB0616	Tomato	1957	Trinidad		

YO191	Ralstonia pseudosolanacearum	I	NCPPB2198	Banana	1968	Trinidad		
YO192	Ralstonia solanacearum	II	NCPPB2201	Tomato	1968	Trinidad		
YO193	Ralstonia pseudosolanacearum	I	NCPPB1486	Peanut	1963	Uganda		
YO194	Ralstonia pseudosolanacearum	I	NCPPB2484	Peanut	1969	Uganda		
YO195				Potato	2015	Uganda		
YO196	Ralstonia solanacearum	II		Potato	1992	UK	Oxfordshire	West
YO197	Ralstonia solanacearum	II		Potato	1992	UK	Oxfordshire	West
YO198	Ralstonia solanacearum	II		Potato	1992	UK	Oxfordshire	West
YO199	Ralstonia solanacearum	II	NCPPB3854	Potato	1992	UK	Oxfordshire	West
YO200	Ralstonia solanacearum	II	NCPPB3855	Potato	1992	UK	Oxfordshire	West
YO201	Ralstonia solanacearum	II	NCPPB3856	Potato	1992	UK	Oxfordshire	West
YO202	Ralstonia solanacearum	II	NCPPB3858	Potato	1992	UK	Oxfordshire	West
YO203	Ralstonia solanacearum	II	NCPPB3815	Potato	1992	UK	Oxfordshire	West
YO204	Ralstonia solanacearum	II		Solanum	1993	UK	Oxfordshire	West

				dulcamara				
YO205	Ralstonia solanacearum	II		Solanum dulcamara	1993	UK	Oxfordshire	West
YO206	Ralstonia solanacearum	II		Solanum dulcamara	1993	UK	Oxfordshire	West
YO207	Ralstonia solanacearum	II		Solanum dulcamara	1993	UK	Oxfordshire	West
YO208	Ralstonia solanacearum	II		Solanum dulcamara	1993	UK	Oxfordshire	West
YO209	Ralstonia solanacearum	II		Solanum dulcamara	1993	UK	Oxfordshire	West
YO210	Ralstonia solanacearum	II		Solanum dulcamara	1993	UK	Oxfordshire	West
YO211	Ralstonia solanacearum	II		Solanum dulcamara	1993	UK	Oxfordshire	West
YO212	Ralstonia solanacearum	II		Solanum dulcamara	1993	UK	Oxfordshire	West
YO213	Ralstonia solanacearum	II		Solanum dulcamara	1993	UK	Oxfordshire	West
YO214	Ralstonia solanacearum	II		Solanum dulcamara	1993	UK	Oxfordshire	West
YO215	Ralstonia solanacearum	II		Solanum dulcamara	1993	UK	Oxfordshire	West
YO216	Ralstonia solanacearum	II		Solanum dulcamara	1993	UK	Oxfordshire	West

YO217	Ralstonia solanacearum	II		Solanum dulcamara	1993	UK	Oxfordshire	West
YO218	Ralstonia solanacearum	II		Solanum dulcamara	1993	UK	Oxfordshire	West
YO219	Ralstonia solanacearum	II		Solanum dulcamara	1993	UK	Oxfordshire	West
YO220	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Wiltshire	West
YO221	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Berkshire	South
YO222	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO223	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Berkshire	South
YO224	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO225	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO226	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO227	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO228	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West

YO229	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO230	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Berkshire	South
YO231	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO232	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO233	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO234	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO235	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO236	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO237	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO238	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO239	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO240	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West

YO241	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO242	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO243	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO244	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO245	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO246	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Berkshire	South
YO247	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Berkshire	South
YO248	Ralstonia solanacearum	II		Solanum dulcamara	1994	UK	Oxfordshire	West
YO249	Ralstonia solanacearum	II		Water	1995	UK	Hertfordshire	Severn & West. Mid.
YO250	Ralstonia solanacearum	II		Water	1995	UK	Hertfordshire	Severn & West. Mid.
YO251	Ralstonia solanacearum	II		Water	1995	UK	Hertfordshire	Severn & West. Mid.
YO252	Ralstonia solanacearum	II		Water	1995	UK	Wiltshire	West
YO253	Ralstonia solanacearum	II		Water	1995	UK	Wiltshire	West

YO254	Ralstonia solanacearum	II		Water	1995	UK	Wiltshire	West
YO255	Ralstonia solanacearum	II		Water	1995	UK	Wiltshire	West
YO256	Ralstonia solanacearum	II		Water	1995	UK	Wiltshire	West
YO257	Ralstonia solanacearum	II		Solanum dulcamara	1995	UK	Oxfordshire	West
YO258	Ralstonia solanacearum	II		Water	1995	UK	Hertfordshire	Severn & West. Mid.
YO259	Ralstonia solanacearum	II		Water	1996	UK	Oxfordshire	West
YO260	Ralstonia solanacearum	II		Solanum dulcamara	1996	UK	Greater London	Essex & London
YO261	Ralstonia solanacearum	II		Solanum dulcamara	1996	UK	Buckinghamshire	Essex & London
YO262	Ralstonia solanacearum	II		Solanum dulcamara	1996	UK	Surrey	South
YO263	Ralstonia solanacearum	II		Solanum dulcamara	1996	UK	Surrey	South
YO264	Ralstonia solanacearum	II		Solanum dulcamara	1996	UK	Surrey	South
YO265	Ralstonia solanacearum	II		Solanum dulcamara	1996	UK	Berkshire	South
YO266	Ralstonia solanacearum	II		Solanum dulcamara	1996	UK	Greater London	Essex & London

YO267	Ralstonia solanacearum	II		Potato	1996	UK	Berkshire	South
YO268	Ralstonia solanacearum	II		Water	1996	UK	Greater London	Essex & London
YO269	Ralstonia solanacearum	II		Water	1997	UK	Surrey	South
YO270	Ralstonia solanacearum	II		Water	1997	UK	Berkshire	South
YO271	Ralstonia solanacearum	II		Water	1997	UK	Hampshire	South
YO272	Ralstonia solanacearum	II		Water	1997	UK	Surrey	South
YO273	Ralstonia solanacearum	II		Tomato	1997	UK	Bedfordshire	East Anglia
YO274	Ralstonia solanacearum	II		Tomato	1997	UK	Bedfordshire	East Anglia
YO275	Ralstonia solanacearum	II		Water	1997	UK	Bedfordshire	East Anglia
YO276	Ralstonia solanacearum	II		Water	1997	UK	Bedfordshire	East Anglia
YO277	Ralstonia solanacearum	II		Water	1997	UK	Bedfordshire	East Anglia
YO278	Ralstonia solanacearum	II		Tomato	1997	UK	Bedfordshire	East Anglia
YO279	Ralstonia solanacearum	II		Tomato	1997	UK	Bedfordshire	East Anglia
YO280	Ralstonia solanacearum	II		Water	1997	UK	Bedfordshire	East Anglia
YO281	Ralstonia solanacearum	II		Water	1998	UK	Cambridgeshire	East Anglia
YO282	Ralstonia solanacearum	II		Water	1998	UK	Greater London	Essex &

								London
YO283	Ralstonia solanacearum	II		Water	1998	UK	Bedfordshire	East Anglia
YO284	Ralstonia solanacearum	II		Water	1999	UK	Northamptonshire	East Anglia
YO285	Ralstonia solanacearum	II		Water	1999	UK	Northamptonshire	East Anglia
YO286	Ralstonia solanacearum	II		Potato	1999	UK	Northamptonshire	East Anglia
YO287	Ralstonia solanacearum	II		Water	2000	UK	Kent	South East
YO288	Ralstonia solanacearum	II		Water	2000	UK	Scotland	Scotland
YO289	Ralstonia solanacearum	II		Water	2005	UK	Bedfordshire	East Anglia
YO290	Ralstonia solanacearum	II		Water	2005	UK	Warwickshire	East Midlands
YO291	Ralstonia solanacearum	II		Water	2005	UK	Gloucestershire	West
YO292	Ralstonia solanacearum	II		Water	2005	UK	Gloucestershire	West
YO293	Ralstonia solanacearum	II		Water	2005	UK	Gloucestershire	West
YO294	Ralstonia solanacearum	II		Water	2005	UK	Worcestershire	Severn & West. Mid.
YO295	Ralstonia solanacearum	II		Water	2005	UK	Worcestershire	Severn & West. Mid.
YO296	Ralstonia solanacearum	II		Water	2005	UK	Gloucestershire	West
YO297	Ralstonia solanacearum	II		Water	2005	UK	Bedfordshire	East Anglia

YO298	Ralstonia solanacearum	II		Water	2005	UK	Cambridgeshire	East Anglia
YO299	Ralstonia solanacearum	II		Water	2005	UK	Bedfordshire	East Anglia
YO300	Ralstonia solanacearum	II		Water	2005	UK	Cambridgeshire	East Anglia
YO301	Ralstonia solanacearum	II		Potato	2005	UK	Nottinghamshire	East Midlands
YO302	Ralstonia solanacearum	II		Water	2006	UK	Lincolnshire	Humberside
YO303	Ralstonia solanacearum	II		Water	2006	UK	Carmarthenshire	Wales
YO304	Ralstonia solanacearum	II		Water	2006	UK	Warwickshire	East Midlands
YO305	Ralstonia solanacearum	II		Water	2006	UK	Mid Glamorgan	Wales
YO306	Ralstonia solanacearum	II		Water	2006	UK	Cambridgeshire	East Anglia
YO307	Ralstonia solanacearum	II		Water	2006	UK	Cambridgeshire	East Anglia
YO308	Ralstonia solanacearum	II		Water	2006	UK	Dorset	South West
YO309	Ralstonia solanacearum	II		Water	2006	UK	Dorset	South West
YO310	Ralstonia solanacearum	II		Water	2006	UK	Carmarthenshire	Wales
YO311	Ralstonia solanacearum	II		Water	2006	UK	Carmarthenshire	Wales
YO312	Ralstonia solanacearum	II		Water	2006	UK	West Midlands	Severn & West. Mid.
YO313	Ralstonia solanacearum	II		Water	2006	UK	West Midlands	Severn &

								West. Mid.
YO314	Ralstonia solanacearum	II		Water	2006	UK	Warwickshire	East Midlands
YO315	Ralstonia solanacearum	II		Water	2006	UK	Warwickshire	East Midlands
YO316	Ralstonia solanacearum	II		Water	2006	UK	Carmarthenshire	Wales
YO317	Ralstonia solanacearum	II		Water	2006	UK	Carmarthenshire	Wales
YO318	Ralstonia solanacearum	II		Water	2006	UK	Dorset	South West
YO319	Ralstonia solanacearum	II		Water	2006	UK	Mid Glamorgan	Wales
YO320	Ralstonia solanacearum	II		Water	2006	UK	Bedfordshire	East Anglia
YO321	Ralstonia solanacearum	II		Water	2006	UK	Warwickshire	East Midlands
YO322	Ralstonia solanacearum	II		Water	2006	UK	Staffordshire	Severn & West. Mid.
YO323	Ralstonia solanacearum	II		Water	2006	UK	Carmarthenshire	Wales
YO324	Ralstonia solanacearum	II		Water	2006	UK	Bedfordshire	East Anglia
YO325	Ralstonia solanacearum	II		Water	2006	UK	Warwickshire	East Midlands
YO326	Ralstonia solanacearum	II		Water	2006	UK	Bedfordshire	East Anglia
YO327	Ralstonia solanacearum	II		Water	2006	UK	Dorset	South West
YO328	Ralstonia solanacearum	II		Water	2006	UK	Carmarthenshire	Wales

YO329	Ralstonia solanacearum	II		Water	2006	UK	Mid Glamorgan	Scotland
YO330	Ralstonia solanacearum	II		Water	2006	UK	Warwickshire	East Midlands
YO331	Ralstonia solanacearum	II		Water	2006	UK	Carmarthenshire	Wales
YO332	Ralstonia solanacearum	II		Water	2006	UK	Cambridgeshire	East Anglia
YO333	Ralstonia solanacearum	II		Water	2006	UK	Carmarthenshire	Wales
YO334	Ralstonia solanacearum	II		Water	2006	UK	Cambridgeshire	East Anglia
YO335	Ralstonia solanacearum	II		Water	2007	UK	Hampshire	South
YO336	Ralstonia solanacearum	II		Water	2007	UK	Cambridgeshire	East Anglia
YO337	Ralstonia solanacearum	II		Water	2007	UK	Cambridgeshire	East Anglia
YO338	Ralstonia solanacearum	II		Water	2007	UK	Surrey	South
YO339	Ralstonia solanacearum	II		Water	2007	UK	Hampshire	South
YO340	Ralstonia solanacearum	II		Water	2007	UK	Berkshire	South
YO341	Ralstonia solanacearum	II		Water	2007	UK	Berkshire	South
YO342	Ralstonia solanacearum	II		Water	2007	UK	Berkshire	South
YO343	Ralstonia solanacearum	II		Water	2007	UK	Berkshire	South
YO344	Ralstonia solanacearum	II		Water	2008	UK	Oxfordshire	West

YO345	Ralstonia solanacearum	II		Solanum dulcamara	2008	UK	Oxfordshire	West
YO346	Ralstonia solanacearum	II		Solanum dulcamara	2008	UK	Oxfordshire	West
YO347	Ralstonia solanacearum	II		Solanum dulcamara	2008	UK	Oxfordshire	West
YO348	Ralstonia solanacearum	II		Water	2008	UK	Oxfordshire	West
YO349	Ralstonia solanacearum	II		Potato	2009	UK	Cornwall	South West
YO350	Ralstonia solanacearum	II		Water	2013	UK	Berkshire	South
YO351	Ralstonia solanacearum	II		Water	2013	UK	Berkshire	South
YO352	Ralstonia solanacearum	II		Water	2014	UK	Berkshire	South
YO353	Ralstonia solanacearum	II		Water	2014	UK	Berkshire	South
YO354	Ralstonia solanacearum	II		Water	2015	UK	Berkshire	South
YO355	Ralstonia solanacearum	II		Water	2015	UK	Berkshire	South
YO356	Ralstonia solanacearum	II		Water	2016	UK	Cambridgeshire	East Anglia
YO357	Ralstonia solanacearum	II		Water	2016	UK	Cambridgeshire	East Anglia
YO358	Ralstonia solanacearum	II		Water	2016	UK	Norfolk	East Anglia
YO359	Ralstonia solanacearum	II		Water	2016	UK	Norfolk	East Anglia

YO360	Ralstonia solanacearum	II		Water	2017	UK	Norfolk	East Anglia
YO361	Ralstonia solanacearum	II		Water	2017	UK	Norfolk	East Anglia
YO362	Ralstonia solanacearum	II		Water	2017	UK	Cambridgeshire	East Anglia
YO363	Ralstonia solanacearum	II		Water	2017	UK	Norfolk	East Anglia
YO364	Ralstonia solanacearum	II		Water	2018	UK	Norfolk	East Anglia
YO365	Ralstonia solanacearum	II		Water	2018	UK	Cambridgeshire	East Anglia
YO366	Ralstonia solanacearum	II		Water	2018	UK	Norfolk	East Anglia
YO367	Ralstonia solanacearum	II		Water	2018	UK	Cambridgeshire	East Anglia
YO368	Ralstonia solanacearum	II		Water	2018	UK	Norfolk	East Anglia
YO369	Ralstonia solanacearum	II		Water	2018	UK	Cambridgeshire	East Anglia
YO370	Ralstonia solanacearum	II		Water		UK		Scotland
YO371	Ralstonia solanacearum	II		Water		UK		Scotland
YO372	Ralstonia solanacearum	II	NCPPB325, NCPPB3973	Tomato	1953	USA	USA	
YO373	Ralstonia solanacearum	II	NCPPB0337	Tobacco	1954	USA		

YO374	Ralstonia solanacearum	II	NCPPB0338	Tobacco	1954	USA		
YO375	Ralstonia pseudosolanacearum	I	NCPPB4009	Tobacco	1955	USA		
YO376	Ralstonia pseudosolanacearum	I	NCPPB1580	Tomato	1959	USA		
YO377			NCPPB927	Sugarcane	1961	USA		
YO378	Ralstonia pseudosolanacearum	I	NCPPB1579	Ginger	1961	USA		
YO379	Ralstonia pseudosolanacearum	I	NCPPB1581	Strelitzia reginae	1961	USA		
YO380	Ralstonia solanacearum	II	NCPPB4362	Pelargonium hortorum	2003	USA		
YO381	Ralstonia solanacearum	II	NCPPB3969	Banana		Venezuela		
YO382	Ralstonia pseudosolanacearum	III	NCPPB0283	Solanum panduraforme	1950	Zimbabwe		
YO383	Ralstonia pseudosolanacearum	III	NCPPB0332	Potato	1954	Zimbabwe		
YO384	Ralstonia pseudosolanacearum	III	NCPPB0505	Comfrey	1956	Zimbabwe		
YO385				Solanum dulcamara	2019	UK		West
YO386				Water	2019	UK		West

YO387				Solanum dulcamara	2019	UK		West
YO388				Water	2019	UK		West
YO389				Solanum dulcamara	2019	UK		West
YO390				Water	2019	UK		West

7. APPENDIX 2

ID	name	Species York	Phylotype York	IIB-1	Assigned Phylotype web
GCA_000009125.1	GMI1000	Ralstonia pseudosolanacearum	I		
GCA_000020205.1	12J	Ralstonia pickettii			
GCA_000215325.1	Po82	Ralstonia solanacearum	II		
GCA_000283475.1	PSI07	Ralstonia syzygii	IV		
GCA_001267515.1	YC45	Ralstonia pseudosolanacearum	I		
GCA_001299555.1	UY031	Ralstonia solanacearum	II	IIB-1	
GCA_001586135.1	KACC10722	Ralstonia syzygii	IV		

GCA_001587135.1	UW163	Ralstonia solanacearum	II		IIB (https://datamed.org/display-item.php?repository=0008&idName=ID&id=5914e5525152c67771b5b7d8)
GCA_001587155.1	IBSBF1503	Ralstonia solanacearum	II		
GCA_001663415.1	YC40-M	Ralstonia pseudosolanacearum	I		
GCA_001708525.1	KACC10709	Ralstonia pseudosolanacearum	I		
GCA_001879565.1	OE1-1	Ralstonia pseudosolanacearum	I		
GCA_001887535.1	FJAT-1458	Ralstonia pseudosolanacearum	I		
GCA_001891105.1	EP1	Ralstonia pseudosolanacearum	I		
GCA_002155245.1	FJAT-91	Ralstonia pseudosolanacearum	I		
GCA_002162015.1	SEPPX05	Ralstonia pseudosolanacearum	I		
GCA_002220465.1	CQPS-1	Ralstonia pseudosolanacearum	I		
GCA_002501565.1	RS488	Ralstonia solanacearum	II	IIB-1	
GCA_002549815.1	RS489	Ralstonia solanacearum	II		
GCA_002894285.1	RSCM	Ralstonia pseudosolanacearum	I		
GCA_003515145.1	T51	Ralstonia syzygii	IV		
GCA_003515165.1	T11	Ralstonia syzygii	IV		

GCA_0035151 85.1	SL3175	Ralstonia syzygii	IV		
GCA_0035152 05.1	SL3103	Ralstonia pseudosolan acearum	I		
GCA_0035152 25.1	SL2330	Ralstonia pseudosolan acearum	I		
GCA_0035152 45.1	T117	Ralstonia pseudosolan acearum	I		
GCA_0035152 65.1	T98	Ralstonia syzygii	IV		
GCA_0035152 85.1	T78	Ralstonia pseudosolan acearum	I		
GCA_0035153 05.1	T25	Ralstonia pseudosolan acearum	I		
GCA_0035153 25.1	T12	Ralstonia syzygii	IV		
GCA_0035153 45.1	SL3755	Ralstonia pseudosolan acearum	I		
GCA_0035153 65.1	SL3730	Ralstonia pseudosolan acearum	I		
GCA_0035153 85.1	SL3022	Ralstonia syzygii	IV		
GCA_0035154 05.1	SL2729	Ralstonia pseudosolan acearum	I		
GCA_0035154 25.1	SL2312	Ralstonia syzygii	IV		
GCA_0035154 45.1	SL2064	Ralstonia syzygii	IV		
GCA_0035154 65.1	T110	Ralstonia pseudosolan acearum	I		

GCA_003515485.1	T101	Ralstonia syzygii	IV		
GCA_003515505.1	T95	Ralstonia syzygii	IV		
GCA_003515525.1	T82	Ralstonia syzygii	IV		
GCA_003515545.1	T60	Ralstonia pseudosolanacearum	I		
GCA_003515565.1	T42	Ralstonia pseudosolanacearum	I		
GCA_003515585.1	SL3882	Ralstonia pseudosolanacearum	I		
GCA_003515605.1	SL3822	Ralstonia pseudosolanacearum	I		
GCA_003515625.1	SL3300	Ralstonia pseudosolanacearum	I		
GCA_003590585.1	IBSBF 2570	Ralstonia solanacearum	II		
GCA_003595305.1	RS 476	Ralstonia pseudosolanacearum	I		
GCA_003612975.1	CRMRs218	Ralstonia solanacearum	I		

8. APPENDIX 3

Component	Gene	Annotation
210	group_10907	hypothetical protein
210	group_4078	hypothetical protein
209	group_5148	hypothetical protein
209	group_853	hypothetical protein
208	group_8075	hypothetical protein
208	group_7870	IS110 family transposase ISPye16
208	group_7304	hypothetical protein
208	group_5103	hypothetical protein
207	group_5138	hypothetical protein
207	estP	Esterase EstP;hypothetical protein
206	group_6368	hypothetical protein
206	group_1854	hypothetical protein
205	dmlR_15	HTH-type transcriptional regulator DmlR
205	group_8749	hypothetical protein
205	polS_2	Sorbitol dehydrogenase;Galactitol 2-dehydrogenase
204	group_415	hypothetical protein
204	group_433	hypothetical protein
203	group_12047	Outer membrane porin protein;hypothetical protein
203	yphB~~~yphB_1~~~yphB_2	putative protein YphB
203	group_10422	hypothetical protein
203	group_3466	Outer membrane porin protein 32;hypothetical protein
203	group_1431	hypothetical protein
203	araG_1	Arabinose import ATP-binding protein AraG
203	pat~~~pat_1	Phosphinothricin N-acetyltransferase;hypothetical protein
203	group_11865	hypothetical protein
203	group_11717	hypothetical protein
203	rbn_2~~~rbn_3~~~rbn_1	Ribonuclease BN
203	group_1358	hypothetical protein
203	umaA	S-adenosylmethionine-dependent methyltransferase UmaA
202	group_10028	HTH-type transcriptional regulator HdfR

202	msrP_3	Protein-methionine-sulfoxide reductase catalytic subunit MsrP
202	group_8186	Nitrilase
202	metC_2	Cystathionine beta-lyase MetC
202	mdcG	Phosphoribosyl-dephospho-CoA transferase
202	madA	Acetyl-S-ACP:malonate ACP transferase
202	mdcC	Malonate decarboxylase acyl carrier protein
202	group_1698	hypothetical protein
202	group_950	hypothetical protein
201	group_955	hypothetical protein
201	group_771	hypothetical protein
200	group_4211	hypothetical protein
200	group_4210	hypothetical protein
200	group_2888	hypothetical protein
199	group_3738	Transposon Tn7 transposition protein TnsB;hypothetical protein
199	group_3736	hypothetical protein
198	tsr_3~~~tsr_1~~~tar_1~~~tar_2	Methyl-accepting chemotaxis protein I;hypothetical protein;Methyl-accepting chemotaxis protein II
198	pgl_4~~~pgl_3	6-phosphogluconolactonase
197	group_2411	putative protein;hypothetical protein
197	acrR	hypothetical protein;HTH-type transcriptional regulator AcrR
196	group_10659	hypothetical protein
196	group_10655	IS21 family transposase ISRso6
195	group_2051	hypothetical protein
195	group_2050	hypothetical protein
195	yraJ	hypothetical protein;Outer membrane usher protein YraJ
195	group_1959	hypothetical protein
195	group_1887	hypothetical protein
195	group_1843	hypothetical protein
194	group_9124	Cystathionine beta-lyase PatB
194	group_7630	hypothetical protein
194	sqr	hypothetical protein;Sulfide-quinone reductase
193	copB	Copper resistance protein B
193	copA_3~~~copA_2~~~copA_4~~~copA_5	Copper resistance protein A;hypothetical protein

193	cusR_1~~~cusR_2~~~rcsC_7~~~cusR	Transcriptional regulatory protein CusR;Sensor histidine kinase RcsC
193	sasA_12~~~cusS~~~sasA_8~~~cusS_2~~~cusS_1	Adaptive-response sensory-kinase SasA;Sensor histidine kinase CusS;hypothetical protein
192	group_8066	hypothetical protein
192	group_6981	hypothetical protein;Vitamin B12 import ATP-binding protein BtuD
192	group_5464	hypothetical protein
192	group_1473	hypothetical protein
192	group_901	hypothetical protein
191	cmpR_2	hypothetical protein;HTH-type transcriptional activator CmpR
191	group_2062	hypothetical protein
191	group_1984	putative HTH-type transcriptional regulator
191	naiP_1~~~yjhB~~~naiP_2~~~naiP	Putative niacin/nicotinamide transporter NaiP;Putative metabolite transport protein YjhB;hypothetical protein
190	group_10934	hypothetical protein
190	group_4385	hypothetical protein
189	group_11788	hypothetical protein
189	group_7113	hypothetical protein
189	group_5015	hypothetical protein
189	fecI_2	putative RNA polymerase sigma factor FecI
189	ttdB_2	L(+)-tartrate dehydratase subunit beta
189	group_3783	hypothetical protein
189	pglA_2	Polygalacturonase
188	copD	Copper resistance protein D
188	copC_1~~~copC_2~~~copC	Copper resistance protein C
187	dctA2_3~~~dctA2_2	C4-dicarboxylate transport protein 2
187	pcaB_1~~~pcaB_2	3-carboxy-ciscis-muconate cycloisomerase
187	dasR	HTH-type transcriptional repressor DasR
186	group_11553	hypothetical protein
186	group_11097	hypothetical protein
186	group_10337	hypothetical protein
186	group_9442	hypothetical protein
185	group_6703	hypothetical protein
185	group_4022	hypothetical protein
185	rep_3	ATP-dependent DNA helicase Rep

184	group_9260	hypothetical protein
184	group_9259	hypothetical protein
184	group_7171	hypothetical protein
184	group_1132	3'-5' exoribonuclease
183	sutR_2~~~sutR_3~~~sutR_1~~~sutR_4	hypothetical protein;HTH-type transcriptional regulator SutR
183	ttr	Acetyltransferase
182	group_10911	hypothetical protein
182	group_4024	hypothetical protein
181	ompR_3	hypothetical protein;Transcriptional regulatory protein OmpR
181	hopD2	Effector protein hopD2
181	group_4152	hypothetical protein
180	group_7631	hypothetical protein
180	nemA_3	N-ethylmaleimide reductase
180	group_4700	hypothetical protein
180	acuR	Transcriptional regulator AcuR;hypothetical protein
179	group_6094	hypothetical protein
179	group_8221	hypothetical protein
178	intA_1~~~intA_2	Prophage integrase IntA
178	group_94	hypothetical protein
177	group_10997	hypothetical protein
177	group_6262	hypothetical protein
177	group_4839	hypothetical protein
177	group_3054	hypothetical protein
176	xerC_4~~~xerC_1	Tyrosine recombinase XerC;hypothetical protein
176	group_747	hypothetical protein
175	group_8175	IS30 family transposase ISHar4;hypothetical protein;IS30 family transposase ISHar5
175	group_4611	hypothetical protein
174	group_3211	hypothetical protein
174	virS	hypothetical protein;HTH-type transcriptional regulator VirS
174	actIII~~~hcaB_2	Putative ketoacyl reductase;hypothetical protein;3-phenylpropionate-dihydrodiol/cinnamic acid-dihydrodiol dehydrogenase
173	group_3182	hypothetical protein
173	fprA_2~~~fprA_1~~~fprB~~~fprA	NADPH-ferredoxin reductase FprA;putative

		ferredoxin/ferredoxin--NADP reductase
173	moaA_2~~~moaA_3~~~moaA_1	GTP 3'8-cyclase
173	ynfF~~~ynfE~~~napA_2	putative dimethyl sulfoxide reductase chain YnfF;hypothetical protein;Putative dimethyl sulfoxide reductase chain YnfE;Periplasmic nitrate reductase
173	fdxA	Ferredoxin-1
173	rhaR_1~~~rhaR_2	HTH-type transcriptional activator RhaR
172	group_3731	hypothetical protein
172	hxuB	Heme/hemopexin transporter protein HuxB
171	group_1146	hypothetical protein
171	group_631	hypothetical protein
170	group_10040	hypothetical protein
170	group_8993	hypothetical protein
170	virB1	hypothetical protein;Type IV secretion system protein virB1
170	group_8076	hypothetical protein
170	group_6727	hypothetical protein
170	group_6725	hypothetical protein
170	group_5865	hypothetical protein
169	xerD_3~~~xerC_5~~~xerC_11	Tyrosine recombinase XerD;Tyrosine recombinase XerC
169	group_3024	hypothetical protein;Chromosome partition protein Smc
168	group_1156	hypothetical protein
168	group_1268	IS21 family transposase ISRso19
167	group_7876	hypothetical protein
167	group_11945	hypothetical protein
166	phnZ	hypothetical protein;2-amino-1-hydroxyethylphosphonate dioxygenase (glycine-forming)
166	group_7067	hypothetical protein
165	group_2944	putative HTH-type transcriptional regulator
165	hcaB_2~~~hcaB_1~~~ydfG_3	3-phenylpropionate-dihydrodiol/cinnamic acid-dihydrodiol dehydrogenase;hypothetical protein;NADP-dependent 3-hydroxy acid dehydrogenase YdfG

164	cutC~~~pflD	Choline trimethylamine-lyase;hypothetical protein;Trans-4-hydroxy-L-proline dehydratase
164	group_1683	hypothetical protein
163	group_1532	hypothetical protein
163	group_1476	hypothetical protein
162	group_2243	hypothetical protein
162	group_2208	hypothetical protein
162	group_2197	hypothetical protein
162	menH_3~~~menH_2~~~menH_4	2-succinyl-6-hydroxy-24-cyclohexadiene-1-carboxylate synthase;hypothetical protein
162	group_2181	hypothetical protein
162	novN_1	hypothetical protein;Decarbamoylnovobiocin carbamoyltransferase
162	group_2163	hypothetical protein;4-hydroxyphenylalkanoate adenylyltransferase
162	group_2162	hypothetical protein
162	group_2141	hypothetical protein
162	group_2139	hypothetical protein
162	valS_2	Valine--tRNA ligase;hypothetical protein
162	metZ~~~metZ_1	O-succinylhomoserine sulfhydrylase;hypothetical protein
162	anoI_2~~~anoI_1	Acyl-homoserine-lactone synthase;hypothetical protein
162	ycaC_2	hypothetical protein;putative hydrolase YcaC
162	anoR_2~~~anoR_1	Transcriptional activator protein AnoR
162	dmlR_13~~~dmlR_19~~~dmlR_12	hypothetical protein;HTH-type transcriptional regulator DmlR
162	sucA_2~~~sucA_3	2-oxoglutarate dehydrogenase E1 component
162	tar_7	hypothetical protein;Methyl-accepting chemotaxis protein II
161	dmlR_18~~~dmlR_5~~~dmlR_10~~~cynR_2~~~dmlR_6~~~dmlR_2	HTH-type transcriptional regulator DmlR;HTH-type transcriptional regulator CynR;hypothetical protein
161	menH_2	hypothetical protein;2-succinyl-6-hydroxy-24-cyclohexadiene-1-carboxylate synthase
160	group_9896	hypothetical protein
160	group_8426	hypothetical protein

159	group_7300	hypothetical protein
159	group_6277	hypothetical protein
158	hchA~~~hchA_1	Protein/nucleic acid deglycase HchA
158	ligJ	hypothetical protein;2-keto-4-carboxy-3-hexenedioate hydratase
157	group_3228	hypothetical protein
157	group_3253	hypothetical protein
157	group_3252	hypothetical protein
157	group_3233	hypothetical protein
157	group_3229	hypothetical protein
157	group_3218	hypothetical protein
156	group_3222	hypothetical protein
156	group_3570	hypothetical protein
155	stp~~~emrY_4~~~emrB_3~~~bmr3~~~stp_2~~~stp_1~~~emrB_4	Multidrug resistance protein Stp;putative multidrug resistance protein EmrY;Multidrug export protein EmrB;Multidrug resistance protein 3;hypothetical protein
155	dhaA~~~dhaA_2~~~rutD_2	Haloalkane dehalogenase;Arylesterase;Putative aminoacrylate hydrolase RutD;hypothetical protein
155	pksJ~~~menE_3~~~menE_4~~~menE_2	Polyketide synthase PksJ;2-succinylbenzoate--CoA ligase;hypothetical protein
154	group_8974	hypothetical protein
154	virB9	Type IV secretion system protein virB9
154	group_8316	hypothetical protein
154	group_8137	Type IV secretion system protein VirB11
154	virB4_2~~~virB4	Type IV secretion system protein virB4
154	group_8077	hypothetical protein
154	group_7530	hypothetical protein
154	group_6687	hypothetical protein
154	virB8	Type IV secretion system protein virB8;hypothetical protein
154	group_6194	hypothetical protein
154	group_5958	hypothetical protein
154	group_5957	hypothetical protein
154	virB5	Type IV secretion system protein virB5
154	group_4321	hypothetical protein
154	group_3953	hypothetical protein

154	group_3435	hypothetical protein
154	group_3434	hypothetical protein
154	group_3431	hypothetical protein
154	group_3134	hypothetical protein
154	group_2512	hypothetical protein
153	group_6121	hypothetical protein
153	hipA_2	Serine/threonine-protein kinase toxin HipA
152	group_10318	HTH-type transcriptional regulator GltC
152	cuyA	L-cysteate sulfo-lyase
152	sotB_2	sugar efflux transporter
151	xerC_1	hypothetical protein;Tyrosine recombinase XerC
151	group_7894	hypothetical protein
151	group_995	hypothetical protein
150	group_1572	hypothetical protein
150	group_1677	hypothetical protein
149	group_1197	hypothetical protein
149	group_1299	hypothetical protein
148	group_11290	hypothetical protein
148	group_10641	hypothetical protein
147	group_7554	hypothetical protein
147	group_2858	hypothetical protein
147	group_7555	hypothetical protein
146	group_10752	hypothetical protein
146	group_10258	hypothetical protein
146	group_9300	hypothetical protein
146	group_9254	hypothetical protein
146	group_7327	hypothetical protein
146	group_7258	hypothetical protein
146	group_7163	hypothetical protein
145	rhsC_3~~~rhsD	Putative deoxyribonuclease RhsC;Protein RhsD;hypothetical protein
145	group_11777	hypothetical protein
145	group_6009	hypothetical protein
145	sasA_15	Adaptive-response sensory-kinase SasA
145	rssB_4	Regulator of RpoS
145	group_4680	hypothetical protein
144	fliI_2~~~fliI_1	Flagellum-specific ATP synthase
144	group_3643	hypothetical protein

144	fliE_2~~~fliE_1	Flagellar hook-basal body complex protein FliE
144	flgC_2~~~flgC_1	Flagellar basal-body rod protein FlgC
144	group_3626	hypothetical protein
144	flgI_2~~~flgI_1	Flagellar P-ring protein
144	flgH_2~~~flgH_1	Flagellar L-ring protein
144	group_3569	hypothetical protein
144	flgF_2~~~flgF_1	Flagellar basal-body rod protein FlgF
144	flgG_2~~~flgG_3~~~flgG_1	Flagellar basal-body rod protein FlgG
144	flhA_1~~~flhA_2	Flagellar biosynthesis protein FlhA
144	flhB_1	Flagellar biosynthetic protein FlhB
144	fliR_1	hypothetical protein;Flagellar biosynthetic protein FliR
144	group_3096	hypothetical protein
143	group_582	4-hydroxyphenylalkanoate adenylyltransferase;Putative fatty-acid--CoA ligase FadD21;Long-chain-fatty-acid--CoA ligase FadD23;Long-chain-fatty-acid--AMP ligase FadD30
143	group_551	hypothetical protein
143	ppsA_2	hypothetical protein;Phthiocerol synthesis polyketide synthase type I PpsA
143	ppsB	hypothetical protein;Phthiocerol/phenolphthiocerol synthesis polyketide synthase type I PpsB
143	group_347	hypothetical protein
143	lgrB~~~ppsE	Linear gramicidin synthase subunit B;hypothetical protein;Phthiocerol synthesis polyketide synthase type I PpsE
143	pikAV~~~pikAV_2	Thioesterase PikA5
142	group_5844	hypothetical protein
142	group_286	hypothetical protein
141	group_2103	hypothetical protein
141	group_7675	hypothetical protein
140	group_10366	hypothetical protein
140	group_9473	hypothetical protein
139	group_2516	hypothetical protein
139	group_2364	hypothetical protein
138	group_11617	hypothetical protein
138	group_11895	hypothetical protein

137	group_2104	hypothetical protein
137	group_2069	hypothetical protein
136	mshA_2~~~mshA_3~~~mshA_4	D-inositol-3-phosphate glycosyltransferase;hypothetical protein
136	cya~~~cya_2~~~cya_1	hypothetical protein;Bifunctional hemolysin/adenylate cyclase
135	group_1811	hypothetical protein
135	group_1521	hypothetical protein
135	group_1029	hypothetical protein
134	group_11638	hypothetical protein
134	group_10515	hypothetical protein
133	group_9987	hypothetical protein
133	group_9672	hypothetical protein
133	group_8880	hypothetical protein
133	group_6566	hypothetical protein
133	ttdB~~~ttdB_1~~~ttdB_2	L(+)-tartrate dehydratase subunit beta
132	mmgC_5	hypothetical protein;Acyl-CoA dehydrogenase fadE12;Acyl-CoA dehydrogenase
132	group_3051	hypothetical protein
132	yofA	HTH-type transcriptional regulator YofA;hypothetical protein
131	group_2310	hypothetical protein
131	group_3410	hypothetical protein
130	group_7491	hypothetical protein
130	group_1844	hypothetical protein
129	sutR_3~~~sutR_4	hypothetical protein;HTH-type transcriptional regulator SutR
129	rhtB_2~~~rhtB_3	Homoserine/homoserine lactone efflux protein
128	group_4825	hypothetical protein
128	group_1932	hypothetical protein
128	group_7495	hypothetical protein
127	group_9159	hypothetical protein
127	group_8998	Protein kinase YegI
127	group_6994	hypothetical protein
127	group_6733	Serine/threonine-protein phosphatase 3
126	group_3385	hypothetical protein
126	group_11901	hypothetical protein
126	group_285	hypothetical protein
126	group_58	hypothetical protein
125	group_11976	hypothetical protein

125	group_11735	hypothetical protein
125	group_11397	hypothetical protein
125	soj_2~~~soj_1	Chromosome-partitioning ATPase Soj
125	group_10851	hypothetical protein
125	group_11975	hypothetical protein
125	group_11736	hypothetical protein
125	group_122	hypothetical protein
125	group_746	hypothetical protein
125	group_2810	hypothetical protein
125	group_2809	hypothetical protein
125	group_1750	hypothetical protein
125	group_1445	hypothetical protein
125	group_1475	hypothetical protein
124	yqcF	Antitoxin YqcF
124	group_9151	hypothetical protein
124	group_5443	hypothetical protein
123	cnrA_2~~~cnrA_1~~~cnrA	Nickel and cobalt resistance protein CnrA
123	cnrB	Nickel and cobalt resistance protein CnrB
123	cnrC	Nickel and cobalt resistance protein CnrC
123	cnrH~~~cnrH_1	RNA polymerase sigma factor CnrH
123	cnrR	Nickel and cobalt resistance protein CnrR
123	cnrY	Nickel and cobalt resistance protein CnrY
122	group_5040	hypothetical protein
122	group_7201	hypothetical protein
122	group_2735	hypothetical protein
121	group_2338	hypothetical protein
121	group_2989	hypothetical protein
121	group_2504	hypothetical protein
120	group_371	hypothetical protein
120	group_302	hypothetical protein
119	group_8127	hypothetical protein
119	group_3913	hypothetical protein
118	group_10292	hypothetical protein
118	group_9375	hypothetical protein
117	group_7697	hypothetical protein
117	group_4008	hypothetical protein
116	group_3337	hypothetical protein
116	group_3053	hypothetical protein

	mdtA_9~~~mdtA_4~~~mdtA_7~~~mdtA_3	Multidrug resistance protein MdtA;Cation efflux system protein
116	~~~cusB~~~mdtA_2	CusB
116	cusA~~~cusA_1~~~cusA_2	Cation efflux system protein CusA
116	group_2790	hypothetical protein
115	creB	Transcriptional regulatory protein CreB
115	creC	Sensor protein CreC
115	creD	Inner membrane protein CreD;hypothetical protein
114	emrB_1~~~stp~~~stp_1~~~emrB_3~~~mdtD_1~~~mdtD_2~~~mdtD~~~emrB_5	hypothetical protein;Multidrug export protein EmrB;Multidrug resistance protein Stp;Putative multidrug resistance protein MdtD
114	group_3265	hypothetical protein
113	group_2884	hypothetical protein
113	group_2829	hypothetical protein
112	group_11001	hypothetical protein
112	tar_2~~~tar_1	Methyl-accepting chemotaxis protein II
112	group_929	hypothetical protein
111	oprJ~~~oprM_7	Outer membrane protein OprJ;Outer membrane protein OprM
111	hcaB_3~~~hcaB_1~~~hcaB	3-phenylpropionate-dihydrodiol/cinnamic acid-dihydrodiol dehydrogenase
110	group_3036	hypothetical protein
110	group_3746	IS3 family transposase ISButh1
109	xerC_5~~~xerC_10	Tyrosine recombinase XerC
109	group_6604	hypothetical protein
109	group_8992	hypothetical protein
109	group_8918	hypothetical protein
109	group_10003	hypothetical protein
109	ttuB_1~~~ttuB_2	Putative tartrate transporter
109	group_8214	hypothetical protein
109	group_6154	hypothetical protein
108	group_2946	hypothetical protein;RutC family protein YjgH
108	lrpC~~~lrpC_3~~~lrpC_2	hypothetical protein;HTH-type transcriptional regulator LrpC
108	hpd_2~~~lly~~~hpd_1	4-hydroxyphenylpyruvate dioxygenase;hypothetical protein
107	pgrR_5~~~pgrR_1	HTH-type transcriptional regulator PgrR

107	group_3364	hypothetical protein
107	group_3342	hypothetical protein
107	ycjY_1~~~ycjY	putative protein YcjY;hypothetical protein
107	yvgN	putative oxidoreductase/MSMEI_2346;hypothetical protein;Glyoxal reductase
107	efpA	putative MFS-type transporter EfpA;hypothetical protein
106	group_1558	hypothetical protein
106	group_1559	hypothetical protein
105	group_821	hypothetical protein
105	group_532	hypothetical protein
105	group_10664	hypothetical protein
105	group_9516	hypothetical protein
105	group_7666	hypothetical protein
105	group_4996	hypothetical protein
105	group_3326	hypothetical protein
104	group_5833	hypothetical protein
104	group_1158	hypothetical protein
103	group_1464	hypothetical protein
103	group_1429	hypothetical protein
102	group_3505	hypothetical protein
102	group_3291	hypothetical protein
102	group_835	hypothetical protein
101	group_9433	hypothetical protein
101	group_6578	hypothetical protein
100	group_3289	hypothetical protein
100	group_2567	hypothetical protein
100	group_2573	hypothetical protein
99	group_89	hypothetical protein
99	group_1168	hypothetical protein
98	group_7890	hypothetical protein
98	group_2161	hypothetical protein
98	group_2160	hypothetical protein
98	group_2144	hypothetical protein
98	clpP_1~~~clpP_2	ATP-dependent Clp protease proteolytic subunit
98	group_2109	hypothetical protein
98	group_2077	hypothetical protein
97	group_3733	hypothetical protein
97	xerC_1~~~xerC_4	Tyrosine recombinase XerC;hypothetical protein
96	group_8220	hypothetical protein

96	group_3204	hypothetical protein
96	gltI_6~~~gltI_7	Glutamate/aspartate import solute-binding protein
95	group_1180	hypothetical protein
95	group_1133	hypothetical protein
94	group_9545	hypothetical protein
94	group_6870	hypothetical protein
93	group_727	hypothetical protein
93	group_10458	hypothetical protein;Type I restriction enzyme EcoKI M protein
92	group_10414	hypothetical protein
92	higB~~~higB_2	Endoribonuclease HigB
92	group_7682	hypothetical protein
92	group_2799	hypothetical protein
92	group_1712	hypothetical protein
92	group_1631	hypothetical protein
92	group_1161	hypothetical protein
92	group_1274	hypothetical protein
91	ccr~~~pikAII	Crotonyl-CoA reductase;Narbonolide/10-deoxymethynolide synthase PikA2 modules 3 and 4
91	group_4783	hypothetical protein
91	fabG_4~~~fabG_2	hypothetical protein;3-oxoacyl-[acyl-carrier-protein] reductase FabG
91	group_9456	hypothetical protein
90	group_14014	hypothetical protein
90	group_14012	hypothetical protein
89	group_2362	hypothetical protein
89	group_890	hypothetical protein
89	group_10120	hypothetical protein
89	group_10268	hypothetical protein
89	group_8771	ATP-dependent zinc metalloprotease FtsH
89	group_3809	hypothetical protein
89	group_7246	hypothetical protein
88	group_2824	hypothetical protein
88	group_1170	hypothetical protein
87	group_9938	hypothetical protein
87	group_838	hypothetical protein
86	group_508	hypothetical protein
86	group_552	hypothetical protein
86	group_11904	hypothetical protein

86	group_11633	hypothetical protein
86	epsE~~~hxcR_2	hypothetical protein;Type II secretion system protein E;putative type II secretion system protein HxcR
86	epsG~~~xcpT_3~~~xcpT_1~~~xcpT_2	Type II secretion system protein G;hypothetical protein
86	group_53	hypothetical protein
86	sctC_6~~~sctC_7~~~sctC_4~~~sctC_5~~~sctC_8~~~sctC_2~~~sctC_1	Type 3 secretion system secretin;hypothetical protein
86	xcpT_3~~~xcpT_5~~~xcpT_2~~~xcpT_4	Type II secretion system protein G;hypothetical protein
86	epsF_5~~~epsF_4~~~epsF_3	Type II secretion system protein F
85	dnaK_2~~~dnaK_1~~~dnaK_3~~~dnaK_4	Chaperone protein DnaK
85	group_2615	hypothetical protein
85	group_10361	hypothetical protein
85	groS	10 kDa chaperonin
85	groL_2~~~groL_3	60 kDa chaperonin
85	cynT	Carbonic anhydrase 1
85	group_4330	hypothetical protein
85	group_2678	hypothetical protein
85	phrA_2~~~phrA	Deoxyribodipyrimidine photo-lyase
85	group_10900	hypothetical protein
85	group_10362	hypothetical protein
85	group_3335	hypothetical protein
85	group_3306	hypothetical protein
84	group_1833	hypothetical protein
84	group_1593	hypothetical protein
83	group_2105	hypothetical protein
83	group_2147	hypothetical protein
82	casD	CRISPR system Cascade subunit CasD;hypothetical protein
82	group_11499	hypothetical protein
82	casC	CRISPR system Cascade subunit CasC
82	ygbT~~~ygbF~~~cas1	CRISPR-associated endonuclease Cas1;CRISPR-associated endoribonuclease Cas2
82	group_10999	hypothetical protein
82	casA	hypothetical protein;CRISPR system Cascade subunit CasA
82	ygcB	CRISPR-associated endonuclease/helicase Cas3;hypothetical protein
81	group_9073	hypothetical protein

81	group_6894	hypothetical protein
80	group_2629	IS3 family transposase IS1416;IS3 family transposase IS401
80	group_2620	IS3 family transposase ISBt3;IS3 family transposase IS401
79	group_7003	hypothetical protein
79	group_9162	hypothetical protein
78	group_9644	hypothetical protein
78	group_10520	hypothetical protein
77	group_1839	hypothetical protein
77	group_2837	DNA-binding protein Bv3F
76	group_1836	hypothetical protein
76	group_3729	hypothetical protein
76	group_7310	hypothetical protein
75	group_2855	hypothetical protein
75	group_1076	hypothetical protein
74	group_10035	hypothetical protein
74	group_10498	hypothetical protein
74	fabI_1~~~fabI_2	Enoyl-[acyl-carrier-protein] reductase [NADH] FabI
74	ackA	Acetate kinase
74	pta	Phosphate acetyltransferase
74	group_3756	hypothetical protein
74	group_3156	hypothetical protein
74	group_9584	hypothetical protein
74	group_2115	hypothetical protein
74	group_11506	hypothetical protein
74	group_10199	hypothetical protein
74	group_9022	hypothetical protein
74	yecD~~~yecD_2~~~yecD_1	Isochorismatase family protein YecD
74	nopX	hypothetical protein;Nodulation outer protein X
74	group_1766	hypothetical protein
74	yycB~~~nimT	putative transporter YycB;hypothetical protein;2-nitroimidazole transporter
74	group_9765	hypothetical protein
74	group_7240	hypothetical protein
74	group_7239	hypothetical protein
74	group_2706	hypothetical protein
74	cmpR_1~~~cmpR_3~~~cmpR_2~~~hdfR_1	HTH-type transcriptional activator CmpR;hypothetical protein;HTH-type transcriptional regulator HdfR
74	group_7567	Outer membrane protein TolC

74	group_691	hypothetical protein
74	group_290	hypothetical protein
74	group_262	hypothetical protein
74	group_81	putative glycosyltransferase
74	group_11917	hypothetical protein
74	sfnC	putative FMNH2-dependent monooxygenase SfnC;hypothetical protein
74	group_11049	hypothetical protein
74	group_11005	hypothetical protein
74	argE_2~~~argE_3	Acetylornithine deacetylase
74	group_10386	queuosine precursor transporter
74	qseB_2~~~qseB_3	Transcriptional regulatory protein QseB
74	phaB_1~~~phaB_2	Acetoacetyl-CoA reductase
74	group_9923	hypothetical protein
74	davD_2~~~davD_1	Glutarate-semialdehyde dehydrogenase
74	sasA_3~~~sasA_16~~~sasA_5~~~sasA_12	Adaptive-response sensory-kinase SasA
74	group_9301	hypothetical protein
74	iolX	scyllo-inositol 2-dehydrogenase (NAD(+))
74	phzS	5-methylphenazine-1-carboxylate 1-monooxygenase
74	phnM	Alpha-D-ribose 1-methylphosphonate 5-triphosphate diphosphatase
74	group_5101	hypothetical protein
74	oxyR_2	Hydrogen peroxide-inducible genes activator;hypothetical protein
74	group_2196	hypothetical protein
74	group_2100	hypothetical protein
74	sacX_2~~~sacX_1~~~sacX	Negative regulator of SacY activity;hypothetical protein
74	scrB	hypothetical protein;Sucrose-6-phosphate hydrolase
74	lamB	Maltoporin;hypothetical protein
74	group_1051	hypothetical protein
74	scrK	Fructokinase
74	benC	Benzoate 12-dioxygenase electron transfer component;hypothetical protein
74	baiA~~~fabG2~~~ucpA_2~~~ucpA_1~~~baiC_1	hypothetical protein;3-alpha-hydroxycholanate dehydrogenase

		(NADP(+));putative oxidoreductase;Oxidoreductase UcpA;Dihydroanticapsin 7-dehydrogenase 3-alpha-hydroxycholanate dehydrogenase (NADP(+));3-oxoacyl-[acyl-carrier-protein] reductase FabG;1-deoxy-11-beta-hydroxypentalenate dehydrogenase;hypothetical protein;3-phenylpropionate-dihydrodiol/cinnamic acid-dihydrodiol dehydrogenase
74	baiA~~~fabG_5~~~ptlF~~~hcaB_1~~~baiA_2	
74	araG_2~~~araG	hypothetical protein;Arabinose import ATP-binding protein AraG
74	cusB_3~~~cusB~~~cusB_2	Cation efflux system protein CusB
74	xerC_4~~~xerC_7~~~xerC_1~~~xerC_5	Tyrosine recombinase XerC;hypothetical protein
74	group_2201	hypothetical protein
74	ccpA_3~~~purR_1~~~ccpA_1	hypothetical protein;IS5 family transposase IS1421;Catabolite control protein A;HTH-type transcriptional repressor PurR
74	group_1412	hypothetical protein
74	group_576	hypothetical protein
74	group_533	hypothetical protein
74	group_487	hypothetical protein
74	group_435	hypothetical protein
74	czcA_6~~~czcA_4	Cobalt-zinc-cadmium resistance protein CzcA
74	group_219	hypothetical protein
74	bamA_2~~~bamA_1	hypothetical protein;Outer membrane protein assembly factor BamA
74	group_65	hypothetical protein
74	group_11979	Ribose import ATP-binding protein RbsA
74	iolE	Inososе dehydratase
74	thpA	D-threitol-binding protein
74	iolB	5-deoxy-glucuronate isomerase
74	group_11405	hypothetical protein
74	group_11253	hypothetical protein
74	group_10866	hypothetical protein
74	idhA~~~iolG_2	Inositol 2-dehydrogenase;Myo-inositol 2-dehydrogenase

74	doxA	Naphthalene 12-dioxygenase system ferredoxin component
74	group_9984	hypothetical protein
74	cbdA	2-halobenzoate 12-dioxygenase large subunit
74	murR_2~~~murR_1	HTH-type transcriptional regulator MurR
74	group_8864	hypothetical protein
74	catA	Catechol 12-dioxygenase
74	cbdB	2-halobenzoate 12-dioxygenase small subunit
74	andAa	Anthranilate 12-dioxygenase system ferredoxin--NAD(+) reductase component
74	catB	Muconate cycloisomerase 1
74	catC	Muconolactone Delta-isomerase
74	benM~~~benM_1	HTH-type transcriptional regulator BenM
74	argC_3~~~argC_2~~~argC_1	N-acetyl-gamma-glutamyl-phosphate reductase
74	group_1972	hypothetical protein
74	group_934	hypothetical protein
74	iolC	5-dehydro-2-deoxygluconokinase
74	rbsC_1~~~rbsC_2	hypothetical protein;Ribose import permease protein RbsC
74	iolD	3D-(35/4)-trihydroxycyclohexane-12-dione hydrolase
74	group_8239	hypothetical protein
74	group_2619	hypothetical protein
74	group_2583	hypothetical protein
74	group_2082	hypothetical protein
74	glnH_2~~~mltF_2~~~glnH_3	ABC transporter glutamine-binding protein GlnH;Membrane-bound lytic murein transglycosylase F;hypothetical protein
74	glnM_2	hypothetical protein;putative glutamine ABC transporter permease protein GlnM
74	occM	Octopine transport system permease protein OccM
74	btuD_9~~~glnQ_5~~~tcyC_2~~~tcyC_1	Vitamin B12 import ATP-binding protein BtuD;Glutamine transport ATP-binding protein GlnQ;hypothetical protein;L-cystine import ATP-binding protein TcyC

74	crp~~~queE_2~~~queE_1	CRP-like cAMP-activated global transcriptional regulator;7-carboxy-7-deazaguanine synthase
74	group_3624	hypothetical protein
74	nrdD	hypothetical protein;Anaerobic ribonucleoside-triphosphate reductase
74	narT~~~nasA_2~~~nasA_1	hypothetical protein;putative nitrate transporter NarT;Nitrate transporter
74	narG	hypothetical protein;Respiratory nitrate reductase 1 alpha chain
74	degU	Transcriptional regulatory protein DegU
74	yihG	putative acyltransferase YihG
74	group_3126	hypothetical protein
74	argP_2~~~pgrR_3~~~pgrR_4~~~pgrR_2~~~pgrR_1~~~pgrR_5	HTH-type transcriptional regulator ArgP;HTH-type transcriptional regulator PgrR
74	group_11160	hypothetical protein
74	yhbU	putative protease YhbU
74	group_3650	hypothetical protein
74	group_3642	hypothetical protein
74	group_3640	hypothetical protein
74	narK	Nitrate/nitrite transporter NarK
74	group_3325	hypothetical protein;putative oxidoreductase
74	group_3119	hypothetical protein
74	group_529	putative protein;hypothetical protein
74	group_503	putative protein;hypothetical protein
74	group_467	hypothetical protein
74	puuC~~~puuC_1~~~puuC_2	NADP/NAD-dependent aldehyde dehydrogenase PuuC;hypothetical protein
74	group_10101	hypothetical protein
74	abo_1~~~abo_2	4-methylaminobutanoate oxidase (formaldehyde-forming)
74	hcnB_2~~~hcnB	Hydrogen cyanide synthase subunit HcnB
74	ghrA_2~~~ghrA_1	Glyoxylate/hydroxypyruvate reductase A
74	group_6891	hypothetical protein
74	spuC_2~~~spuC_1~~~spuC	Putrescine--pyruvate aminotransferase

74	lrp_5	Leucine-responsive regulatory protein
74	narH	Respiratory nitrate reductase 1 beta chain
74	group_3544	hypothetical protein
74	narI	Respiratory nitrate reductase 1 gamma chain
74	mobB_2	Molybdopterin-guanine dinucleotide biosynthesis adapter protein
74	narX	Nitrate/nitrite sensor protein NarX
74	group_953	hypothetical protein
74	group_623	hypothetical protein
74	hipO_3~~~hipO_1	hypothetical protein;Hippurate hydrolase
74	group_565	hypothetical protein
74	group_228	hypothetical protein
74	group_10541	hypothetical protein
74	glpE_3~~~glpE_2~~~glpE_1	Thiosulfate sulfurtransferase GlpE
74	metC_2~~~metC_3~~~metC_1	Cystathionine beta-lyase
74	group_3028	hypothetical protein
74	sgrR	hypothetical protein;HTH-type transcriptional regulator SgrR
74	puuB_2~~~puuB_3	Gamma-glutamylputrescine oxidoreductase
74	group_11465	hypothetical protein
74	proP_3~~~proP_1	Proline/betaine transporter;hypothetical protein
74	group_1961	hypothetical protein
74	gsiA_2	Glutathione import ATP-binding protein GsiA
74	group_251	hypothetical protein
74	gsiC_3~~~gsiC_2	Glutathione transport system permease protein GsiC
74	infA_2~~~infA_3	Translation initiation factor IF-1
74	aspC_3~~~aspC_2	Aspartate aminotransferase;hypothetical protein
74	bacC_2~~~gdh	Dihydroanticapsin 7-dehydrogenase;Galactitol 2-dehydrogenase
74	cefD_2~~~cefD	Isopenicillin N epimerase
74	group_3328	hypothetical protein
74	TDO2_1~~~TDO2_2~~~TDO2	Tryptophan 23-dioxygenase
74	group_3105	hypothetical protein

74	group_183	hypothetical protein
74	group_2701	hypothetical protein
74	glpR_1~~~glcR	Glycerol-3-phosphate regulon repressor;hypothetical protein;HTH-type transcriptional repressor GlcR
74	fabG1~~~fabG_8	3-oxoacyl-[acyl-carrier-protein] reductase FabG1;3-oxoacyl-[acyl-carrier-protein] reductase FabG
74	psd_2~~~psd_1	Phosphatidylserine decarboxylase proenzyme
74	group_602	hypothetical protein
74	cdhR~~~cdhR_2	HTH-type transcriptional regulator CdhR
74	group_11852	hypothetical protein
74	group_3159	hypothetical protein
74	group_10883	hypothetical protein
74	dmlR_17~~~dmlR_4~~~dmlR_6~~~dmlR_3~~~dmlR_9	HTH-type transcriptional regulator DmlR
74	jefA_2~~~stp~~~jefA	Drug efflux pump JefA;Multidrug resistance protein Stp
74	group_9649	hypothetical protein
74	group_7045	hypothetical protein
74	group_1446	hypothetical protein
74	group_597	hypothetical protein
74	group_2157	hypothetical protein
74	group_3535	hypothetical protein
74	tam_1~~~tam_2	Trans-aconitate 2-methyltransferase
74	group_3421	hypothetical protein
74	group_1624	hypothetical protein
74	group_1080	hypothetical protein
74	group_2780	hypothetical protein
74	group_1711	hypothetical protein
74	group_1668	hypothetical protein
74	group_1343	hypothetical protein
74	group_10522	hypothetical protein
74	amiD	N-acetylmuramoyl-L-alanine amidase AmiD;hypothetical protein
74	group_2848	hypothetical protein
74	cycA	D-serine/D-alanine/glycine transporter;hypothetical protein
74	group_2690	hypothetical protein
74	group_1399	hypothetical protein
74	group_10725	hypothetical protein
74	gcvA_2~~~gcvA_1~~~gcvA_4~~~gcvA_3	Glycine cleavage system transcriptional activator

74	group_9148	hypothetical protein
74	argO_2~~~argO_1	Arginine exporter protein ArgO
74	group_8039	Inner membrane protein YbiR
74	group_7417	hypothetical protein
74	yafQ	hypothetical protein;mRNA interferase toxin YafQ
74	dapA_1~~~dapA_5	4-hydroxy-tetrahydrodipicolinate synthase
74	adh1~~~adhB	Long-chain-alcohol dehydrogenase 1;Alcohol dehydrogenase 2
74	pdxA_2~~~pdxA_1	4-hydroxythreonine-4-phosphate dehydrogenase
74	putP_2~~~putP_1	High-affinity proline transporter PutP;hypothetical protein
74	group_2713	hypothetical protein
74	kimA	hypothetical protein;Potassium transporter KimA
74	group_2175	hypothetical protein
74	group_12038	hypothetical protein
74	group_11460	hypothetical protein
74	tsaR~~~argP_2~~~cynR_4	HTH-type transcriptional regulator TsaR;HTH-type transcriptional regulator ArgP;hypothetical protein;HTH-type transcriptional regulator CynR
74	group_10661	hypothetical protein
74	group_10185	hypothetical protein
74	group_10039	hypothetical protein
74	dapA_2~~~dapA_4~~~dapA_1	4-hydroxy-tetrahydrodipicolinate synthase
74	group_8327	hypothetical protein
74	cmpR_3~~~cmpR_2	hypothetical protein;HTH-type transcriptional activator CmpR
74	nox~~~drgA~~~rutE_2	NADH dehydrogenase;Protein DrgA;hypothetical protein;malonic semialdehyde reductase RutE
74	ilvE_2	Branched-chain-amino-acid aminotransferase
74	dmlR_12~~~dmlR_11	HTH-type transcriptional regulator DmlR;hypothetical protein
74	group_7070	hypothetical protein
74	group_6217	hypothetical protein
74	kynB_1~~~kynB_2	Kynurenine formamidase
74	sasA_14~~~sasA_13~~~sasA_12	Adaptive-response sensory-kinase SasA;hypothetical protein
74	group_9931	hypothetical protein

74	group_9273	hypothetical protein
74	group_9195	hypothetical protein
74	group_9025	hypothetical protein
74	group_7204	hypothetical protein
74	group_7199	hypothetical protein
74	group_7083	hypothetical protein
74	group_6779	hypothetical protein
74	group_6296	hypothetical protein
74	group_5907	hypothetical protein
74	group_3324	hypothetical protein
74	group_2146	hypothetical protein
74	group_2133	hypothetical protein
74	lodB	Putative FAD-dependent oxidoreductase LodB
	rhsB_1~~~rhsA_2~~~rhsB_2~~~rhsA_1~~~	putative deoxyribonuclease RhsB;putative deoxyribonuclease
74	rhsB~~~rhsA	RhsA;hypothetical protein
74	group_6334	hypothetical protein
74	group_1577	hypothetical protein
74	group_11656	hypothetical protein
74	group_8865	hypothetical protein
74	tar_2	Methyl-accepting chemotaxis protein II;hypothetical protein
74	group_6558	hypothetical protein
74	citB	Aconitate/2-methylaconitate hydratase
74	mdtC_4~~~mdtC_5~~~mdtC_2	Multidrug resistance protein MdtC
74	mdtA_7~~~mdtA_3~~~ttgA	Multidrug resistance protein MdtA;Toluene efflux pump periplasmic linker protein TtgA
74	ddl_2	D-alanine--D-alanine ligase
74	group_1772	hypothetical protein
74	group_1263	hypothetical protein
74	group_12041	hypothetical protein
74	group_12000	hypothetical protein
74	hmuV~~~btuD_5	Hemin import ATP-binding protein HmuV;hypothetical protein;Vitamin B12 import ATP-binding protein BtuD
74	prpC_2~~~prpC_1	2-methylcitrate synthase
74	group_9849	hypothetical protein
74	group_9196	hypothetical protein
74	group_7167	hypothetical protein
74	group_6709	2-amino-3-ketobutyrate coenzyme A ligase

74	group_6068	putative epimerase/dehydratase
74	sutR_2	HTH-type transcriptional regulator SutR
74	czcA_1~~~czcA_2~~~czcA_3	Cobalt-zinc-cadmium resistance protein CzcA
74	mdtA_4~~~mdtA_3~~~cusB_1	Multidrug resistance protein MdtA;Cation efflux system protein CusB
74	oprM_3~~~oprM_2	Outer membrane protein OprM
74	group_1290	hypothetical protein
74	group_1137	hypothetical protein
74	group_753	hypothetical protein
74	group_11849	hypothetical protein
74	adh_1~~~adhT~~~adh_3	Alcohol dehydrogenase;hypothetical protein
74	group_10671	hypothetical protein
74	dmlR_2~~~dmlR_4~~~dmlR_9~~~dmlR_8	HTH-type transcriptional regulator DmlR;hypothetical protein
74	baiA~~~baiA_1~~~baiA_3	3-alpha-hydroxycholanate dehydrogenase (NADP(+))
74	group_8119	hypothetical protein
74	group_8093	hypothetical protein
74	group_7608	hypothetical protein
74	group_5963	hypothetical protein
74	group_5926	hypothetical protein
74	adhC2	hypothetical protein;NADP-dependent alcohol dehydrogenase C 2
74	moeZ	putative adenylyltransferase/sulfurtransferase MoeZ
74	mec	CysO-cysteine peptidase
74	cysO_1~~~cysO	Sulfur carrier protein CysO
74	frc_2	Formyl-CoA:oxalate CoA-transferase
74	trsA	Triostin synthetase I
74	bbsG	(R)-benzylsuccinyl-CoA dehydrogenase
74	group_3587	hypothetical protein
74	group_3586	hypothetical protein
74	group_3513	hypothetical protein
74	btuB_4~~~btuB_2	Vitamin B12 transporter BtuB
74	group_2729	hypothetical protein
74	creB~~~rssB_2	Transcriptional regulatory protein CreB;Regulator of RpoS
74	group_1148	hypothetical protein

74	group_873	hypothetical protein
74	group_11995	hypothetical protein
74	mdtA_2~~~mdtA_7	hypothetical protein;Multidrug resistance protein MdtA
74	group_11610	hypothetical protein
74	group_11448	Cobalt-zinc-cadmium resistance protein CzcA
74	group_9467	hypothetical protein
74	cpdA_3~~~cpdA_2~~~cpdA_4~~~cpdA_1	3'5'-cyclic adenosine monophosphate phosphodiesterase CpdA
74	group_1746	hypothetical protein
74	group_1472	hypothetical protein
74	dmlR_9~~~dmlR_19~~~dmlR_14~~~dmlR_4	HTH-type transcriptional regulator DmlR
74	yhcG_1~~~yhcG~~~yhcG_2	Putative nuclease YhcG;hypothetical protein
74	ttuD_2~~~ttuD_1	Putative hydroxypyruvate reductase
74	nemA_2~~~nemA_1	N-ethylmaleimide reductase
74	group_8082	hypothetical protein
74	group_7545	hypothetical protein
74	group_7535	hypothetical protein
74	group_7161	hypothetical protein
74	azoR	FMN-dependent NADH-azoreductase
74	group_6287	hypothetical protein
74	group_6030	hypothetical protein
74	group_5401	hypothetical protein
74	group_3648	hypothetical protein
74	ycgJ	putative methyltransferase YcgJ;hypothetical protein
74	group_11955	hypothetical protein
74	group_10271	hypothetical protein
74	group_6525	hypothetical protein
74	group_11080	hypothetical protein
74	group_9925	hypothetical protein
74	group_9511	hypothetical protein
74	dhmA~~~oleB	Haloalkane dehalogenase;Cis-3-alkyl-4-alkyloxetan-2-one decarboxylase;hypothetical protein
74	yjiB	Putative cytochrome P450 YjiB;hypothetical protein
74	ptlI	Putative cytochrome P450 120;Pentalenene oxygenase;hypothetical protein

74	btuC~~~hmuU~~~btuC_2	Vitamin B12 import system permease protein BtuC;Hemin transport system permease protein HmuU
74	group_9763	hypothetical protein
74	pcpR_6~~~pcpR_5~~~pcpR_7	PCP degradation transcriptional activation protein
74	gltR_2	HTH-type transcriptional regulator GltR
74	group_6206	hypothetical protein
74	pksR~~~pksM~~~pksM_3~~~pksN~~~pksL_1~~~fadA_2~~~pksL~~~fadA_3	hypothetical protein;Polyketide synthase PksR;Polyketide synthase PksM;Polyketide synthase PksN;Polyketide synthase PksL;3-ketoacyl-CoA thiolase
74	acpP_4~~~acpP_3~~~acpP_1~~~pksL_1~~~acpP_2	Acyl carrier protein;hypothetical protein;Polyketide synthase PksL
74	group_513	hypothetical protein
74	group_222	hypothetical protein
74	fadI	3-ketoacyl-CoA thiolase FadI;3-ketoacyl-CoA thiolase
74	ku	hypothetical protein;Non-homologous end joining protein Ku
74	pksL~~~pksL_2~~~pksN~~~pksJ~~~pksR	Polyketide synthase PksL;Polyketide synthase PksN;Polyketide synthase PksJ;Polyketide synthase PksR
74	baeD	hypothetical protein;Polyketide biosynthesis acyltransferase BaeD
74	group_12009	hypothetical protein
74	higB	hypothetical protein;mRNA interferase HigB
74	chrA1_2~~~chrA1_1	Chromate transport protein;hypothetical protein
74	group_6819	hypothetical protein
74	group_5797	hypothetical protein
74	group_1823	hypothetical protein
74	group_10270	hypothetical protein
74	group_8166	hypothetical protein
74	group_7990	hypothetical protein
74	group_7925	hypothetical protein
74	ligD_2~~~ligD_1~~~ligD	Multifunctional non-homologous end joining protein LigD
74	group_3739	hypothetical protein
74	group_2178	hypothetical protein
74	group_1479	hypothetical protein
74	group_11857	hypothetical protein

74	group_10446	hypothetical protein
74	rtcB_1	hypothetical protein;RNA-splicing ligase RtcB
74	group_8397	Cardiolipin synthase;hypothetical protein;Cardiolipin synthase B
74	pqiB_2~~~pqiB_3	Intermembrane transport protein PqiB
74	group_7824	hypothetical protein
74	group_7561	hypothetical protein
74	group_5796	hypothetical protein
74	group_4421	hypothetical protein
74	group_3873	hypothetical protein
74	cmoM~~~cmoM_1	tRNA 5-carboxymethoxyuridine methyltransferase;hypothetical protein
74	COQ5_3~~~COQ5_2~~~COQ5_1	2-methoxy-6-polyprenyl-14-benzoquinol methylase mitochondrial
74	group_2937	hypothetical protein
74	group_2348	hypothetical protein
74	fliY_1~~~fliY_3	L-cystine-binding protein FliY
74	glnM_1~~~glnM_3~~~glnM_2	putative glutamine ABC transporter permease protein GlnM
74	metXA_2~~~metXA_1	Homoserine O-acetyltransferase
74	group_9914	hypothetical protein
74	group_9468	hypothetical protein
74	group_9444	hypothetical protein
74	group_8991	hypothetical protein
74	group_8028	hypothetical protein
74	pntB_2	NAD(P) transhydrogenase subunit beta
74	metR_2~~~argP_2	hypothetical protein;HTH-type transcriptional regulator MetR;HTH-type transcriptional regulator ArgP
74	group_5951	hypothetical protein
74	group_401	hypothetical protein
74	group_47	hypothetical protein
74	group_9244	hypothetical protein
74	group_6309	hypothetical protein
74	group_5891	hypothetical protein
74	group_1628	hypothetical protein
74	group_663	hypothetical protein
74	group_9858	hypothetical protein
74	group_9101	hypothetical protein
74	group_6587	hypothetical protein

74	rhaS_3~~~rhaS_2~~~rhaS_4	HTH-type transcriptional activator RhaS;hypothetical protein
74	group_9878	hypothetical protein
74	group_9528	hypothetical protein
74	group_8100	hypothetical protein
74	group_7789	hypothetical protein
74	group_6902	hypothetical protein
74	group_4957	hypothetical protein
74	group_2557	hypothetical protein;IS5 family transposase ISAzo11
74	licC	hypothetical protein;Lichenan permease IIC component
74	group_10697	hypothetical protein
74	group_3282	hypothetical protein
74	group_3003	hypothetical protein
74	sadH~~~hcaB_2	hypothetical protein;Putative oxidoreductase SadH;3-phenylpropionate-dihydrodiol/cinnamic acid-dihydrodiol dehydrogenase;Baeyer-Villiger monooxygenase
74	group_9830	hypothetical protein
74	group_8343	hypothetical protein
74	group_2835	hypothetical protein
74	sigJ	ECF RNA polymerase sigma factor SigJ
74	group_10972	hypothetical protein
74	group_10847	hypothetical protein
74	group_9416	hypothetical protein
74	group_9393	hypothetical protein
74	group_9387	hypothetical protein
74	group_8730	hypothetical protein
74	group_8167	hypothetical protein
74	group_7689	hypothetical protein
74	group_7268	hypothetical protein
74	group_7262	hypothetical protein
74	decR_4~~~lrp_4	DNA-binding transcriptional activator DecR;Leucine-responsive regulatory protein
74	dhaA~~~dhaA_1	Haloalkane dehalogenase
74	group_6423	hypothetical protein
74	group_6293	hypothetical protein;Leucine efflux protein
74	phnX	Phosphonoacetaldehyde hydrolase
74	group_5814	hypothetical protein

74	hdfR_3~~~hdfR_4	HTH-type transcriptional regulator HdfR
74	group_4494	hypothetical protein
74	group_4115	hypothetical protein
74	group_1727	hypothetical protein
74	group_10341	hypothetical protein
74	group_10299	hypothetical protein
74	group_8420	hypothetical protein
74	group_8189	hypothetical protein
74	group_7278	hypothetical protein;Putative thymidine phosphorylase
74	group_6271	hypothetical protein
74	group_5920	hypothetical protein
74	maoA~~~tynA	Primary amine oxidase
74	nicB	6-hydroxy-3-succinoylpyridine 3-monooxygenase HspA;hypothetical protein
74	group_4465	hypothetical protein
74	group_858	hypothetical protein
74	group_12043	hypothetical protein
74	group_12042	hypothetical protein
74	ligJ_2~~~ligJ_1~~~ligJ	2-keto-4-carboxy-3-hexenedioate hydratase
74	proA_2	4-hydroxy-4-methyl-2-oxoglutarate aldolase/4-carboxy-4-hydroxy-2-oxoadipate aldolase
74	group_11387	hypothetical protein
74	group_10327	hypothetical protein
74	yabJ_2~~~yabJ_4~~~yabJ_5~~~yabJ_1~~~yabJ_3	hypothetical protein;2-iminobutanoate/2-iminopropanoate deaminase
74	dksA_2~~~dksA_1	RNA polymerase-binding transcription factor DksA
74	group_8048	hypothetical protein
74	group_6377	hypothetical protein
74	group_6215	hypothetical protein
74	group_5799	hypothetical protein
74	group_5328	hypothetical protein
74	group_3534	hypothetical protein
74	group_1859	hypothetical protein
74	group_619	hypothetical protein
74	hdfR_4~~~hdfR_3	HTH-type transcriptional regulator HdfR
74	ligU	(4E)-oxalomesaconate Delta-isomerase

74	group_11449	hypothetical protein
74	aqpZ_2~~~aqpZ~~~aqpZ_1	Aquaporin Z
74	group_3121	hypothetical protein
74	group_2659	hypothetical protein
74	group_2604	hypothetical protein
74	group_1640	hypothetical protein
74	mhpB	23-dihydroxyphenylpropionate/23-dihydroxicinnamic acid 12-dioxygenase
74	group_10524	hypothetical protein
74	group_10409	hypothetical protein
74	nodD2_2	Nodulation protein D 2
74	group_8745	hypothetical protein
74	group_8354	hypothetical protein
74	group_8353	hypothetical protein
74	group_8277	hypothetical protein
74	group_7680	hypothetical protein
74	group_6968	hypothetical protein
74	group_6478	hypothetical protein
74	group_6477	hypothetical protein
74	group_6248	hypothetical protein
74	arsC_3~~~arsC_2~~~arsC	Arsenate reductase
74	cadI~~~cadI_1	Cadmium-induced protein CadI
74	group_3677	hypothetical protein
74	group_3125	hypothetical protein
74	group_2738	hypothetical protein
74	group_7	hypothetical protein;Putative antitoxin VapB45
74	group_11876	hypothetical protein
74	uppP_2	Undecaprenyl-diphosphatase
74	group_2853	hypothetical protein
74	group_905	hypothetical protein
74	group_9720	hypothetical protein
74	group_8094	hypothetical protein
74	group_7902	hypothetical protein
74	group_5986	hypothetical protein
74	group_5909	hypothetical protein
74	group_4273	hypothetical protein
74	group_3649	hypothetical protein
74	mdtA_1	Multidrug resistance protein MdtA
74	group_1301	hypothetical protein
74	group_1010	hypothetical protein
74	group_465	hypothetical protein
74	group_425	hypothetical protein

74	dadA_1~~~dadA_4~~~dadA_3	D-amino acid dehydrogenase
74	group_10834	hypothetical protein
74	group_10462	HTH-type transcriptional activator CmpR
74	group_10370	hypothetical protein
74	plcB~~~plcC	Phospholipase C 2;hypothetical protein;Phospholipase C 3
74	group_9847	hypothetical protein
74	group_9567	hypothetical protein
74	group_9392	hypothetical protein
74	group_8812	hypothetical protein
74	desVI	dTDP-3-amino-346-trideoxy-alpha-D-glucopyranose
74	group_7781	hypothetical protein
74	group_7528	hypothetical protein
74	group_7480	hypothetical protein
74	group_7364	hypothetical protein
74	group_7269	hypothetical protein
74	group_7236	hypothetical protein
74	group_7234	hypothetical protein
74	group_7214	hypothetical protein
74	group_6882	hypothetical protein
74	group_6562	hypothetical protein;IS5 family transposase IS1421
74	group_6507	hypothetical protein
74	group_6506	hypothetical protein
74	group_5989	hypothetical protein
74	group_4707	hypothetical protein
74	group_4077	hypothetical protein
74	group_3894	hypothetical protein
74	group_3479	hypothetical protein
74	cdiA	hypothetical protein;Toxin CdiA
74	group_2640	IS3 family transposase ISRso10;IS3 family transposase ISButh1
74	group_2636	hypothetical protein
74	group_2035	hypothetical protein
74	inhA_2	Isonitrile hydratase
74	group_937	hypothetical protein
74	dadA1_2~~~dadA_3~~~thiO	D-amino acid dehydrogenase 1;D-amino acid dehydrogenase;Glycine oxidase
74	yabJ_5~~~yabJ_4	2-iminobutanoate/2-iminopropanoate deaminase;hypothetical protein

74	aam	Acylamidase
74	proP_4~~~ousA~~~proP_5	Proline/betaine transporter;Glycine betaine/proline/ectoine/pipecolic acid transporter OusA
74	group_562	hypothetical protein
74	group_470	hypothetical protein
74	group_6853	hypothetical protein
74	group_6135	hypothetical protein
74	arsC_2~~~arsC_1	Arsenate reductase
74	group_3498	hypothetical protein
74	group_3483	hypothetical protein
74	group_3072	hypothetical protein
74	group_1962	hypothetical protein
74	group_1888	hypothetical protein
74	group_1723	hypothetical protein
74	group_1387	hypothetical protein
74	group_1103	hypothetical protein
74	group_255	hypothetical protein
74	group_11323	hypothetical protein
74	group_10832	hypothetical protein
74	group_10194	hypothetical protein
74	group_10093	hypothetical protein
74	group_9831	hypothetical protein
74	group_9429	hypothetical protein
74	group_9183	Putative universal stress protein
74	group_7991	hypothetical protein
74	dmlA~~~dmlA_1~~~dmlA_3	D-malate dehydrogenase [decarboxylating]
74	group_6825	hypothetical protein
74	group_6824	hypothetical protein
74	group_6802	hypothetical protein
74	group_5935	hypothetical protein
74	group_4444	hypothetical protein
74	group_4344	hypothetical protein
74	aioA	Arsenite oxidase subunit AioA
74	aioB	Arsenite oxidase subunit AioB
74	xerC_7~~~xerC_1~~~xerD_2~~~xerC_5	Tyrosine recombinase XerC;Tyrosine recombinase XerD
74	group_2176	hypothetical protein
74	group_2137	hypothetical protein
74	group_10971	hypothetical protein
74	group_10585	hypothetical protein
74	group_10340	hypothetical protein
74	group_10005	hypothetical protein

74	group_7746	hypothetical protein
74	group_6399	hypothetical protein
74	group_5956	hypothetical protein
74	group_3371	hypothetical protein
74	group_2821	hypothetical protein
74	group_2032	hypothetical protein
74	group_1509	hypothetical protein
74	group_10902	hypothetical protein
74	group_10159	hypothetical protein
74	group_9775	hypothetical protein
74	group_9730	hypothetical protein
74	group_9307	hypothetical protein
74	group_7683	hypothetical protein
74	group_7672	hypothetical protein
74	group_7408	hypothetical protein
74	group_7267	hypothetical protein
74	group_7228	hypothetical protein
74	group_7227	hypothetical protein
74	group_6781	hypothetical protein
74	group_5127	hypothetical protein
74	group_4948	hypothetical protein
74	group_3855	hypothetical protein
74	group_3647	hypothetical protein
74	group_3300	hypothetical protein
74	group_3292	hypothetical protein
74	group_3234	hypothetical protein
74	group_2746	hypothetical protein
74	group_1994	hypothetical protein
74	group_1481	hypothetical protein
74	group_235	hypothetical protein
74	group_67	hypothetical protein
74	group_45	hypothetical protein
74	group_11261	Putative phosphoribosyl transferase
74	group_10726	hypothetical protein
74	group_10090	hypothetical protein
74	sbnD_1	Staphyloferrin B transporter
74	group_9313	hypothetical protein
74	group_8935	hypothetical protein
74	group_7595	hypothetical protein
74	group_7560	hypothetical protein
74	group_7231	hypothetical protein
74	group_7151	hypothetical protein
74	group_7054	hypothetical protein

74	oprM_7~~~oprM_1	Outer membrane protein OprM
74	group_6164	hypothetical protein
74	group_5850	hypothetical protein
74	ytrE_1	ABC transporter ATP-binding protein YtrE
74	fabD_2	hypothetical protein;Malonyl CoA-acyl carrier protein transacylase
74	madD	Malonyl-S-ACP:biotin-protein carboxyltransferase MADD;hypothetical protein
74	accD_2	hypothetical protein;Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta chloroplastic
74	group_2924	hypothetical protein
74	group_2906	hypothetical protein
74	mdcB~~~citG	2-(5''-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase
74	group_2693	hypothetical protein;Universal stress protein
74	group_2566	hypothetical protein
74	group_1650	hypothetical protein
74	group_11272	hypothetical protein
74	group_8150	hypothetical protein
74	group_8145	hypothetical protein
74	group_6196	hypothetical protein
74	group_5019	hypothetical protein
74	nepI~~~nepI_2	Purine ribonucleoside efflux pump NepI
74	group_3514	hypothetical protein
74	group_3216	hypothetical protein
74	kdpB_1	Potassium-transporting ATPase ATP-binding subunit
74	group_2915	hypothetical protein
74	group_2774	hypothetical protein
74	group_2488	hypothetical protein
74	fatA_1	Ferric-anguibactin receptor FatA
74	group_2312	hypothetical protein
74	group_2173	hypothetical protein
74	group_1971	hypothetical protein
74	group_603	hypothetical protein
74	group_10956	hypothetical protein
74	group_10307	hypothetical protein
74	mltF_2~~~mltF_1	Membrane-bound lytic murein transglycosylase F

74	rarA_2	Replication-associated recombination protein A
74	rspR_1	hypothetical protein;HTH-type transcriptional repressor RspR
74	group_9391	hypothetical protein
74	group_9156	hypothetical protein
74	group_9087	hypothetical protein
74	group_9086	hypothetical protein
74	group_8644	hypothetical protein
74	yecD_2~~~yecD_1	Isochorismatase family protein YecD
74	pks5	Mycocerosic acid synthase-like polyketide synthase;hypothetical protein
74	group_6887	hypothetical protein
74	group_6886	hypothetical protein
74	group_6718	hypothetical protein
74	group_6535	hypothetical protein
74	group_6499	hypothetical protein
74	group_6411	hypothetical protein
74	group_5703	hypothetical protein
74	group_3290	hypothetical protein
74	group_1657	hypothetical protein
74	group_1229	hypothetical protein
74	group_1225	hypothetical protein
74	group_8362	hypothetical protein
74	group_7673	hypothetical protein
74	group_7107	hypothetical protein
74	group_6828	hypothetical protein
74	group_6070	hypothetical protein
74	group_5877	hypothetical protein
74	group_5469	hypothetical protein
74	group_5071	hypothetical protein
74	group_4945	hypothetical protein
74	pinR_2~~~pinR	Serine recombinase PinR
74	group_3115	hypothetical protein
74	group_2528	hypothetical protein
74	group_1334	hypothetical protein
74	group_11427	hypothetical protein
74	group_11424	hypothetical protein
74	hupB_3~~~hup	DNA-binding protein HU-beta;DNA-binding protein HU
74	group_11307	hypothetical protein
74	group_10106	IS5 family transposase ISAzo23
74	group_9311	hypothetical protein

74	group_9105	hypothetical protein
74	group_7690	hypothetical protein
74	group_7309	hypothetical protein
74	group_7308	hypothetical protein
74	group_6904	hypothetical protein
74	group_6903	hypothetical protein
74	group_6832	hypothetical protein
74	group_6004	hypothetical protein
74	cynR_4~~~cynR_3	HTH-type transcriptional regulator CynR;hypothetical protein
74	group_5835	hypothetical protein
74	group_4624	hypothetical protein
74	group_4232	hypothetical protein
74	group_3486	hypothetical protein
74	group_3334	hypothetical protein
74	group_3333	hypothetical protein
74	group_2955	hypothetical protein
74	group_2777	hypothetical protein
74	group_2092	hypothetical protein
74	group_2052	hypothetical protein
74	group_2044	hypothetical protein
74	group_1965	hypothetical protein
74	group_1908	hypothetical protein
74	group_1661	hypothetical protein
74	group_1660	hypothetical protein
74	group_1618	hypothetical protein
74	group_1587	hypothetical protein
74	group_1234	hypothetical protein
74	group_1159	hypothetical protein
74	group_1122	hypothetical protein
74	cbbZC_1~~~gph_1~~~cbbZC_2	Phosphoglycolate phosphatase chromosomal;Phosphoglycolate phosphatase;hypothetical protein
74	group_755	hypothetical protein
74	group_11478	IS3 family transposase ISRso20
74	group_10109	IS5 family transposase ISAzo23
74	group_9729	hypothetical protein
74	group_7614	IS5 family transposase ISAzo23
74	group_904	hypothetical protein
74	group_10091	IS5 family transposase IS1021
74	group_10080	IS5 family transposase IS1021
74	group_10079	IS5 family transposase IS1021
74	group_9844	hypothetical protein

74	group_6748	hypothetical protein;IS3 family transposase ISRso20
74	group_6198	hypothetical protein
74	group_5453	hypothetical protein
74	group_3174	hypothetical protein
74	group_1533	IS5 family transposase IS1021
74	mbtH~~~mbtH_1	Protein MbtH
74	mdtK_1~~~mdtK_2	Multidrug resistance protein MdtK
74	group_10108	IS5 family transposase ISAzo23
74	group_10076	IS5 family transposase IS1021
74	group_9310	hypothetical protein
74	egl_2~~~egl	Endoglucanase
74	group_4301	hypothetical protein
74	group_2925	IS3 family transposase ISRso10
73	group_1607	hypothetical protein
73	group_1096	hypothetical protein
72	group_617	hypothetical protein
72	group_606	hypothetical protein
71	group_3474	hypothetical protein
71	group_2849	hypothetical protein
71	group_2700	hypothetical protein
71	group_2598	hypothetical protein
71	group_2597	hypothetical protein;Type IV secretion system protein virB10
71	group_2517	hypothetical protein
71	group_12031	hypothetical protein
71	group_12028	hypothetical protein
71	traG	Conjugal transfer protein TraG;hypothetical protein
71	group_3601	hypothetical protein
71	group_3600	hypothetical protein
71	virB4~~~virB4_2~~~virB4_1	Type IV secretion system protein virB4
71	hcaR~~~hdfR_3~~~hdfR_2~~~benM_1~~~hdfR_1~~~benM_3	Hca operon transcriptional activator HcaR;HTH-type transcriptional regulator HdfR;HTH-type transcriptional regulator BenM
71	group_2331	hypothetical protein
70	group_1656	hypothetical protein
70	group_1654	hypothetical protein
70	intA_2	Prophage integrase IntA
70	group_2382	hypothetical protein
70	group_1416	hypothetical protein

70	group_7362	hypothetical protein;Actin cross-linking toxin VgrG1
70	group_7004	hypothetical protein
70	group_7248	hypothetical protein
70	group_1345	hypothetical protein
70	group_1673	hypothetical protein
70	group_7709	hypothetical protein
70	group_2916	hypothetical protein
70	group_7249	hypothetical protein
70	group_7197	hypothetical protein
70	group_6854	hypothetical protein
70	group_4314	hypothetical protein
70	group_4117	hypothetical protein
70	group_342	hypothetical protein
69	group_9325	hypothetical protein
69	group_7294	hypothetical protein
68	group_7628	hypothetical protein
68	group_7457	Cyanate hydratase
68	triA_1	Melamine deaminase
68	group_3819	hypothetical protein
68	hisC_4	Histidinol-phosphate aminotransferase
67	group_5967	hypothetical protein
67	group_10768	hypothetical protein
67	group_7849	Tyrosine recombinase XerC
66	group_1560	hypothetical protein
66	group_1134	hypothetical protein
65	group_11259	hypothetical protein
65	intS~~~intS_2~~~intS_1~~~intA_3~~~intA	Prophage integrase IntS;Prophage integrase IntA
65	group_10591	hypothetical protein
65	group_1629	hypothetical protein
64	group_1281	hypothetical protein
64	shlB_1~~~shlB_7~~~shlB_3	Hemolysin transporter protein ShlB;hypothetical protein
63	group_9875	HTH-type transcriptional regulator DmlR
63	novR_1~~~novR_2	Decarboxylase NovR
63	nicT_2	Putative metabolite transport protein NicT
62	group_1883	hypothetical protein
62	group_1881	hypothetical protein
62	group_1719	hypothetical protein
62	group_1389	hypothetical protein

62	group_1360	hypothetical protein
62	group_10417	hypothetical protein
62	group_2089	hypothetical protein
62	group_2060	hypothetical protein
62	group_2023	hypothetical protein
62	group_2022	hypothetical protein
62	group_1979	hypothetical protein
62	group_1977	hypothetical protein
62	group_10864	hypothetical protein
62	group_1970	hypothetical protein
62	group_1969	hypothetical protein
62	group_1866	hypothetical protein
62	group_1694	hypothetical protein
62	group_1323	hypothetical protein
62	group_9879	hypothetical protein
62	group_9802	hypothetical protein
62	group_9353	hypothetical protein
62	group_9176	hypothetical protein
62	group_9038	hypothetical protein
62	group_9033	hypothetical protein
62	group_8345	hypothetical protein
62	group_7850	hypothetical protein
62	group_7387	hypothetical protein
62	group_7257	hypothetical protein
62	group_7028	hypothetical protein
62	group_6803	hypothetical protein
62	group_6789	hypothetical protein
62	group_6241	hypothetical protein
62	group_5440	hypothetical protein
62	group_3981	hypothetical protein
62	group_2572	hypothetical protein
62	group_9298	hypothetical protein
61	group_12049	hypothetical protein
61	group_7610	hypothetical protein
61	group_3670	hypothetical protein
60	group_1769	hypothetical protein
60	group_1635	hypothetical protein
59	group_2283	hypothetical protein
59	group_2355	hypothetical protein
58	group_3476	hypothetical protein
58	group_1293	hypothetical protein
58	group_801	hypothetical protein
58	group_2451	hypothetical protein

58	group_2851	hypothetical protein
58	group_806	hypothetical protein
57	group_3015	hypothetical protein;IS3 family transposase ISAisp2
57	group_1615	IS3 family transposase ISAisp2
56	group_11761	hypothetical protein
56	group_11431	hypothetical protein
56	group_11426	hypothetical protein
56	group_10904	hypothetical protein
55	prsA_2	putative parvulin-type peptidyl-prolyl cis-trans isomerase;hypothetical protein;Foldase protein PrsA
55	group_1315	hypothetical protein
54	group_2857	hypothetical protein
54	group_1151	hypothetical protein
54	group_7463	hypothetical protein
53	group_2291	hypothetical protein
53	group_2958	hypothetical protein
53	group_2957	hypothetical protein
53	group_2304	hypothetical protein
52	group_3426	hypothetical protein
52	group_3422	hypothetical protein
51	group_2895	hypothetical protein
51	group_10066	hypothetical protein
50	group_10001	hypothetical protein;All-trans-zeta-carotene desaturase
50	group_9630	hypothetical protein
50	group_8922	Isatin hydrolase;hypothetical protein
50	styA	Styrene monooxygenase StyA;hypothetical protein
50	rhaS_1~~~pchR~~~feaR	hypothetical protein;HTH-type transcriptional activator RhaS;Regulatory protein PchR;Transcriptional activator FeaR
50	ntaB	FMN reductase (NADH) NtaB
50	group_9629	hypothetical protein
50	group_8088	hypothetical protein
50	yisK	putative protein YisK
50	maiA_1~~~maiA_2~~~maiA	Maleate isomerase
50	sdgC~~~namA_1~~~namA_2	Salicyloyl-CoA 5-hydroxylase;NADPH dehydrogenase
50	group_6637	hypothetical protein
50	group_6636	hypothetical protein

50	nitA	Aliphatic nitrilase;hypothetical protein
50	group_5880	14-dihydroxy-2-naphthoyl-CoA hydrolase
50	sdgD_2~~~sdgD_1	Gentisate 12-dioxygenase
49	group_10861	hypothetical protein
49	group_2839	hypothetical protein
48	group_10420	hypothetical protein
48	group_10400	IS3 family transposase ISAisp2
47	group_3445	hypothetical protein
47	group_1287	hypothetical protein
46	group_2979	hypothetical protein
46	group_9524	Putative defective protein IntQ
45	group_9243	hypothetical protein
45	group_9242	hypothetical protein
45	group_8917	hypothetical protein
45	group_9999	hypothetical protein
44	group_1691	hypothetical protein
44	esiB	hypothetical protein;Secretory immunoglobulin A-binding protein EsiB
43	group_11495	hypothetical protein
43	gpFl~~~gpFl_2~~~gpFl_1~~~gpFl_3	Putative prophage major tail sheath protein
43	group_1487	hypothetical protein
43	group_123	hypothetical protein
42	group_2437	Putative transposase InsK for insertion sequence element IS150;IS3 family transposase ISAisp2
42	group_2392	hypothetical protein;IS3 family transposase ISAisp2
41	xerC_5~~~xerC_3~~~xerC_7	hypothetical protein;Tyrosine recombinase XerC
41	group_5769	Tn3 family transposase ISPa43;hypothetical protein
40	group_3764	hypothetical protein
40	group_3767	hypothetical protein
39	dltA_2~~~dltA_1~~~dltA_3	hypothetical protein;D-alanine--D-alanyl carrier protein ligase
39	group_2425	hypothetical protein
39	rhaR_1~~~rhaR_3	HTH-type transcriptional activator RhaR
39	acpP_1~~~acpP_2~~~ppsE~~~ppsE_1	Acyl carrier protein;hypothetical protein;Phthiocerol synthesis polyketide synthase type I PpsE

39	irtA	putative ABC transporter ATP-binding protein;Iron import ATP-binding/permease protein IrtA
39	btuD_3~~~btuD_1~~~btuD_5	Vitamin B12 import ATP-binding protein BtuD
39	fyuA	Pesticin receptor;hypothetical protein
39	acs~~~eryA~~~dltA_1~~~dltA_2~~~dltA_4~~~menE_1~~~acs_2~~~menE_3	Acetyl-coenzyme A synthetase;4-hydroxyphenylalkanoate adenylyltransferase;6-deoxyerythronolide-B synthase EryA2 modules 3 and 4;D-alanine--D-alanyl carrier protein ligase;Phenolphthiocerol synthesis polyketide synthase type I Pks15/1;hypothetical protein;2-succinylbenzoate--CoA ligase
39	group_3012	hypothetical protein
39	COQ5_2~~~sfp~~~COQ5_1	hypothetical protein;2-methoxy-6-polyprenyl-14-benzoquinol methylase mitochondrial;4'-phosphopantetheinyl transferase Sfp
39	group_11370	N-acetylmuramoyl-L-alanine amidase AmiD;hypothetical protein
38	noc_1~~~noc_4~~~noc_2	hypothetical protein;Nucleoid occlusion protein
38	group_3742	hypothetical protein
37	group_3734	hypothetical protein
37	group_2704	hypothetical protein
37	group_2650	hypothetical protein
37	group_2761	hypothetical protein
37	group_1626	hypothetical protein
36	group_11402	hypothetical protein
36	group_1818	hypothetical protein
36	group_11741	hypothetical protein
35	group_55	hypothetical protein
35	group_22	hypothetical protein
34	group_2827	hypothetical protein
34	group_2830	hypothetical protein
34	group_4835	hypothetical protein
33	group_3230	hypothetical protein
33	group_132	hypothetical protein
33	group_2387	hypothetical protein
33	group_599	IS630 family transposase ISCARN39

33	group_11985	hypothetical protein
33	group_11756	hypothetical protein
33	group_11422	hypothetical protein
33	group_407	hypothetical protein
33	group_10389	hypothetical protein
33	group_4318	hypothetical protein
33	group_3457	hypothetical protein
33	group_10388	D-serine/D-alanine/glycine transporter
33	group_8091	hypothetical protein
33	group_6126	hypothetical protein
33	group_10313	hypothetical protein
32	group_3086	hypothetical protein
32	group_4630	hypothetical protein
32	group_1520	hypothetical protein
32	group_1518	hypothetical protein
32	group_4196	hypothetical protein
31	group_2930	hypothetical protein
31	group_12015	hypothetical protein
31	group_11813	hypothetical protein
31	group_11493	hypothetical protein
31	group_10992	hypothetical protein
31	group_10153	hypothetical protein
31	group_10993	hypothetical protein
31	group_10155	hypothetical protein
31	group_405	hypothetical protein
31	group_2568	hypothetical protein
31	gpFI_1~~~gpFI~~~gpFI_2	Putative prophage major tail sheath protein
30	group_2671	IS3 family transposase ISAtu5;IS3 family transposase ISRso14;hypothetical protein
30	group_7602	IS3 family transposase ISBam2;IS3 family transposase IS407
29	group_1255	hypothetical protein
29	group_10213	hypothetical protein
28	group_10892	hypothetical protein
28	group_10006	hypothetical protein
28	group_1829	hypothetical protein
28	group_9342	hypothetical protein
28	group_6627	hypothetical protein
28	group_1350	hypothetical protein
27	group_152	hypothetical protein
27	group_313	hypothetical protein

26	intA_1~~~intA~~~intA_2	Prophage integrase IntA
26	group_2653	hypothetical protein
26	group_1647	hypothetical protein
26	group_1645	hypothetical protein
25	hin_1~~~hin~~~hin_2	DNA-invertase hin;hypothetical protein
25	group_11821	hypothetical protein
25	group_1447	hypothetical protein
25	group_10227	hypothetical protein
25	group_9261	hypothetical protein
25	group_7173	hypothetical protein
25	group_7172	hypothetical protein
25	group_2264	hypothetical protein
25	group_1243	hypothetical protein
24	group_296	hypothetical protein
24	group_11637	hypothetical protein
23	group_9544	hypothetical protein
23	group_10428	hypothetical protein
22	group_3409	IS3 family transposase IS401;hypothetical protein
22	group_2398	IS3 family transposase IS401;Putative transposase InsK for insertion sequence element IS150
21	group_3215	hypothetical protein
21	noc_2~~~noc_4~~~noc_3	Nucleoid occlusion protein;hypothetical protein
21	noc_1~~~noc_2~~~noc_3	Nucleoid occlusion protein;hypothetical protein
20	group_10995	hypothetical protein
20	group_1271	hypothetical protein
20	group_2964	hypothetical protein
20	group_1086	hypothetical protein
20	group_2963	hypothetical protein
19	group_5966	hypothetical protein
19	group_3397	hypothetical protein
19	group_2480	hypothetical protein
18	group_1009	IS66 family transposase IS1313
18	group_1054	IS66 family transposase ISAeh1;IS66 family transposase ISPa82;hypothetical protein
18	group_800	hypothetical protein
17	group_1762	hypothetical protein
17	group_1526	hypothetical protein
17	group_852	hypothetical protein

17	clsB_3	Cardiolipin synthase B;Cardiolipin synthase;hypothetical protein
17	group_845	hypothetical protein
17	group_2862	hypothetical protein
17	group_1672	hypothetical protein
17	group_1643	hypothetical protein
17	group_1525	hypothetical protein
17	group_1270	hypothetical protein
17	group_9198	hypothetical protein
17	group_10795	hypothetical protein
17	group_10190	hypothetical protein
17	group_10131	hypothetical protein
17	group_9905	hypothetical protein
17	group_9904	hypothetical protein
17	group_9130	hypothetical protein
17	group_8475	hypothetical protein
17	group_3280	hypothetical protein
17	ydiP_2~~~ydiP	hypothetical protein;putative BsuMI modification methylase subunit YdiP
17	group_10985	hypothetical protein
17	group_9199	hypothetical protein
17	group_11797	hypothetical protein
17	group_7072	hypothetical protein
16	group_11070	hypothetical protein
16	group_3166	hypothetical protein
15	group_11812	hypothetical protein
15	group_11826	hypothetical protein
15	group_1418	hypothetical protein
15	group_11550	Modification methylase DpnIIB;hypothetical protein
15	group_7175	hypothetical protein
15	group_1279	hypothetical protein
15	xerC_1~~~xerC_2~~~xerC_5~~~xerC_10~~~xerC_3	Tyrosine recombinase XerC
15	group_787	hypothetical protein
15	group_10880	hypothetical protein
15	group_172	hypothetical protein
15	group_2677	hypothetical protein
15	group_500	hypothetical protein
15	group_460	hypothetical protein
15	group_375	hypothetical protein
15	group_315	hypothetical protein
15	group_155	hypothetical protein
15	group_11530	hypothetical protein

15	group_11055	hypothetical protein
15	group_7174	hypothetical protein
15	group_870	hypothetical protein
15	group_266	hypothetical protein
15	group_9263	hypothetical protein
15	group_9274	hypothetical protein
15	group_10332	hypothetical protein
15	group_8787	hypothetical protein
15	group_7922	hypothetical protein
14	smc_2~~~smc_1	Chromosome partition protein Smc
14	group_10693	putative protein/MSMEI_1241
14	group_9792	hypothetical protein
14	group_8210	hypothetical protein
14	group_1203	hypothetical protein
14	group_923	hypothetical protein
14	group_11314	hypothetical protein
14	group_8232	hypothetical protein
13	group_8279	hypothetical protein
13	group_3461	hypothetical protein
13	group_6106	hypothetical protein
13	group_11340	hypothetical protein
13	group_7413	hypothetical protein
13	group_7116	hypothetical protein
13	group_5897	hypothetical protein
13	group_3398	hypothetical protein
13	group_174	hypothetical protein
13	group_7187	hypothetical protein
13	group_5729	hypothetical protein
13	group_5710	hypothetical protein
13	group_10100	hypothetical protein
13	group_8916	hypothetical protein
13	group_7406	hypothetical protein
13	group_6548	hypothetical protein
13	group_6522	hypothetical protein
13	group_6071	hypothetical protein
13	group_4356	hypothetical protein
13	group_3307	hypothetical protein
13	group_2167	hypothetical protein
13	group_11837	hypothetical protein
13	group_10246	Glycine cleavage system transcriptional activator
13	group_6127	Glycine betaine/proline/ectoine/pipecolic acid transporter OusA

12	group_1741	hypothetical protein
12	group_1743	hypothetical protein
12	group_1444	hypothetical protein
12	group_7723	hypothetical protein
11	group_2694	hypothetical protein
11	teaD	Putative universal stress protein;TRAP-T-associated universal stress protein TeaD
11	group_2763	hypothetical protein
11	group_1929	hypothetical protein
11	group_8230	hypothetical protein
11	ccpA_2~~~ccpA_3~~~treR~~~ccpA_1	Catabolite control protein A;HTH-type transcriptional regulator TreR;hypothetical protein
11	sacX	Negative regulator of SacY activity
11	group_11352	Sucrose-6-phosphate hydrolase
11	group_11165	hypothetical protein
11	nemA_2~~~nemA_1~~~nemA_3	N-ethylmaleimide reductase
11	group_10746	Maltoporin
11	citB~~~citB_1~~~acn	Aconitate/2-methylaconitate hydratase;Aconitate hydratase A
11	group_10472	hypothetical protein
11	hicd	Homoisocitrate dehydrogenase
11	araG_1~~~mglA_3~~~araG~~~araG_2	Arabinose import ATP-binding protein AraG;Galactose/methyl galactoside import ATP-binding protein MglA;hypothetical protein
11	group_10058	hypothetical protein
11	ptsI_3~~~ptsI_2	hypothetical protein;Phosphoenolpyruvate-protein phosphotransferase
11	group_9715	hypothetical protein
11	group_9664	hypothetical protein
11	group_9615	hypothetical protein
11	group_9071	hypothetical protein
11	group_9027	hypothetical protein
11	group_9026	hypothetical protein
11	group_8315	hypothetical protein
11	resA_2~~~resA_3~~~resA_4	Thiol-disulfide oxidoreductase ResA
11	group_8215	hypothetical protein
11	acnB~~~leuC_2~~~acnB_2~~~acnB_3~~~acnB_1	Aconitate hydratase B;3-isopropylmalate dehydratase large subunit

11	nanR~~~lutR_2~~~lutR_1	HTH-type transcriptional repressor NanR;HTH-type transcriptional regulator LutR
11	emrB_2~~~emrB_3~~~emrY_2	Multidrug export protein EmrB;hypothetical protein;putative multidrug resistance protein EmrY
11	scrK~~~RBKS	Fructokinase;Ribokinase
11	glnP_1~~~glnP_2	Glutamine transport system permease protein GlnP;hypothetical protein
11	group_6780	hypothetical protein
11	group_6762	hypothetical protein
11	bauC_1~~~iolA_1	Putative 3-oxopropanoate dehydrogenase;Malonate-semialdehyde dehydrogenase
11	group_6133	hypothetical protein
11	group_6052	hypothetical protein
11	dmlR_11~~~dmlR_12	HTH-type transcriptional regulator DmlR
11	glnH_1	ABC transporter glutamine-binding protein GlnH
11	folE	GTP cyclohydrolase 1
11	glnM_1~~~yecS_2	putative glutamine ABC transporter permease protein GlnM;L-cystine transport system permease protein YecS
11	group_2519	Glutamine transport ATP-binding protein GlnQ
11	group_2342	hypothetical protein
11	group_1213	hypothetical protein
11	group_11924	hypothetical protein
11	group_11885	hypothetical protein
11	group_11670	hypothetical protein
11	group_11299	hypothetical protein
11	narH_1~~~narH_2~~~narY~~~narH	Respiratory nitrate reductase 1 beta chain;Respiratory nitrate reductase 2 beta chain
11	iolD~~~iolD_1~~~iolD_2	3D-(35/4)-trihydroxycyclohexane-12-dione hydrolase
11	group_11039	hypothetical protein
11	group_11007	hypothetical protein
11	iolG_1	Inositol 2-dehydrogenase/D-chiro-inositol 3-dehydrogenase
11	group_10685	hypothetical protein
11	iolE_2~~~iolE_1	Inosose dehydratase;hypothetical protein

11	group_10330	hypothetical protein
11	group_10037	2-amino-3-ketobutyrate coenzyme A ligase;hypothetical protein
11	group_10036	hypothetical protein
11	group_9970	hypothetical protein
11	group_9954	hypothetical protein
11	thpA~~~thpA_1	D-threitol-binding protein;hypothetical protein
11	iolG_2~~~iolG_1	hypothetical protein;Inositol 2-dehydrogenase/D-chiro-inositol 3-dehydrogenase
11	narJ	hypothetical protein;Nitrate reductase molybdenum cofactor assembly chaperone NarJ
11	sutR_2~~~sutR_3	HTH-type transcriptional regulator SutR;hypothetical protein
11	cmr	HTH-type transcriptional regulator Cmr;Cyclic AMP receptor protein;hypothetical protein
11	group_8988	putative epimerase/dehydratase;hypothetical protein
11	ytfE	Iron-sulfur cluster repair protein YtfE
11	degU_2~~~cheB_2~~~degU_1	Transcriptional regulatory protein DegU;Protein-glutamate methylesterase/protein-glutamine glutaminase
11	mobB_2~~~mobB_1	Molybdopterin-guanine dinucleotide biosynthesis adapter protein;hypothetical protein
11	dmlR_4~~~dmlR_14~~~dmlR_12~~~dmlR_17	HTH-type transcriptional regulator DmlR;hypothetical protein
11	group_8742	hypothetical protein
11	mglA_1~~~btuD_4~~~mglA_2	Galactose/methyl galactoside import ATP-binding protein MglA;Vitamin B12 import ATP-binding protein BtuD
11	murR_1~~~murR_2	HTH-type transcriptional regulator MurR
11	narI~~~narI_1~~~narI_2	Respiratory nitrate reductase 1 gamma chain;hypothetical protein
11	iolC~~~iolC_1	5-dehydro-2-deoxygluconokinase
11	iolE_1~~~iolE_2	Inosose dehydratase
11	queE_2~~~queE_1	7-carboxy-7-deazaguanine synthase;hypothetical protein

11	group_7386	Anaerobic ribonucleoside-triphosphate reductase;hypothetical protein
11	norR_5~~~norR2	hypothetical protein;Anaerobic nitric oxide reductase transcription regulator NorR;Nitric oxide reductase transcription regulator NorR2
11	group_6652	hypothetical protein
11	rbsC_2~~~rbsC_3	Ribose import permease protein RbsC
11	tolC	hypothetical protein;Outer membrane protein TolC
11	nasA_1~~~nasA_2	hypothetical protein;Nitrate transporter
11	narK_1~~~narK_2~~~narK	hypothetical protein;Nitrate/nitrite transporter NarK
11	group_2727	hypothetical protein
11	drgA	Protein DrgA
11	narZ_1~~~narG_2~~~narZ_2~~~narG_3~~~narG_1~~~narG	Respiratory nitrate reductase 2 alpha chain;Respiratory nitrate reductase 1 alpha chain
11	group_1385	hypothetical protein
11	group_1000	hypothetical protein
11	group_291	hypothetical protein
11	group_202	hypothetical protein
11	group_83	hypothetical protein
11	hdfR~~~hdfR_1~~~hdfR_2	HTH-type transcriptional regulator HdfR;hypothetical protein
11	yihG~~~yihG_1~~~yihG_2	putative acyltransferase YihG
11	group_11403	hypothetical protein
11	group_11177	hypothetical protein
11	apbE_2~~~apbE_1	FAD:protein FMN transferase;hypothetical protein
11	group_10534	hypothetical protein;HTH-type transcriptional regulator CdhR;HTH-type transcriptional regulator
11	group_10358	hypothetical protein;Cytochrome c-555
11	group_10352	hypothetical protein
11	group_10134	hypothetical protein
11	mmgF	putative protein;2-methylisocitrate lyase
11	group_9481	hypothetical protein
11	nosZ	Nitrous-oxide reductase
11	group_9388	hypothetical protein

11	yhbU_1~~~yhbU_2~~~yhbU	putative protease YhbU
11	group_9133	hypothetical protein
11	nosL	Copper-binding lipoprotein NosL
11	yahK_2~~~adhC2~~~adhC	Aldehyde reductase YahK;hypothetical protein;NADP-dependent alcohol dehydrogenase C 2;NADP-dependent alcohol dehydrogenase C
11	nosD~~~nosD_1~~~nosD_2	putative ABC transporter binding protein NosD
11	group_7486	hypothetical protein
11	group_6966	hypothetical protein
11	nosY	hypothetical protein;putative ABC transporter permease protein NosY
11	group_6712	Nitrate/nitrite sensor protein NarX;hypothetical protein
11	group_6341	hypothetical protein
11	pgrR_3~~~pgrR_4	HTH-type transcriptional regulator PgrR
11	nosF	putative ABC transporter ATP-binding protein NosF
11	group_5621	hypothetical protein
11	group_5504	hypothetical protein
11	aiiA	N-acyl homoserine lactonase;hypothetical protein
11	group_2819	hypothetical protein
11	aspC_2~~~aspC_1~~~hisC_6	Aspartate aminotransferase;hypothetical protein;Histidinol-phosphate aminotransferase
11	ucpA_2~~~gdh_1	Oxidoreductase UcpA;Galactitol 2-dehydrogenase
11	group_11599	hypothetical protein
11	TDO2	Tryptophan 23-dioxygenase
11	cefD~~~egtE	Isopenicillin N epimerase;hercynylcysteine sulfoxide lyase
11	group_10963	hypothetical protein
11	group_10737	hypothetical protein
11	group_9299	hypothetical protein
11	group_7900	hypothetical protein
11	group_7256	hypothetical protein
11	group_7254	hypothetical protein
11	group_5024	hypothetical protein

11	cyoD_1~~~cyoD_2	Cytochrome bo(3) ubiquinol oxidase subunit 4;hypothetical protein
11	tar_1	Methyl-accepting chemotaxis protein II;hypothetical protein
11	group_9418	hypothetical protein
11	group_7884	Outer membrane porin protein 32;hypothetical protein
11	cyoC_2~~~cyoC_1	Cytochrome bo(3) ubiquinol oxidase subunit 3;hypothetical protein
11	cyoA~~~cyoA_2~~~cyoA_1	hypothetical protein;Cytochrome bo(3) ubiquinol oxidase subunit 2
11	cyoB_2~~~cyoB_3~~~cyoB_4	Cytochrome bo(3) ubiquinol oxidase subunit 1
11	group_3283	hypothetical protein
11	group_9	hypothetical protein
11	group_11194	hypothetical protein
11	group_11157	hypothetical protein
11	group_10947	Outer membrane porin protein 32;Outer membrane porin protein;hypothetical protein
11	triA~~~atzA~~~dadD_1~~~dadD~~~dadD_2	Melamine deaminase;Atrazine chlorohydrolase;5'-deoxyadenosine deaminase;hypothetical protein
11	iolB~~~iolB_1	5-deoxy-glucuronate isomerase
11	ptsJ~~~lysN_2	Vitamin B6 salvage pathway transcriptional repressor PtsJ;hypothetical protein;2-aminoadipate transaminase
11	purU_2	Formyltetrahydrofolate deformylase;hypothetical protein
11	bcr_3~~~bcr_2~~~mdtL~~~bcr_1	Bicyclomycin resistance protein;Multidrug resistance protein MdtL;hypothetical protein
11	group_8096	hypothetical protein
11	group_7987	hypothetical protein
11	soxG	Sarcosine oxidase subunit gamma;hypothetical protein
11	soxD	Sarcosine oxidase subunit delta
11	cdhR_1~~~cdhR_2~~~cdhR_3	HTH-type transcriptional regulator CdhR;hypothetical protein
11	group_6526	hypothetical protein
11	group_6208	Aldo-keto reductase IolS
11	soxA_1~~~soxA_2~~~soxA_3	Sarcosine oxidase subunit alpha;hypothetical protein
11	soxB_1~~~soxB_2~~~soxB	Sarcosine oxidase subunit beta

11	sdaA_1~~~sdaA_2~~~sdaB	L-serine dehydratase;L-serine dehydratase 2
11	group_10663	hypothetical protein
11	group_10662	hypothetical protein
11	sir_2~~~sir_1	Sulfite reductase [ferredoxin];hypothetical protein
11	group_9618	hypothetical protein
11	group_9265	hypothetical protein
11	group_9264	hypothetical protein
11	group_9042	hypothetical protein
11	group_3374	hypothetical protein
11	group_3173	hypothetical protein
11	group_3078	Ribonuclease
11	group_2940	hypothetical protein
11	group_2911	hypothetical protein
11	group_11774	HTH-type transcriptional regulator TsaR;hypothetical protein
11	group_11703	hypothetical protein
11	group_11441	hypothetical protein
11	group_11353	hypothetical protein
11	group_10748	hypothetical protein
11	group_9864	hypothetical protein
11	group_9852	hypothetical protein
11	group_9652	hypothetical protein
11	group_8318	hypothetical protein
11	group_7957	hypothetical protein
11	group_6597	hypothetical protein
11	group_6210	hypothetical protein
11	dkgA	putative oxidoreductase/MSMEI_2347;2,5-diketo-D-gluconic acid reductase A
11	group_2171	hypothetical protein
11	cdhR_2~~~cdhR_1	HTH-type transcriptional regulator CdhR;hypothetical protein
11	group_9504	hypothetical protein
11	aatA_1~~~aatB	Aspartate/prephenate aminotransferase;Aspartate aminotransferase
11	murR_3	HTH-type transcriptional regulator MurR
11	group_8125	hypothetical protein
11	group_6607	hypothetical protein
11	group_6401	hypothetical protein
11	group_5955	hypothetical protein

11	group_4818	hypothetical protein
11	group_4757	hypothetical protein
11	group_4236	hypothetical protein
11	group_2733	hypothetical protein
11	group_2610	hypothetical protein
11	group_640	hypothetical protein
11	group_11802	hypothetical protein
11	group_11534	hypothetical protein
11	group_11094	hypothetical protein
11	group_10249	hypothetical protein
11	group_9469	hypothetical protein
11	group_9401	hypothetical protein
11	group_8683	hypothetical protein
11	group_8064	hypothetical protein
11	group_7466	hypothetical protein
11	group_7464	hypothetical protein;Outer membrane porin protein
11	group_7431	hypothetical protein
11	group_7193	hypothetical protein
11	group_7179	hypothetical protein
11	group_6549	hypothetical protein
11	group_6312	hypothetical protein
11	group_6113	hypothetical protein
11	lifO	Lipase chaperone
11	mtlK	Mannitol 2-dehydrogenase
11	sadA~~~sadA_2	hypothetical protein;Autotransporter adhesin SadA
11	group_3984	hypothetical protein
11	group_2996	hypothetical protein
11	pdp	Putative thymidine phosphorylase;Pyrimidine-nucleoside phosphorylase
11	group_2872	hypothetical protein
11	group_1206	hypothetical protein
11	group_945	hypothetical protein
11	group_628	hypothetical protein
11	xylB	Xylulose kinase
11	eryD	Erythritol catabolism regulatory protein EryD
11	group_9610	hypothetical protein
11	group_9533	hypothetical protein;Catalase
11	group_9443	hypothetical protein
11	group_9373	hypothetical protein

11	group_7864	hypothetical protein
11	group_7804	hypothetical protein
11	group_7787	hypothetical protein
11	gph_1~~~gph_2~~~yieH_2	Phosphoglycolate phosphatase;6-phosphogluconate phosphatase
11	group_7642	hypothetical protein
11	group_6980	hypothetical protein
11	group_6820	hypothetical protein
11	group_6056	hypothetical protein
11	kbaZ	D-tagatose-16-bisphosphate aldolase subunit KbaZ
11	group_4931	L-aspartate/glutamate-specific racemase;hypothetical protein
11	sugB	hypothetical protein;Trehalose transport system permease protein SugB
11	group_4810	hypothetical protein
11	group_4782	hypothetical protein
11	pgl_4~~~pgl_1	6-phosphogluconolactonase;hypothetical protein
11	malK	Maltose/maltodextrin import ATP-binding protein MalK;hypothetical protein
11	group_3889	hypothetical protein
11	group_2854	hypothetical protein
11	melD	Melibiose/raffinose/stachyose import permease protein MelD;hypothetical protein
11	group_2121	hypothetical protein
11	polS	Sorbitol dehydrogenase
11	group_339	Tagatose kinase
11	group_11892	hypothetical protein
11	group_10484	hypothetical protein
11	group_9348	hypothetical protein
11	group_8153	hypothetical protein
11	pgrR_7~~~argP~~~argP_1~~~pgrR_2	hypothetical protein;HTH-type transcriptional regulator PgrR;HTH-type transcriptional regulator ArgP
11	group_7770	hypothetical protein
11	dapH_3~~~cysE	2345-tetrahydropyridine-26-dicarboxylate N-acetyltransferase;Serine acetyltransferase;hypothetical protein

11	group_7423	hypothetical protein
11	group_6717	putative HTH-type transcriptional regulator
11	group_6716	hypothetical protein
11	group_6413	hypothetical protein
11	ppsB~~~entF	Plipastatin synthase subunit B;Enterobactin synthase component F
11	group_5113	hypothetical protein
11	group_5026	hypothetical protein;putative protein
11	gltR_2~~~gltR_1	HTH-type transcriptional regulator GltR
11	group_4800	hypothetical protein
11	ectB_1~~~ectB~~~ectB_2	Diaminobutyrate--2-oxoglutarate transaminase
11	pvdA	hypothetical protein;L-ornithine N(5)-monooxygenase
11	group_4648	hypothetical protein
11	nepI	Purine ribonucleoside efflux pump NepI
11	group_2555	hypothetical protein
11	soxA_3~~~soxA_4	Monomeric sarcosine oxidase
11	group_387	hypothetical protein
11	group_11178	hypothetical protein
11	group_9757	hypothetical protein
11	pgi_2~~~pgi_1	Glucose-6-phosphate isomerase
11	group_6599	hypothetical protein
11	group_5681	hypothetical protein
11	yodB	Cytochrome b561
11	group_595	hypothetical protein
11	group_9853	hypothetical protein
11	group_8154	hypothetical protein
11	group_7717	hypothetical protein
11	group_7544	hypothetical protein
11	group_2737	hypothetical protein
11	group_743	hypothetical protein
11	group_9527	hypothetical protein
11	group_2289	hypothetical protein
11	group_1963	hypothetical protein
11	group_7866	hypothetical protein
11	group_3723	hypothetical protein
11	group_2591	IS3 family transposase ISBcen15;hypothetical protein
11	group_11161	hypothetical protein

11	group_10763	hypothetical protein
11	group_10531	hypothetical protein
11	group_5046	hypothetical protein
11	group_8961	hypothetical protein
11	group_8375	hypothetical protein
11	group_7951	hypothetical protein
11	group_7946	hypothetical protein
11	group_352	hypothetical protein
11	iolG_1~~~iolG_3~~~iolG_4	Inositol 2-dehydrogenase/D-chiro-inositol 3-dehydrogenase;hypothetical protein
11	gcvA_7~~~gcvA_5	Glycine cleavage system transcriptional activator
11	group_7446	hypothetical protein
11	ssuD_2	Alkanesulfonate monooxygenase
11	dhpH	hypothetical protein;2,6-dihydroxypyridine 3-monooxygenase
11	dmlR_16	HTH-type transcriptional regulator DmlR
11	group_10487	hypothetical protein
11	group_7345	Cystathionine beta-lyase PatB
11	cyt	Cytochrome c-552;hypothetical protein
11	group_3128	hypothetical protein
11	group_3102	Cytochrome c-552
11	group_2960	hypothetical protein
11	group_10429	hypothetical protein
11	group_5516	hypothetical protein
11	jefA~~~stp~~~mdtD_1	hypothetical protein;Drug efflux pump JefA;Multidrug resistance protein Stp;Putative multidrug resistance protein MdtD
11	group_11168	hypothetical protein
11	group_3868	hypothetical protein
11	group_2769	IS3 family transposase ISRso14;IS3 family transposase ISAtu5;hypothetical protein
11	group_265	hypothetical protein
11	group_9197	hypothetical protein
11	group_7412	hypothetical protein
11	group_7071	hypothetical protein
11	lrpC_3~~~lrpC_4	hypothetical protein;HTH-type transcriptional regulator LrpC
11	group_2447	hypothetical protein

11	group_2275	hypothetical protein
11	group_7863	hypothetical protein
11	group_3808	hypothetical protein
11	group_1614	hypothetical protein
11	adhC2~~~yahK_1~~~tdh_2~~~adhA_2~~~adhA_3	NADP-dependent alcohol dehydrogenase C 2;Aldehyde reductase YahK;L-threonine 3-dehydrogenase;hypothetical protein;putative formaldehyde dehydrogenase AdhA
11	group_2617	hypothetical protein
11	group_7956	hypothetical protein
11	group_7737	hypothetical protein
11	group_7563	hypothetical protein
11	group_4127	hypothetical protein
11	group_2973	Putative universal stress protein;Universal stress protein/MSMEI_3859
11	group_2751	hypothetical protein;IS3 family transposase IS222
11	group_10714	Vitamin B12-dependent ribonucleoside-diphosphate reductase;hypothetical protein
11	group_8236	hypothetical protein
11	group_7050	hypothetical protein
11	group_6923	hypothetical protein
11	group_5856	Alcohol dehydrogenase
11	emrY_2~~~emrY_1~~~emrB_4~~~emrB_2~~~emrB_1	hypothetical protein;putative multidrug resistance protein EmrY;Multidrug export protein EmrB
11	iorA_1~~~iorA_3~~~iorA_2	Isoquinoline 1-oxidoreductase subunit alpha
11	group_3751	Isoquinoline 1-oxidoreductase subunit beta;hypothetical protein
11	paoD_2~~~pucA_2~~~paoD~~~pucA	Molybdenum cofactor insertion chaperone PaoD;putative xanthine dehydrogenase subunit A
11	group_3630	hypothetical protein
11	dksA_3~~~dksA_2~~~dksA_1	RNA polymerase-binding transcription factor DksA
11	group_1852	hypothetical protein
11	group_1162	hypothetical protein
11	prs_2~~~prs_1	Ribose-phosphate pyrophosphokinase;Putative ribose-phosphate pyrophosphokinase

11	adh_2~~~adh_1	Alcohol dehydrogenase
11	group_7775	hypothetical protein
11	group_5199	hypothetical protein
11	group_5013	IS3 family transposase ISBmu5;IS3 family transposase ISRme12;hypothetical protein
11	mas~~~tdh	hypothetical protein;Mycocerosic acid synthase;L-threonine 3-dehydrogenase
11	group_4049	hypothetical protein
11	group_10479	hypothetical protein
11	farA	Fatty acid resistance protein FarA
11	oprM_7~~~oprM_1~~~oprM_9	Outer membrane protein OprM;hypothetical protein
11	group_4410	hypothetical protein
11	group_2424	hypothetical protein
11	group_2383	hypothetical protein
11	group_2136	hypothetical protein
11	group_2122	hypothetical protein
11	group_2101	hypothetical protein
11	group_2057	hypothetical protein
11	group_1978	hypothetical protein
11	group_10310	IS3 family transposase ISRso11
11	group_7917	hypothetical protein
11	group_7915	hypothetical protein
11	group_7892	hypothetical protein
11	mrpX	Methionine-rich peptide X;hypothetical protein
11	group_3219	hypothetical protein
11	group_2894	hypothetical protein
11	group_2779	putative multidrug-efflux transporter
11	group_2215	hypothetical protein
11	group_2166	hypothetical protein
11	group_2145	hypothetical protein
11	group_1892	hypothetical protein
11	yedY1	Putative protein-methionine-sulfoxide reductase subunit YedZ1
11	group_9578	hypothetical protein
11	group_9384	hypothetical protein
11	rhaS_2~~~rhaS_3~~~rhaS_4	HTH-type transcriptional activator RhaS
11	group_3753	putative zinc-binding alcohol dehydrogenase;hypothetical protein

11	group_3247	hypothetical protein;IS3 family transposase ISPsy24
11	group_3150	hypothetical protein
11	yecD_1~~~yecD_2~~~yecD~~~sbnD_2~~~sbnD_1~~~mdtG	Isochorismatase family protein YecD;Staphyloferrin B transporter;hypothetical protein;N-carbamoylsarcosine amidase;Multidrug resistance protein MdtG
11	rspR_3~~~rspR_1	hypothetical protein;HTH-type transcriptional repressor RspR
11	yedZ1	Putative protein-methionine-sulfoxide reductase subunit YedZ1
11	group_2005	hypothetical protein
11	group_1922	hypothetical protein
11	group_1921	hypothetical protein
11	group_1862	hypothetical protein
11	group_1803	hypothetical protein
11	birA_1	hypothetical protein;Bifunctional ligase/repressor BirA
11	group_1773	hypothetical protein
11	group_1740	hypothetical protein
11	lrpC_2	hypothetical protein;HTH-type transcriptional regulator LrpC
11	group_1251	hypothetical protein
11	group_3592	hypothetical protein
11	group_3148	hypothetical protein
11	group_2841	hypothetical protein
11	group_2531	hypothetical protein
11	group_1664	hypothetical protein
11	group_7766	Flavin-dependent monooxygenase oxygenase subunit HsaA
11	group_11862	hypothetical protein
11	group_7658	hypothetical protein
11	group_2058	hypothetical protein
11	group_9494	hypothetical protein
11	group_8007	hypothetical protein;ATP-dependent RecD-like DNA helicase
11	group_7869	hypothetical protein
11	group_3144	hypothetical protein
11	group_1454	hypothetical protein
11	yqcF~~~yqcF_2	Antitoxin YqcF;hypothetical protein
11	group_11565	hypothetical protein
11	group_10188	hypothetical protein

11	nuoF_1~~~nuoF_3~~~nuoF_2	NADH-quinone oxidoreductase subunit F
11	aspT	Aspartate/alanine antiporter;hypothetical protein
11	sdhA_1~~~sdhA_2	Succinate dehydrogenase flavoprotein subunit;hypothetical protein
11	rcsC_11~~~rcsC_5~~~lldD~~~dctB_1	Sensor histidine kinase RcsC;L-lactate dehydrogenase;C4-dicarboxylate transport sensor protein DctB
11	group_4961	hypothetical protein
11	group_3223	hypothetical protein
11	maeA~~~maeA_2~~~maeA_1~~~mleS	NAD-dependent malic enzyme;putative NAD-dependent malic enzyme 2;Malolactic enzyme
11	dmlR_19	HTH-type transcriptional regulator DmlR
11	group_7809	hypothetical protein
11	yidZ~~~pcpR_7	hypothetical protein;HTH-type transcriptional regulator YidZ;PCP degradation transcriptional activation protein
11	group_6202	hypothetical protein
11	group_4831	hypothetical protein
11	group_9592	Transcriptional regulator SlyA;hypothetical protein
11	group_9256	hypothetical protein
11	dctD_1~~~dctD_2	C4-dicarboxylate transport transcriptional regulatory protein DctD
11	group_7266	hypothetical protein
11	group_4833	hypothetical protein
11	group_1976	hypothetical protein
11	group_1089	hypothetical protein
11	group_11541	hypothetical protein
11	group_7326	hypothetical protein
11	group_6657	hypothetical protein;IS5 family transposase ISRso1
11	group_5408	hypothetical protein
11	group_5314	hypothetical protein
11	group_2656	hypothetical protein
11	group_11922	hypothetical protein
11	group_2575	IS5 family transposase ISRso1
11	group_12046	hypothetical protein

11	mhpD	2-keto-4-pentenoate hydratase;hypothetical protein
11	dsbA_2~~~dsbA_1	Thiol:disulfide interchange protein DsbA;hypothetical protein
11	amnE	4-oxalocrotonate decarboxylase
11	dmpI	2-hydroxymuconate tautomerase
11	group_9817	hypothetical protein
11	group_9169	RNA 2'3'-cyclic phosphodiesterase
11	amnF	2-oxopent-4-enoate hydratase
11	group_8702	3-hydroxy-2-methylpyridine-45-dicarboxylate 4-decarboxylase;hypothetical protein
11	mhpF~~~dmpF	Acetaldehyde dehydrogenase
11	mhbT_2~~~mhbT_1	3-hydroxybenzoate transporter MhbT
11	xylG_2~~~xylG_1~~~betB_5	2-hydroxymuconic semialdehyde dehydrogenase;NAD/NADP-dependent betaine aldehyde dehydrogenase
11	hsaC	Iron-dependent extradiol dioxygenase
11	mhpE_2~~~mhpE_1~~~mhpE	4-hydroxy-2-oxovalerate aldolase
11	group_2535	hypothetical protein
11	intS_1~~~intS~~~intS_2	Prophage integrase IntS;hypothetical protein
11	group_2008	hypothetical protein
11	group_531	hypothetical protein
11	group_507	hypothetical protein
11	group_356	hypothetical protein
11	group_274	hypothetical protein
11	group_198	hypothetical protein
11	group_12045	hypothetical protein
11	rstA_2	Transcriptional regulatory protein RstA;hypothetical protein
11	group_7921	hypothetical protein
11	group_7885	hypothetical protein
11	group_7883	hypothetical protein
11	group_7662	hypothetical protein
11	group_3482	hypothetical protein
11	amnD~~~rutC_2	2-aminomuconate deaminase;RutC family protein;Putative aminoacrylate peracid reductase RutC
11	group_2190	hypothetical protein
11	group_2150	hypothetical protein

11	group_2149	hypothetical protein
11	group_2108	hypothetical protein
11	group_2078	hypothetical protein
11	group_9341	hypothetical protein
11	group_9329	hypothetical protein
11	group_7740	hypothetical protein
11	kgtP_5~~~nanT_2	hypothetical protein;IS5 family transposase ISRso1;Alpha-ketoglutarate permease;Sialic acid transporter NanT
11	group_2952	hypothetical protein
11	cya_1	hypothetical protein;Bifunctional hemolysin/adenylate cyclase
11	slyA_1~~~slyA_3	Transcriptional regulator SlyA
11	group_6396	hypothetical protein
11	group_3781	hypothetical protein
11	livQ	6'''-hydroxyparomomycin C oxidase
11	group_3775	hypothetical protein
11	group_3023	hypothetical protein
11	group_785	hypothetical protein
11	group_78	hypothetical protein
11	group_10316	hypothetical protein
11	group_7207	hypothetical protein
11	group_1254	hypothetical protein
11	group_1069	hypothetical protein
11	group_575	hypothetical protein
11	group_9396	hypothetical protein
11	group_2075	hypothetical protein
11	group_4096	hypothetical protein
11	group_9150	hypothetical protein
11	group_7633	hypothetical protein
11	group_3697	hypothetical protein
11	group_3447	hypothetical protein
11	group_2551	hypothetical protein
11	group_426	hypothetical protein
11	group_7895	hypothetical protein
11	group_2249	hypothetical protein
11	group_607	hypothetical protein
11	group_11748	hypothetical protein
11	group_9438	hypothetical protein
11	group_9064	hypothetical protein
11	group_7911	hypothetical protein
11	group_6858	hypothetical protein
11	group_2621	IS3 family transposase ISAisp2

11	group_2510	IS110 family transposase ISBcen4
11	group_11923	hypothetical protein
11	gcvA_6	Glycine cleavage system transcriptional activator
11	group_8043	Alcohol dehydrogenase
11	group_7051	hypothetical protein
11	group_6429	hypothetical protein
11	rfnT_2~~~rfnT	Riboflavin transporter RfnT
11	glnQ_4~~~glnQ_5	Glutamine transport ATP-binding protein GlnQ
11	group_4453	hypothetical protein
11	group_2860	hypothetical protein;IS5 family transposase ISRso18;IS5 family transposase IS1405
11	group_925	hypothetical protein
11	group_776	hypothetical protein
11	xerC_2~~~xerC_3~~~xerC_4~~~xerC_1	Tyrosine recombinase XerC;hypothetical protein
11	group_8187	Competence protein ComM
11	group_7910	hypothetical protein
11	group_5039	hypothetical protein
11	lldD_1~~~lldD_2~~~lldD	L-lactate dehydrogenase
11	group_4006	hypothetical protein
11	group_3765	hypothetical protein
11	group_11800	hypothetical protein
11	group_10427	Carboxymethylenebutenolidase
11	group_10273	hypothetical protein
11	group_9328	hypothetical protein
11	group_7927	hypothetical protein
11	group_3712	hypothetical protein
11	group_2635	hypothetical protein
11	group_2223	IS3 family transposase ISBcen15;IS3 family transposase ISRso12
11	group_954	hypothetical protein
11	group_9541	hypothetical protein
11	speE_2~~~speE_3	Polyamine aminopropyltransferase
11	group_1351	hypothetical protein
11	group_10626	hypothetical protein
11	group_9248	hypothetical protein;IS4 family transposase ISCro3
11	group_7320	hypothetical protein
11	group_6965	hypothetical protein
11	group_1840	hypothetical protein
11	group_1731	hypothetical protein

11	group_898	hypothetical protein
11	group_2905	hypothetical protein
11	group_2838	hypothetical protein
11	group_2401	hypothetical protein
11	group_2396	hypothetical protein
11	group_2293	IS3 family transposase ISBcen15;hypothetical protein
11	group_3754	hypothetical protein
11	group_2113	hypothetical protein
11	group_11234	hypothetical protein
11	group_10557	hypothetical protein
11	group_3770	hypothetical protein
11	group_2440	hypothetical protein;IS4 family transposase ISCro3
11	group_2330	IS30 family transposase IS1382;IS30 family transposase IST3091;IS30 family transposase ISRme10
11	group_10279	hypothetical protein
11	group_9326	hypothetical protein
11	group_8323	hypothetical protein
11	group_7721	hypothetical protein
11	group_2397	hypothetical protein
11	group_2274	hypothetical protein
11	group_1248	hypothetical protein
11	group_1044	hypothetical protein
11	group_765	hypothetical protein
11	group_10334	hypothetical protein
11	group_9319	hypothetical protein
11	group_7120	hypothetical protein
11	group_2987	IS66 family transposase ISPpu19;hypothetical protein;IS66 family transposase ISBcen19
11	group_2549	hypothetical protein
11	group_11000	hypothetical protein
11	group_10223	IS4 family transposase ISCro3
11	group_10163	hypothetical protein
11	group_7703	IS5 family transposase IS1405;hypothetical protein
11	group_2832	IS4 family transposase ISCro3;hypothetical protein
11	group_2804	hypothetical protein
11	group_2391	IS3 family transposase ISAisp2;Putative transposase InsK for insertion sequence element IS150

11	group_1328	hypothetical protein
11	group_1098	IS3 family transposase ISRso11
11	group_1050	hypothetical protein
11	group_9437	hypothetical protein
11	group_8921	IS3 family transposase ISAisp2
11	group_7727	hypothetical protein
11	group_6609	hypothetical protein;IS3 family transposase ISAisp2
11	group_5344	hypothetical protein;IS4 family transposase ISCro3
11	xerC_6~~~xerC_5~~~xerD_1~~~xerD_2~~~xerC_4	Tyrosine recombinase XerC;Tyrosine recombinase XerD
11	group_3217	hypothetical protein
11	group_2896	hypothetical protein
11	group_2624	hypothetical protein
11	group_2587	hypothetical protein
11	group_2586	hypothetical protein
11	group_1437	hypothetical protein
11	group_697	hypothetical protein
11	group_10218	IS4 family transposase ISCro3
11	group_9141	hypothetical protein
11	group_7265	hypothetical protein
11	group_4864	IS5 family transposase IS1405
11	group_1663	hypothetical protein
11	group_11447	hypothetical protein
11	group_9949	hypothetical protein
11	group_8824	hypothetical protein
11	group_7140	hypothetical protein
11	group_7073	hypothetical protein
11	group_4583	hypothetical protein
11	group_4047	hypothetical protein
11	group_3686	hypothetical protein
11	group_2699	hypothetical protein
11	group_2388	hypothetical protein
11	group_1380	hypothetical protein
11	group_730	hypothetical protein
11	group_497	hypothetical protein
11	group_62	hypothetical protein
11	group_11226	hypothetical protein
11	group_11149	hypothetical protein
11	group_10222	IS4 family transposase ISCro3
11	group_8180	IS4 family transposase ISPhsp1;hypothetical protein;IS4 family transposase ISCro3

11	group_7415	hypothetical protein
11	group_1870	hypothetical protein
11	group_1634	IS3 family transposase ISBmu5;hypothetical protein
11	group_1621	hypothetical protein
11	group_1066	hypothetical protein
11	mntH_1	Divalent metal cation transporter MntH
11	group_11166	hypothetical protein
11	group_10822	hypothetical protein
11	group_10431	hypothetical protein
11	group_9546	hypothetical protein
11	group_8005	hypothetical protein
11	group_7865	IS5 family transposase ISRso18;IS5 family transposase IS1405
11	group_7729	hypothetical protein
11	group_7728	hypothetical protein
11	group_7330	hypothetical protein
11	group_5659	hypothetical protein
11	group_5545	hypothetical protein
11	group_4874	hypothetical protein;IS110 family transposase ISBcen4
11	group_4375	hypothetical protein
11	group_4374	hypothetical protein
11	group_2358	IS5 family transposase ISRso18;hypothetical protein;IS5 family transposase IS1405
11	group_1184	IS1182 family transposase ISBusp4;IS1182 family transposase ISPpu16;hypothetical protein
11	group_767	IS256 family transposase ISRso7;hypothetical protein
11	group_224	hypothetical protein
11	group_11540	hypothetical protein
11	group_11067	hypothetical protein
11	group_10825	hypothetical protein
11	group_10809	hypothetical protein
11	group_10284	hypothetical protein
11	group_10231	IS4 family transposase ISCro3
11	group_9359	hypothetical protein
11	group_9121	hypothetical protein
11	group_7329	hypothetical protein
11	group_7322	IS3 family transposase ISRso16;hypothetical protein

11	rhsC_3~~~rhsC_1	hypothetical protein;Putative deoxyribonuclease RhsC
11	group_2977	IS110 family transposase ISBma3;hypothetical protein
11	group_2527	hypothetical protein;IS4 family transposase ISCro3
11	group_2452	hypothetical protein
11	group_1688	hypothetical protein
11	group_989	hypothetical protein
11	group_454	hypothetical protein
11	group_10807	hypothetical protein
11	group_10614	hypothetical protein
11	group_10491	Branched-chain-amino-acid aminotransferase
11	ribZ	Riboflavin transporter RibZ
11	group_5820	hypothetical protein
11	group_5776	hypothetical protein
11	group_4372	hypothetical protein
11	group_3308	hypothetical protein
11	benM_4~~~benM_2	HTH-type transcriptional regulator BenM
11	group_1564	hypothetical protein
11	group_1492	hypothetical protein
11	group_10767	hypothetical protein
11	group_5573	IS4 family transposase ISCro3;hypothetical protein
11	group_4854	hypothetical protein
11	group_4837	hypothetical protein
11	group_4588	hypothetical protein
11	group_4526	hypothetical protein
11	group_4241	hypothetical protein
11	group_4074	hypothetical protein
11	group_3973	hypothetical protein
11	group_334	hypothetical protein
11	group_11390	hypothetical protein
11	group_10267	hypothetical protein
11	group_7964	hypothetical protein
11	group_7017	IS110 family transposase ISBcen4;IS110 family transposase ISPa49;hypothetical protein
11	group_3064	IS1182 family transposase ISBusp4
11	group_2453	IS5 family transposase IS1405
11	group_6571	hypothetical protein
11	group_4355	IS5 family transposase IS1405

11	group_7961	hypothetical protein;IS4 family transposase ISCro3
11	group_7186	IS5 family transposase ISRso1
11	group_4208	hypothetical protein
11	group_3720	hypothetical protein
11	group_3716	IS5 family transposase IS1405
11	group_2552	IS3 family transposase ISAisp2;hypothetical protein
11	group_728	hypothetical protein
11	group_11389	hypothetical protein
11	group_10747	hypothetical protein
11	group_10627	hypothetical protein
11	group_10376	IS110 family transposase ISMno14
11	group_10242	IS3 family transposase ISBam2
11	group_10240	IS3 family transposase ISRso14
11	group_9865	Sarcosine oxidase subunit gamma;hypothetical protein
11	soxA_2~~~soxA_4	Sarcosine oxidase subunit alpha
11	group_7708	hypothetical protein
11	group_6747	hypothetical protein
11	soxB	Sarcosine oxidase subunit beta
11	sdaB	L-serine dehydratase 2
11	group_6193	hypothetical protein
11	group_4781	hypothetical protein
11	group_4369	hypothetical protein
11	group_2553	IS3 family transposase ISAisp2
11	cdhR_1~~~cdhR_4	HTH-type transcriptional regulator CdhR
11	ydeP_1~~~ydeP_2	Protein YdeP
11	group_11004	hypothetical protein
11	group_10940	hypothetical protein
11	group_10639	IS5 family transposase IS1405
11	group_10633	IS5 family transposase IS1405
11	group_10629	hypothetical protein
11	opuAB	Glycine betaine transport system permease protein OpuAB
11	group_10077	hypothetical protein
11	group_9003	hypothetical protein
11	cdhR_6	HTH-type transcriptional regulator CdhR
11	cdhR_4	HTH-type transcriptional regulator CdhR
11	group_7512	hypothetical protein;IS5 family transposase IS1405

11	flhA_3	S-(hydroxymethyl)glutathione dehydrogenase
11	group_7106	NADPH oxidoreductase
11	group_7105	Outer membrane porin protein
11	group_7023	Glycine betaine/proline betaine transport system ATP-binding protein ProV
11	fixB	Protein FixB
11	fdhA	Glutathione-independent formaldehyde dehydrogenase
11	group_6440	Carnitine monooxygenase oxygenase subunit
11	group_6288	Serine hydroxymethyltransferase 2
11	group_6148	hypothetical protein
11	naiP_3	Putative niacin/nicotinamide transporter NaiP
11	group_4365	hypothetical protein
11	group_4313	hypothetical protein
11	group_4010	hypothetical protein
11	stcD	putative N-methylproline demethylase
11	group_3446	IS110 family transposase ISMno14
11	group_2919	hypothetical protein
11	purU_3	Formyltetrahydrofolate deformylase
11	group_1207	hypothetical protein
11	group_1094	IS5 family transposase IS1405
11	group_1093	hypothetical protein
11	group_1018	hypothetical protein
11	moaA_2	GTP 3'8-cyclase;hypothetical protein
11	group_667	hypothetical protein
11	group_11726	IS5 family transposase IS1405
11	group_10645	IS5 family transposase IS1405
11	group_10634	IS5 family transposase IS1405
11	group_10622	hypothetical protein
11	group_10220	IS4 family transposase ISCro3
11	group_9152	hypothetical protein;IS5 family transposase IS1405
11	group_7726	hypothetical protein
11	group_5393	hypothetical protein
11	group_4366	hypothetical protein
11	group_3603	hypothetical protein
11	group_3594	hypothetical protein
11	group_3527	hypothetical protein
11	group_3522	hypothetical protein

11	group_3365	hypothetical protein
11	group_3235	ISNCY family transposase ISBcen27
11	group_1895	TAL effector protein Rip19
11	group_899	hypothetical protein
11	group_10640	IS5 family transposase IS1405
11	group_10617	hypothetical protein
11	group_10379	IS110 family transposase ISMno14
11	group_883	hypothetical protein;IS5 family transposase IS1405
11	group_10692	hypothetical protein
11	group_10656	IS21 family transposase ISRso6
11	amnD	2-aminomuconate deaminase
11	group_8011	hypothetical protein
11	group_7321	hypothetical protein;IS5 family transposase IS1405
11	group_4361	hypothetical protein
11	group_1788	IS5 family transposase IS1405;IS5 family transposase ISRso18
11	rhtB_3	Homoserine/homoserine lactone efflux protein
11	group_10224	IS4 family transposase ISCro3
11	group_7183	IS3 family transposase ISRso10
11	group_5082	hypothetical protein
11	group_9625	hypothetical protein
10	group_3323	hypothetical protein
10	group_3120	hypothetical protein
10	cpdA_5~~~cpdA_4~~~cpdA_3	3'5'-cyclic adenosine monophosphate phosphodiesterase CpdA
10	group_1713	queuosine precursor transporter;hypothetical protein
10	group_1900	hypothetical protein
10	group_1484	hypothetical protein;putative HTH-type transcriptional regulator
9	group_2543	hypothetical protein;IS66 family transposase ISPa82;IS66 family transposase ISAeh1
9	group_2464	IS66 family transposase IS1313
9	group_10466	hypothetical protein
8	group_12012	IS701 family transposase ISRso17
8	group_5593	IS701 family transposase ISRso17
7	group_3715	hypothetical protein
7	group_11742	hypothetical protein
6	group_2238	hypothetical protein
6	group_2237	Lysozyme RrrD

6	group_2236	hypothetical protein
6	xerC_3~~~xerC_5~~~xerC_2~~~xerC_4	Tyrosine recombinase XerC
5	group_2515	IS5 family transposase ISBmu20;IS5 family transposase IS1021;IS5 family transposase ISRso18;IS5 family transposase IS1405;hypothetical protein
5	group_2724	IS5 family transposase IS1405;hypothetical protein;IS5 family transposase ISRso18;IS5 family transposase IS1021
5	group_1046	hypothetical protein;IS5 family transposase IS1021
5	group_3127	IS5 family transposase ISRso18;IS5 family transposase IS1405;IS5 family transposase IS1021;hypothetical protein
5	group_1202	hypothetical protein;IS5 family transposase ISBmu2;IS5 family transposase IS1021
5	group_2889	IS5 family transposase IS1021;hypothetical protein
5	group_1195	hypothetical protein;IS5 family transposase IS1021
5	group_977	IS5 family transposase IS1021;hypothetical protein
5	group_3187	IS5 family transposase IS1021
5	group_828	IS5 family transposase IS1021;hypothetical protein
5	group_1457	IS5 family transposase IS1021;hypothetical protein
5	group_1322	IS5 family transposase IS1021;hypothetical protein
5	group_3768	IS5 family transposase IS1021;hypothetical protein
5	group_3099	IS5 family transposase IS1021
5	group_2320	IS5 family transposase IS1021;hypothetical protein
5	group_1030	IS5 family transposase IS1021
5	group_796	IS5 family transposase IS1021;hypothetical protein
5	group_1527	IS5 family transposase IS1021;hypothetical protein
5	group_1142	IS5 family transposase IS1021
5	group_3709	IS5 family transposase IS1021;hypothetical protein

5	group_2325	IS5 family transposase IS1021;hypothetical protein
5	group_1128	IS5 family transposase IS1021
5	group_2698	IS5 family transposase IS1021;hypothetical protein
4	group_11873	hypothetical protein
4	group_11586	hypothetical protein
4	group_10342	hypothetical protein
4	group_5626	hypothetical protein
4	capV	cGAMP-activated phospholipase;hypothetical protein
4	group_4325	hypothetical protein
4	group_4197	hypothetical protein
3	group_11485	hypothetical protein
3	group_10978	hypothetical protein
3	group_10130	hypothetical protein
2	group_3722	IS3 family transposase ISRso11
2	group_7616	IS3 family transposase ISBcen10;hypothetical protein
1	group_1670	hypothetical protein
1	group_1319	hypothetical protein
210	group_10907	hypothetical protein
210	group_4078	hypothetical protein
209	group_5148	hypothetical protein
209	group_853	hypothetical protein
208	group_8075	hypothetical protein
208	group_7870	IS110 family transposase ISPye16
208	group_7304	hypothetical protein
208	group_5103	hypothetical protein
207	group_5138	hypothetical protein
207	estP	Esterase EstP;hypothetical protein
206	group_6368	hypothetical protein
206	group_1854	hypothetical protein
205	dmlR_15	HTH-type transcriptional regulator DmlR
205	group_8749	hypothetical protein
205	polS_2	Sorbitol dehydrogenase;Galactitol 2-dehydrogenase
204	group_415	hypothetical protein
204	group_433	hypothetical protein
203	group_12047	Outer membrane porin protein;hypothetical protein
203	yphB~~~yphB_1~~~yphB_2	putative protein YphB
203	group_10422	hypothetical protein

203	group_3466	Outer membrane porin protein 32;hypothetical protein
203	group_1431	hypothetical protein
203	araG_1	Arabinose import ATP-binding protein AraG
203	pat~~~pat_1	Phosphinothricin N-acetyltransferase;hypothetical protein
203	group_11865	hypothetical protein
203	group_11717	hypothetical protein
203	rbn_2~~~rbn_3~~~rbn_1	Ribonuclease BN
203	group_1358	hypothetical protein
203	umaA	S-adenosylmethionine-dependent methyltransferase UmaA
202	group_10028	HTH-type transcriptional regulator HdfR
202	msrP_3	Protein-methionine-sulfoxide reductase catalytic subunit MsrP
202	group_8186	Nitrilase
202	metC_2	Cystathionine beta-lyase MetC
202	mdcG	Phosphoribosyl-dephospho-CoA transferase
202	madA	Acetyl-S-ACP:malonate ACP transferase
202	mdcC	Malonate decarboxylase acyl carrier protein
202	group_1698	hypothetical protein
202	group_950	hypothetical protein
201	group_955	hypothetical protein
201	group_771	hypothetical protein
200	group_4211	hypothetical protein
200	group_4210	hypothetical protein
200	group_2888	hypothetical protein
199	group_3738	Transposon Tn7 transposition protein TnsB;hypothetical protein
199	group_3736	hypothetical protein
198	tsr_3~~~tsr_1~~~tar_1~~~tar_2	Methyl-accepting chemotaxis protein I;hypothetical protein;Methyl-accepting chemotaxis protein II
198	pgl_4~~~pgl_3	6-phosphogluconolactonase
197	group_2411	putative protein;hypothetical protein
197	acrR	hypothetical protein;HTH-type transcriptional regulator AcrR
196	group_10659	hypothetical protein

196	group_10655	IS21 family transposase ISRso6
195	group_2051	hypothetical protein
195	group_2050	hypothetical protein
195	yraJ	hypothetical protein;Outer membrane usher protein YraJ
195	group_1959	hypothetical protein
195	group_1887	hypothetical protein
195	group_1843	hypothetical protein
194	group_9124	Cystathionine beta-lyase PatB
194	group_7630	hypothetical protein
194	sqr	hypothetical protein;Sulfide-quinone reductase
193	copB	Copper resistance protein B
193	copA_3~~~copA_2~~~copA_4~~~copA_5	Copper resistance protein A;hypothetical protein
193	cusR_1~~~cusR_2~~~rcsC_7~~~cusR	Transcriptional regulatory protein CusR;Sensor histidine kinase RcsC
193	sasA_12~~~cusS~~~sasA_8~~~cusS_2~~~cusS_1	Adaptive-response sensory-kinase SasA;Sensor histidine kinase CusS;hypothetical protein
192	group_8066	hypothetical protein
192	group_6981	hypothetical protein;Vitamin B12 import ATP-binding protein BtuD
192	group_5464	hypothetical protein
192	group_1473	hypothetical protein
192	group_901	hypothetical protein
191	cmpR_2	hypothetical protein;HTH-type transcriptional activator CmpR
191	group_2062	hypothetical protein
191	group_1984	putative HTH-type transcriptional regulator
191	naiP_1~~~yjhB~~~naiP_2~~~naiP	Putative niacin/nicotinamide transporter NaiP;Putative metabolite transport protein YjhB;hypothetical protein
190	group_10934	hypothetical protein
190	group_4385	hypothetical protein
189	group_11788	hypothetical protein
189	group_7113	hypothetical protein
189	group_5015	hypothetical protein
189	fecI_2	putative RNA polymerase sigma factor FecI
189	ttdB_2	L(+)-tartrate dehydratase subunit beta
189	group_3783	hypothetical protein

189	pglA_2	Polygalacturonase
188	copD	Copper resistance protein D
188	copC_1~~~copC_2~~~copC	Copper resistance protein C
187	dctA2_3~~~dctA2_2	C4-dicarboxylate transport protein 2
187	pcaB_1~~~pcaB_2	3-carboxy-ciscis-muconate cycloisomerase
187	dasR	HTH-type transcriptional repressor DasR
186	group_11553	hypothetical protein
186	group_11097	hypothetical protein
186	group_10337	hypothetical protein
186	group_9442	hypothetical protein
185	group_6703	hypothetical protein
185	group_4022	hypothetical protein
185	rep_3	ATP-dependent DNA helicase Rep
184	group_9260	hypothetical protein
184	group_9259	hypothetical protein
184	group_7171	hypothetical protein
184	group_1132	3'-5' exoribonuclease
183	sutR_2~~~sutR_3~~~sutR_1~~~sutR_4	hypothetical protein;HTH-type transcriptional regulator SutR
183	ttr	Acetyltransferase
182	group_10911	hypothetical protein
182	group_4024	hypothetical protein
181	ompR_3	hypothetical protein;Transcriptional regulatory protein OmpR
181	hopD2	Effector protein hopD2
181	group_4152	hypothetical protein
180	group_7631	hypothetical protein
180	nemA_3	N-ethylmaleimide reductase
180	group_4700	hypothetical protein
180	acuR	Transcriptional regulator AcuR;hypothetical protein
179	group_6094	hypothetical protein
179	group_8221	hypothetical protein
178	intA_1~~~intA_2	Prophage integrase IntA
178	group_94	hypothetical protein
177	group_10997	hypothetical protein
177	group_6262	hypothetical protein
177	group_4839	hypothetical protein
177	group_3054	hypothetical protein
176	xerC_4~~~xerC_1	Tyrosine recombinase XerC;hypothetical protein
176	group_747	hypothetical protein

175	group_8175	IS30 family transposase ISHar4;hypothetical protein;IS30 family transposase ISHar5
175	group_4611	hypothetical protein
174	group_3211	hypothetical protein
174	virS	hypothetical protein;HTH-type transcriptional regulator VirS
174	actIII~~~hcaB_2	Putative ketoacyl reductase;hypothetical protein;3-phenylpropionate-dihydrodiol/cinnamic acid-dihydrodiol dehydrogenase
173	group_3182	hypothetical protein
173	fprA_2~~~fprA_1~~~fprB~~~fprA	NADPH-ferredoxin reductase FprA;putative ferredoxin/ferredoxin--NADP reductase
173	moaA_2~~~moaA_3~~~moaA_1	GTP 3'8-cyclase
173	ynfF~~~ynfE~~~napA_2	putative dimethyl sulfoxide reductase chain YnfF;hypothetical protein;Putative dimethyl sulfoxide reductase chain YnfE;Periplasmic nitrate reductase
173	fdxA	Ferredoxin-1
173	rhaR_1~~~rhaR_2	HTH-type transcriptional activator RhaR
172	group_3731	hypothetical protein
172	hxuB	Heme/hemopexin transporter protein HuxB
171	group_1146	hypothetical protein
171	group_631	hypothetical protein
170	group_10040	hypothetical protein
170	group_8993	hypothetical protein
170	virB1	hypothetical protein;Type IV secretion system protein virB1
170	group_8076	hypothetical protein
170	group_6727	hypothetical protein
170	group_6725	hypothetical protein
170	group_5865	hypothetical protein
169	xerD_3~~~xerC_5~~~xerC_11	Tyrosine recombinase XerD;Tyrosine recombinase XerC
169	group_3024	hypothetical protein;Chromosome partition protein Smc
168	group_1156	hypothetical protein
168	group_1268	IS21 family transposase ISRso19
167	group_7876	hypothetical protein

167	group_11945	hypothetical protein
166	phnZ	hypothetical protein;2-amino-1-hydroxyethylphosphonate dioxygenase (glycine-forming)
166	group_7067	hypothetical protein
165	group_2944	putative HTH-type transcriptional regulator
165	hcaB_2~~~hcaB_1~~~ydfG_3	3-phenylpropionate-dihydrodiol/cinnamic acid-dihydrodiol dehydrogenase;hypothetical protein;NADP-dependent 3-hydroxy acid dehydrogenase YdfG
164	cutC~~~pflD	Choline trimethylamine-lyase;hypothetical protein;Trans-4-hydroxy-L-proline dehydratase
164	group_1683	hypothetical protein
163	group_1532	hypothetical protein
163	group_1476	hypothetical protein
162	group_2243	hypothetical protein
162	group_2208	hypothetical protein
162	group_2197	hypothetical protein
162	menH_3~~~menH_2~~~menH_4	2-succinyl-6-hydroxy-24-cyclohexadiene-1-carboxylate synthase;hypothetical protein
162	group_2181	hypothetical protein
162	novN_1	hypothetical protein;Decarbamoylnovobiocin carbamoyltransferase
162	group_2163	hypothetical protein;4-hydroxyphenylalkanoate adenylyltransferase
162	group_2162	hypothetical protein
162	group_2141	hypothetical protein
162	group_2139	hypothetical protein
162	valS_2	Valine--tRNA ligase;hypothetical protein
162	metZ~~~metZ_1	O-succinylhomoserine sulfhydrylase;hypothetical protein
162	anoI_2~~~anoI_1	Acyl-homoserine-lactone synthase;hypothetical protein
162	ycaC_2	hypothetical protein;putative hydrolase YcaC
162	anoR_2~~~anoR_1	Transcriptional activator protein AnoR

162	dmlR_13~~~dmlR_19~~~dmlR_12	hypothetical protein;HTH-type transcriptional regulator DmlR
162	sucA_2~~~sucA_3	2-oxoglutarate dehydrogenase E1 component
162	tar_7	hypothetical protein;Methyl-accepting chemotaxis protein II
161	dmlR_18~~~dmlR_5~~~dmlR_10~~~cynR_2~~~dmlR_6~~~dmlR_2	HTH-type transcriptional regulator DmlR;HTH-type transcriptional regulator CynR;hypothetical protein
161	menH_2	hypothetical protein;2-succinyl-6-hydroxy-24-cyclohexadiene-1-carboxylate synthase
160	group_9896	hypothetical protein
160	group_8426	hypothetical protein
159	group_7300	hypothetical protein
159	group_6277	hypothetical protein
158	hchA~~~hchA_1	Protein/nucleic acid deglycase HchA
158	ligJ	hypothetical protein;2-keto-4-carboxy-3-hexenedioate hydratase
157	group_3228	hypothetical protein
157	group_3253	hypothetical protein
157	group_3252	hypothetical protein
157	group_3233	hypothetical protein
157	group_3229	hypothetical protein
157	group_3218	hypothetical protein
156	group_3222	hypothetical protein
156	group_3570	hypothetical protein
155	stp~~~emrY_4~~~emrB_3~~~bmr3~~~stp_2~~~stp_1~~~emrB_4	Multidrug resistance protein Stp;putative multidrug resistance protein EmrY;Multidrug export protein EmrB;Multidrug resistance protein 3;hypothetical protein
155	dhaA~~~dhaA_2~~~rutD_2	Haloalkane dehalogenase;Arylesterase;Putative aminoacrylate hydrolase RutD;hypothetical protein
155	pksJ~~~menE_3~~~menE_4~~~menE_2	Polyketide synthase PksJ;2-succinylbenzoate--CoA ligase;hypothetical protein
154	group_8974	hypothetical protein
154	virB9	Type IV secretion system protein virB9
154	group_8316	hypothetical protein
154	group_8137	Type IV secretion system protein VirB11

154	virB4_2~~~virB4	Type IV secretion system protein virB4
154	group_8077	hypothetical protein
154	group_7530	hypothetical protein
154	group_6687	hypothetical protein
154	virB8	Type IV secretion system protein virB8;hypothetical protein
154	group_6194	hypothetical protein
154	group_5958	hypothetical protein
154	group_5957	hypothetical protein
154	virB5	Type IV secretion system protein virB5
154	group_4321	hypothetical protein
154	group_3953	hypothetical protein
154	group_3435	hypothetical protein
154	group_3434	hypothetical protein
154	group_3431	hypothetical protein
154	group_3134	hypothetical protein
154	group_2512	hypothetical protein
153	group_6121	hypothetical protein
153	hipA_2	Serine/threonine-protein kinase toxin HipA
152	group_10318	HTH-type transcriptional regulator GltC
152	cuyA	L-cysteate sulfo-lyase
152	sotB_2	sugar efflux transporter
151	xerC_1	hypothetical protein;Tyrosine recombinase XerC
151	group_7894	hypothetical protein
151	group_995	hypothetical protein
150	group_1572	hypothetical protein
150	group_1677	hypothetical protein
149	group_1197	hypothetical protein
149	group_1299	hypothetical protein
148	group_11290	hypothetical protein
148	group_10641	hypothetical protein
147	group_7554	hypothetical protein
147	group_2858	hypothetical protein
147	group_7555	hypothetical protein
146	group_10752	hypothetical protein
146	group_10258	hypothetical protein
146	group_9300	hypothetical protein
146	group_9254	hypothetical protein
146	group_7327	hypothetical protein

146	group_7258	hypothetical protein
146	group_7163	hypothetical protein
145	rhsC_3~~~rhsD	Putative deoxyribonuclease RhsC;Protein RhsD;hypothetical protein
145	group_11777	hypothetical protein
145	group_6009	hypothetical protein
145	sasA_15	Adaptive-response sensory-kinase SasA
145	rssB_4	Regulator of RpoS
145	group_4680	hypothetical protein
144	fliI_2~~~fliI_1	Flagellum-specific ATP synthase
144	group_3643	hypothetical protein
144	fliE_2~~~fliE_1	Flagellar hook-basal body complex protein FliE
144	flgC_2~~~flgC_1	Flagellar basal-body rod protein FlgC
144	group_3626	hypothetical protein
144	flgI_2~~~flgI_1	Flagellar P-ring protein
144	flgH_2~~~flgH_1	Flagellar L-ring protein
144	group_3569	hypothetical protein
144	flgF_2~~~flgF_1	Flagellar basal-body rod protein FlgF
144	flgG_2~~~flgG_3~~~flgG_1	Flagellar basal-body rod protein FlgG
144	flhA_1~~~flhA_2	Flagellar biosynthesis protein FlhA
144	flhB_1	Flagellar biosynthetic protein FlhB
144	fliR_1	hypothetical protein;Flagellar biosynthetic protein FliR
144	group_3096	hypothetical protein
143	group_582	4-hydroxyphenylalkanoate adenylyltransferase;Putative fatty-acid--CoA ligase FadD21;Long-chain-fatty-acid--CoA ligase FadD23;Long-chain-fatty-acid--AMP ligase FadD30
143	group_551	hypothetical protein
143	ppsA_2	hypothetical protein;Phthiocerol synthesis polyketide synthase type I PpsA
143	ppsB	hypothetical protein;Phthiocerol/phenolphthiocerol synthesis polyketide synthase type I PpsB
143	group_347	hypothetical protein
143	lgrB~~~ppsE	Linear gramicidin synthase subunit B;hypothetical protein;Phthiocerol

		synthesis polyketide synthase type I PpsE
143	pikAV~~~pikAV_2	Thioesterase PikA5
142	group_5844	hypothetical protein
142	group_286	hypothetical protein
141	group_2103	hypothetical protein
141	group_7675	hypothetical protein
140	group_10366	hypothetical protein
140	group_9473	hypothetical protein
139	group_2516	hypothetical protein
139	group_2364	hypothetical protein
138	group_11617	hypothetical protein
138	group_11895	hypothetical protein
137	group_2104	hypothetical protein
137	group_2069	hypothetical protein
136	mshA_2~~~mshA_3~~~mshA_4	D-inositol-3-phosphate glycosyltransferase;hypothetical protein
136	cya~~~cya_2~~~cya_1	hypothetical protein;Bifunctional hemolysin/adenylate cyclase
135	group_1811	hypothetical protein
135	group_1521	hypothetical protein
135	group_1029	hypothetical protein
134	group_11638	hypothetical protein
134	group_10515	hypothetical protein
133	group_9987	hypothetical protein
133	group_9672	hypothetical protein
133	group_8880	hypothetical protein
133	group_6566	hypothetical protein
133	ttdB~~~ttdB_1~~~ttdB_2	L(+)-tartrate dehydratase subunit beta
132	mmgC_5	hypothetical protein;Acyl-CoA dehydrogenase fadE12;Acyl-CoA dehydrogenase
132	group_3051	hypothetical protein
132	yofA	HTH-type transcriptional regulator YofA;hypothetical protein
131	group_2310	hypothetical protein
131	group_3410	hypothetical protein
130	group_7491	hypothetical protein
130	group_1844	hypothetical protein
129	sutR_3~~~sutR_4	hypothetical protein;HTH-type transcriptional regulator SutR
129	rhtB_2~~~rhtB_3	Homoserine/homoserine lactone efflux protein

128	group_4825	hypothetical protein
128	group_1932	hypothetical protein
128	group_7495	hypothetical protein
127	group_9159	hypothetical protein
127	group_8998	Protein kinase YegI
127	group_6994	hypothetical protein
127	group_6733	Serine/threonine-protein phosphatase 3
126	group_3385	hypothetical protein
126	group_11901	hypothetical protein
126	group_285	hypothetical protein
126	group_58	hypothetical protein
125	group_11976	hypothetical protein
125	group_11735	hypothetical protein
125	group_11397	hypothetical protein
125	soj_2~~~soj_1	Chromosome-partitioning ATPase Soj
125	group_10851	hypothetical protein
125	group_11975	hypothetical protein
125	group_11736	hypothetical protein
125	group_122	hypothetical protein
125	group_746	hypothetical protein
125	group_2810	hypothetical protein
125	group_2809	hypothetical protein
125	group_1750	hypothetical protein
125	group_1445	hypothetical protein
125	group_1475	hypothetical protein
124	yqcF	Antitoxin YqcF
124	group_9151	hypothetical protein
124	group_5443	hypothetical protein
123	cnrA_2~~~cnrA_1~~~cnrA	Nickel and cobalt resistance protein CnrA
123	cnrB	Nickel and cobalt resistance protein CnrB
123	cnrC	Nickel and cobalt resistance protein CnrC
123	cnrH~~~cnrH_1	RNA polymerase sigma factor CnrH
123	cnrR	Nickel and cobalt resistance protein CnrR
123	cnrY	Nickel and cobalt resistance protein CnrY
122	group_5040	hypothetical protein
122	group_7201	hypothetical protein
122	group_2735	hypothetical protein

121	group_2338	hypothetical protein
121	group_2989	hypothetical protein
121	group_2504	hypothetical protein
120	group_371	hypothetical protein
120	group_302	hypothetical protein
119	group_8127	hypothetical protein
119	group_3913	hypothetical protein
118	group_10292	hypothetical protein
118	group_9375	hypothetical protein
117	group_7697	hypothetical protein
117	group_4008	hypothetical protein
116	group_3337	hypothetical protein
116	group_3053	hypothetical protein
116	mdtA_9~~~mdtA_4~~~mdtA_7~~~mdtA_3~~~cusB~~~mdtA_2	Multidrug resistance protein MdtA;Cation efflux system protein CusB
116	cusA~~~cusA_1~~~cusA_2	Cation efflux system protein CusA
116	group_2790	hypothetical protein
115	creB	Transcriptional regulatory protein CreB
115	creC	Sensor protein CreC
115	creD	Inner membrane protein CreD;hypothetical protein
114	emrB_1~~~stp~~~stp_1~~~emrB_3~~~mdtD_1~~~mdtD_2~~~mdtD~~~emrB_5	hypothetical protein;Multidrug export protein EmrB;Multidrug resistance protein Stp;Putative multidrug resistance protein MdtD
114	group_3265	hypothetical protein
113	group_2884	hypothetical protein
113	group_2829	hypothetical protein
112	group_11001	hypothetical protein
112	tar_2~~~tar_1	Methyl-accepting chemotaxis protein II
112	group_929	hypothetical protein
111	oprJ~~~oprM_7	Outer membrane protein OprJ;Outer membrane protein OprM
111	hcaB_3~~~hcaB_1~~~hcaB	3-phenylpropionate-dihydrodiol/cinnamic acid-dihydrodiol dehydrogenase
110	group_3036	hypothetical protein
110	group_3746	IS3 family transposase ISButh1
109	xerC_5~~~xerC_10	Tyrosine recombinase XerC
109	group_6604	hypothetical protein
109	group_8992	hypothetical protein

109	group_8918	hypothetical protein
109	group_10003	hypothetical protein
109	ttuB_1~~~ttuB_2	Putative tartrate transporter
109	group_8214	hypothetical protein
109	group_6154	hypothetical protein
108	group_2946	hypothetical protein;RutC family protein YjgH
108	lrpC~~~lrpC_3~~~lrpC_2	hypothetical protein;HTH-type transcriptional regulator LrpC
108	hpd_2~~~lly~~~hpd_1	4-hydroxyphenylpyruvate dioxygenase;hypothetical protein
107	pgrR_5~~~pgrR_1	HTH-type transcriptional regulator PgrR
107	group_3364	hypothetical protein
107	group_3342	hypothetical protein
107	ycjY_1~~~ycjY	putative protein YcjY;hypothetical protein
107	yvgN	putative oxidoreductase/MSMEI_2346;hypothetical protein;Glyoxal reductase
107	efpA	putative MFS-type transporter EfpA;hypothetical protein
106	group_1558	hypothetical protein
106	group_1559	hypothetical protein
105	group_821	hypothetical protein
105	group_532	hypothetical protein
105	group_10664	hypothetical protein
105	group_9516	hypothetical protein
105	group_7666	hypothetical protein
105	group_4996	hypothetical protein
105	group_3326	hypothetical protein
104	group_5833	hypothetical protein
104	group_1158	hypothetical protein
103	group_1464	hypothetical protein
103	group_1429	hypothetical protein
102	group_3505	hypothetical protein
102	group_3291	hypothetical protein
102	group_835	hypothetical protein
101	group_9433	hypothetical protein
101	group_6578	hypothetical protein
100	group_3289	hypothetical protein
100	group_2567	hypothetical protein
100	group_2573	hypothetical protein
99	group_89	hypothetical protein

99	group_1168	hypothetical protein
98	group_7890	hypothetical protein
98	group_2161	hypothetical protein
98	group_2160	hypothetical protein
98	group_2144	hypothetical protein
98	clpP_1~~~clpP_2	ATP-dependent Clp protease proteolytic subunit
98	group_2109	hypothetical protein
98	group_2077	hypothetical protein
97	group_3733	hypothetical protein
97	xerC_1~~~xerC_4	Tyrosine recombinase XerC;hypothetical protein
96	group_8220	hypothetical protein
96	group_3204	hypothetical protein
96	gltI_6~~~gltI_7	Glutamate/aspartate import solute-binding protein
95	group_1180	hypothetical protein
95	group_1133	hypothetical protein
94	group_9545	hypothetical protein
94	group_6870	hypothetical protein
93	group_727	hypothetical protein
93	group_10458	hypothetical protein;Type I restriction enzyme EcoKI M protein
92	group_10414	hypothetical protein
92	higB~~~higB_2	Endoribonuclease HigB
92	group_7682	hypothetical protein
92	group_2799	hypothetical protein
92	group_1712	hypothetical protein
92	group_1631	hypothetical protein
92	group_1161	hypothetical protein
92	group_1274	hypothetical protein
91	ccr~~~pikAII	Crotonyl-CoA reductase;Narbonolide/10-deoxymethynolide synthase PikA2 modules 3 and 4
91	group_4783	hypothetical protein
91	fabG_4~~~fabG_2	hypothetical protein;3-oxoacyl-[acyl-carrier-protein] reductase FabG
91	group_9456	hypothetical protein
90	group_14014	hypothetical protein
90	group_14012	hypothetical protein
89	group_2362	hypothetical protein
89	group_890	hypothetical protein

89	group_10120	hypothetical protein
89	group_10268	hypothetical protein
89	group_8771	ATP-dependent zinc metalloprotease FtsH
89	group_3809	hypothetical protein
89	group_7246	hypothetical protein
88	group_2824	hypothetical protein
88	group_1170	hypothetical protein
87	group_9938	hypothetical protein
87	group_838	hypothetical protein
86	group_508	hypothetical protein
86	group_552	hypothetical protein
86	group_11904	hypothetical protein
86	group_11633	hypothetical protein
86	epsE~~~hxcR_2	hypothetical protein;Type II secretion system protein E;putative type II secretion system protein HxcR
86	epsG~~~xcpT_3~~~xcpT_1~~~xcpT_2	Type II secretion system protein G;hypothetical protein
86	group_53	hypothetical protein
86	sctC_6~~~sctC_7~~~sctC_4~~~sctC_5~~~sctC_8~~~sctC_2~~~sctC_1	Type 3 secretion system secretin;hypothetical protein
86	xcpT_3~~~xcpT_5~~~xcpT_2~~~xcpT_4	Type II secretion system protein G;hypothetical protein
86	epsF_5~~~epsF_4~~~epsF_3	Type II secretion system protein F
85	dnaK_2~~~dnaK_1~~~dnaK_3~~~dnaK_4	Chaperone protein DnaK
85	group_2615	hypothetical protein
85	group_10361	hypothetical protein
85	groS	10 kDa chaperonin
85	groL_2~~~groL_3	60 kDa chaperonin
85	cynT	Carbonic anhydrase 1
85	group_4330	hypothetical protein
85	group_2678	hypothetical protein
85	phrA_2~~~phrA	Deoxyribodipyrimidine photo-lyase
85	group_10900	hypothetical protein
85	group_10362	hypothetical protein
85	group_3335	hypothetical protein
85	group_3306	hypothetical protein
84	group_1833	hypothetical protein
84	group_1593	hypothetical protein
83	group_2105	hypothetical protein
83	group_2147	hypothetical protein

82	casD	CRISPR system Cascade subunit CasD;hypothetical protein
82	group_11499	hypothetical protein
82	casC	CRISPR system Cascade subunit CasC
82	ygbT~~~ygbF~~~cas1	CRISPR-associated endonuclease Cas1;CRISPR-associated endoribonuclease Cas2
82	group_10999	hypothetical protein
82	casA	hypothetical protein;CRISPR system Cascade subunit CasA
82	ygcB	CRISPR-associated endonuclease/helicase Cas3;hypothetical protein
81	group_9073	hypothetical protein
81	group_6894	hypothetical protein
80	group_2629	IS3 family transposase IS1416;IS3 family transposase IS401
80	group_2620	IS3 family transposase ISBt3;IS3 family transposase IS401
79	group_7003	hypothetical protein
79	group_9162	hypothetical protein
78	group_9644	hypothetical protein
78	group_10520	hypothetical protein
77	group_1839	hypothetical protein
77	group_2837	DNA-binding protein Bv3F
76	group_1836	hypothetical protein
76	group_3729	hypothetical protein
76	group_7310	hypothetical protein
75	group_2855	hypothetical protein
75	group_1076	hypothetical protein
74	group_10035	hypothetical protein
74	group_10498	hypothetical protein
74	fabI_1~~~fabI_2	Enoyl-[acyl-carrier-protein] reductase [NADH] FabI
74	ackA	Acetate kinase
74	pta	Phosphate acetyltransferase
74	group_3756	hypothetical protein
74	group_3156	hypothetical protein
74	group_9584	hypothetical protein
74	group_2115	hypothetical protein
74	group_11506	hypothetical protein
74	group_10199	hypothetical protein
74	group_9022	hypothetical protein
74	yecD~~~yecD_2~~~yecD_1	Isochorismatase family protein YecD

74	nopX	hypothetical protein;Nodulation outer protein X
74	group_1766	hypothetical protein
74	yycB~~~nimT	putative transporter YycB;hypothetical protein;2-nitroimidazole transporter
74	group_9765	hypothetical protein
74	group_7240	hypothetical protein
74	group_7239	hypothetical protein
74	group_2706	hypothetical protein
74	cmpR_1~~~cmpR_3~~~cmpR_2~~~hdfR_1	HTH-type transcriptional activator CmpR;hypothetical protein;HTH-type transcriptional regulator HdfR
74	group_7567	Outer membrane protein TolC
74	group_691	hypothetical protein
74	group_290	hypothetical protein
74	group_262	hypothetical protein
74	group_81	putative glycosyltransferase
74	group_11917	hypothetical protein
74	sfnC	putative FMNH2-dependent monooxygenase SfnC;hypothetical protein
74	group_11049	hypothetical protein
74	group_11005	hypothetical protein
74	argE_2~~~argE_3	Acetylornithine deacetylase
74	group_10386	queuosine precursor transporter
74	qseB_2~~~qseB_3	Transcriptional regulatory protein QseB
74	phaB_1~~~phaB_2	Acetoacetyl-CoA reductase
74	group_9923	hypothetical protein
74	davD_2~~~davD_1	Glutarate-semialdehyde dehydrogenase
74	sasA_3~~~sasA_16~~~sasA_5~~~sasA_12	Adaptive-response sensory-kinase SasA
74	group_9301	hypothetical protein
74	iolX	scyllo-inositol 2-dehydrogenase (NAD(+))
74	phzS	5-methylphenazine-1-carboxylate 1-monooxygenase
74	phnM	Alpha-D-ribose 1-methylphosphonate 5-triphosphate diphosphatase
74	group_5101	hypothetical protein
74	oxyR_2	Hydrogen peroxide-inducible genes activator;hypothetical protein

74	group_2196	hypothetical protein
74	group_2100	hypothetical protein
74	sacX_2~~~sacX_1~~~sacX	Negative regulator of SacY activity;hypothetical protein
74	scrB	hypothetical protein;Sucrose-6-phosphate hydrolase
74	lamB	Maltoporin;hypothetical protein
74	group_1051	hypothetical protein
74	scrK	Fructokinase
74	benC	Benzoate 12-dioxygenase electron transfer component;hypothetical protein
74	baiA~~~fabG2~~~ucpA_2~~~ucpA_1~~~baiC_1	hypothetical protein;3-alpha-hydroxycholanate dehydrogenase (NADP(+));putative oxidoreductase;Oxidoreductase UcpA;Dihydroanticapsin 7-dehydrogenase
74	baiA~~~fabG_5~~~ptlF~~~hcaB_1~~~baiA_2	3-alpha-hydroxycholanate dehydrogenase (NADP(+));3-oxoacyl-[acyl-carrier-protein] reductase FabG;1-deoxy-11-beta-hydroxypentalenate dehydrogenase;hypothetical protein;3-phenylpropionate-dihydrodiol/cinnamic acid-dihydrodiol dehydrogenase
74	araG_2~~~araG	hypothetical protein;Arabinose import ATP-binding protein AraG
74	cusB_3~~~cusB~~~cusB_2	Cation efflux system protein CusB
74	xerC_4~~~xerC_7~~~xerC_1~~~xerC_5	Tyrosine recombinase XerC;hypothetical protein
74	group_2201	hypothetical protein
74	ccpA_3~~~purR_1~~~ccpA_1	hypothetical protein;IS5 family transposase IS1421;Catabolite control protein A;HTH-type transcriptional repressor PurR
74	group_1412	hypothetical protein
74	group_576	hypothetical protein
74	group_533	hypothetical protein
74	group_487	hypothetical protein
74	group_435	hypothetical protein
74	czcA_6~~~czcA_4	Cobalt-zinc-cadmium resistance protein CzcA
74	group_219	hypothetical protein

74	bamA_2~~~bamA_1	hypothetical protein;Outer membrane protein assembly factor BamA
74	group_65	hypothetical protein
74	group_11979	Ribose import ATP-binding protein RbsA
74	iolE	Inososе dehydratase
74	thpA	D-threitol-binding protein
74	iolB	5-deoxy-glucuronate isomerase
74	group_11405	hypothetical protein
74	group_11253	hypothetical protein
74	group_10866	hypothetical protein
74	idhA~~~iolG_2	Inositol 2-dehydrogenase;Myo-inositol 2-dehydrogenase
74	doxA	Naphthalene 12-dioxygenase system ferredoxin component
74	group_9984	hypothetical protein
74	cbdA	2-halobenzoate 12-dioxygenase large subunit
74	murR_2~~~murR_1	HTH-type transcriptional regulator MurR
74	group_8864	hypothetical protein
74	catA	Catechol 12-dioxygenase
74	cbdB	2-halobenzoate 12-dioxygenase small subunit
74	andAa	Anthranilate 12-dioxygenase system ferredoxin--NAD(+) reductase component
74	catB	Muconate cycloisomerase 1
74	catC	Muconolactone Delta-isomerase
74	benM~~~benM_1	HTH-type transcriptional regulator BenM
74	argC_3~~~argC_2~~~argC_1	N-acetyl-gamma-glutamyl-phosphate reductase
74	group_1972	hypothetical protein
74	group_934	hypothetical protein
74	iolC	5-dehydro-2-deoxygluconokinase
74	rbsC_1~~~rbsC_2	hypothetical protein;Ribose import permease protein RbsC
74	iolD	3D-(35/4)-trihydroxycyclohexane-12-dione hydrolase
74	group_8239	hypothetical protein
74	group_2619	hypothetical protein
74	group_2583	hypothetical protein
74	group_2082	hypothetical protein

74	glnH_2~~~mltF_2~~~glnH_3	ABC transporter glutamine-binding protein GlnH;Membrane-bound lytic murein transglycosylase F;hypothetical protein
74	glnM_2	hypothetical protein;putative glutamine ABC transporter permease protein GlnM
74	occM	Octopine transport system permease protein OccM
74	btuD_9~~~glnQ_5~~~tcyC_2~~~tcyC_1	Vitamin B12 import ATP-binding protein BtuD;Glutamine transport ATP-binding protein GlnQ;hypothetical protein;L-cystine import ATP-binding protein TcyC
74	crp~~~queE_2~~~queE_1	CRP-like cAMP-activated global transcriptional regulator;7-carboxy-7-deazaguanine synthase
74	group_3624	hypothetical protein
74	nrdD	hypothetical protein;Anaerobic ribonucleoside-triphosphate reductase
74	narT~~~nasA_2~~~nasA_1	hypothetical protein;putative nitrate transporter NarT;Nitrate transporter
74	narG	hypothetical protein;Respiratory nitrate reductase 1 alpha chain
74	degU	Transcriptional regulatory protein DegU
74	yihG	putative acyltransferase YihG
74	group_3126	hypothetical protein
74	argP_2~~~pgrR_3~~~pgrR_4~~~pgrR_2~~~pgrR_1~~~pgrR_5	HTH-type transcriptional regulator ArgP;HTH-type transcriptional regulator PgrR
74	group_11160	hypothetical protein
74	yhbU	putative protease YhbU
74	group_3650	hypothetical protein
74	group_3642	hypothetical protein
74	group_3640	hypothetical protein
74	narK	Nitrate/nitrite transporter NarK
74	group_3325	hypothetical protein;putative oxidoreductase
74	group_3119	hypothetical protein
74	group_529	putative protein;hypothetical protein
74	group_503	putative protein;hypothetical protein

74	group_467	hypothetical protein
74	puuC~~~puuC_1~~~puuC_2	NADP/NAD-dependent aldehyde dehydrogenase PuuC;hypothetical protein
74	group_10101	hypothetical protein
74	abo_1~~~abo_2	4-methylaminobutanoate oxidase (formaldehyde-forming)
74	hcnB_2~~~hcnB	Hydrogen cyanide synthase subunit HcnB
74	ghrA_2~~~ghrA_1	Glyoxylate/hydroxypyruvate reductase A
74	group_6891	hypothetical protein
74	spuC_2~~~spuC_1~~~spuC	Putrescine--pyruvate aminotransferase
74	lrp_5	Leucine-responsive regulatory protein
74	narH	Respiratory nitrate reductase 1 beta chain
74	group_3544	hypothetical protein
74	narI	Respiratory nitrate reductase 1 gamma chain
74	mobB_2	Molybdopterin-guanine dinucleotide biosynthesis adapter protein
74	narX	Nitrate/nitrite sensor protein NarX
74	group_953	hypothetical protein
74	group_623	hypothetical protein
74	hipO_3~~~hipO_1	hypothetical protein;Hippurate hydrolase
74	group_565	hypothetical protein
74	group_228	hypothetical protein
74	group_10541	hypothetical protein
74	glpE_3~~~glpE_2~~~glpE_1	Thiosulfate sulfurtransferase GlpE
74	metC_2~~~metC_3~~~metC_1	Cystathionine beta-lyase
74	group_3028	hypothetical protein
74	sgrR	hypothetical protein;HTH-type transcriptional regulator SgrR
74	puuB_2~~~puuB_3	Gamma-glutamylputrescine oxidoreductase
74	group_11465	hypothetical protein
74	proP_3~~~proP_1	Proline/betaine transporter;hypothetical protein
74	group_1961	hypothetical protein
74	gsiA_2	Glutathione import ATP-binding protein GsiA
74	group_251	hypothetical protein

74	gsiC_3~~~gsiC_2	Glutathione transport system permease protein GsiC
74	infA_2~~~infA_3	Translation initiation factor IF-1
74	aspC_3~~~aspC_2	Aspartate aminotransferase;hypothetical protein
74	bacC_2~~~gdh	Dihydroanticapsin 7-dehydrogenase;Galactitol 2-dehydrogenase
74	cefD_2~~~cefD	Isopenicillin N epimerase
74	group_3328	hypothetical protein
74	TDO2_1~~~TDO2_2~~~TDO2	Tryptophan 23-dioxygenase
74	group_3105	hypothetical protein
74	group_183	hypothetical protein
74	group_2701	hypothetical protein
74	glpR_1~~~glcR	Glycerol-3-phosphate regulon repressor;hypothetical protein;HTH-type transcriptional repressor GlcR
74	fabG1~~~fabG_8	3-oxoacyl-[acyl-carrier-protein] reductase FabG1;3-oxoacyl-[acyl-carrier-protein] reductase FabG
74	psd_2~~~psd_1	Phosphatidylserine decarboxylase proenzyme
74	group_602	hypothetical protein
74	cdhR~~~cdhR_2	HTH-type transcriptional regulator CdhR
74	group_11852	hypothetical protein
74	group_3159	hypothetical protein
74	group_10883	hypothetical protein
74	dmlR_17~~~dmlR_4~~~dmlR_6~~~dmlR_3~~~dmlR_9	HTH-type transcriptional regulator DmlR
74	jefA_2~~~stp~~~jefA	Drug efflux pump JefA;Multidrug resistance protein Stp
74	group_9649	hypothetical protein
74	group_7045	hypothetical protein
74	group_1446	hypothetical protein
74	group_597	hypothetical protein
74	group_2157	hypothetical protein
74	group_3535	hypothetical protein
74	tam_1~~~tam_2	Trans-aconitate 2-methyltransferase
74	group_3421	hypothetical protein
74	group_1624	hypothetical protein
74	group_1080	hypothetical protein
74	group_2780	hypothetical protein
74	group_1711	hypothetical protein

74	group_1668	hypothetical protein
74	group_1343	hypothetical protein
74	group_10522	hypothetical protein
74	amiD	N-acetylmuramoyl-L-alanine amidase AmiD;hypothetical protein
74	group_2848	hypothetical protein
74	cycA	D-serine/D-alanine/glycine transporter;hypothetical protein
74	group_2690	hypothetical protein
74	group_1399	hypothetical protein
74	group_10725	hypothetical protein
74	gcvA_2~~~gcvA_1~~~gcvA_4~~~gcvA_3	Glycine cleavage system transcriptional activator
74	group_9148	hypothetical protein
74	argO_2~~~argO_1	Arginine exporter protein ArgO
74	group_8039	Inner membrane protein YbiR
74	group_7417	hypothetical protein
74	yafQ	hypothetical protein;mRNA interferase toxin YafQ
74	dapA_1~~~dapA_5	4-hydroxy-tetrahydrodipicolinate synthase
74	adh1~~~adhB	Long-chain-alcohol dehydrogenase 1;Alcohol dehydrogenase 2
74	pdxA_2~~~pdxA_1	4-hydroxythreonine-4-phosphate dehydrogenase
74	putP_2~~~putP_1	High-affinity proline transporter PutP;hypothetical protein
74	group_2713	hypothetical protein
74	kimA	hypothetical protein;Potassium transporter KimA
74	group_2175	hypothetical protein
74	group_12038	hypothetical protein
74	group_11460	hypothetical protein
74	tsaR~~~argP_2~~~cynR_4	HTH-type transcriptional regulator TsaR;HTH-type transcriptional regulator ArgP;hypothetical protein;HTH-type transcriptional regulator CynR
74	group_10661	hypothetical protein
74	group_10185	hypothetical protein
74	group_10039	hypothetical protein
74	dapA_2~~~dapA_4~~~dapA_1	4-hydroxy-tetrahydrodipicolinate synthase
74	group_8327	hypothetical protein

74	cmpR_3~~~cmpR_2	hypothetical protein;HTH-type transcriptional activator CmpR
74	nox~~~drgA~~~rutE_2	NADH dehydrogenase;Protein DrgA;hypothetical protein;malonic semialdehyde reductase RutE
74	ilvE_2	Branched-chain-amino-acid aminotransferase
74	dmlR_12~~~dmlR_11	HTH-type transcriptional regulator DmlR;hypothetical protein
74	group_7070	hypothetical protein
74	group_6217	hypothetical protein
74	kynB_1~~~kynB_2	Kynurenine formamidase
74	sasA_14~~~sasA_13~~~sasA_12	Adaptive-response sensory-kinase SasA;hypothetical protein
74	group_9931	hypothetical protein
74	group_9273	hypothetical protein
74	group_9195	hypothetical protein
74	group_9025	hypothetical protein
74	group_7204	hypothetical protein
74	group_7199	hypothetical protein
74	group_7083	hypothetical protein
74	group_6779	hypothetical protein
74	group_6296	hypothetical protein
74	group_5907	hypothetical protein
74	group_3324	hypothetical protein
74	group_2146	hypothetical protein
74	group_2133	hypothetical protein
74	lodB	Putative FAD-dependent oxidoreductase LodB
	rhsB_1~~~rhsA_2~~~rhsB_2~~~rhsA_1~~~	putative deoxyribonuclease RhsB;putative deoxyribonuclease
74	rhsB~~~rhsA	RhsA;hypothetical protein
74	group_6334	hypothetical protein
74	group_1577	hypothetical protein
74	group_11656	hypothetical protein
74	group_8865	hypothetical protein
74	tar_2	Methyl-accepting chemotaxis protein II;hypothetical protein
74	group_6558	hypothetical protein
74	citB	Aconitate/2-methylaconitate hydratase
74	mdtC_4~~~mdtC_5~~~mdtC_2	Multidrug resistance protein MdtC
74	mdtA_7~~~mdtA_3~~~ttgA	Multidrug resistance protein MdtA;Toluene efflux pump periplasmic linker protein TtgA

74	ddl_2	D-alanine--D-alanine ligase
74	group_1772	hypothetical protein
74	group_1263	hypothetical protein
74	group_12041	hypothetical protein
74	group_12000	hypothetical protein
74	hmuV~~~btuD_5	Hemin import ATP-binding protein HmuV;hypothetical protein;Vitamin B12 import ATP-binding protein BtuD
74	prpC_2~~~prpC_1	2-methylcitrate synthase
74	group_9849	hypothetical protein
74	group_9196	hypothetical protein
74	group_7167	hypothetical protein
74	group_6709	2-amino-3-ketobutyrate coenzyme A ligase
74	group_6068	putative epimerase/dehydratase
74	sutR_2	HTH-type transcriptional regulator SutR
74	czcA_1~~~czcA_2~~~czcA_3	Cobalt-zinc-cadmium resistance protein CzcA
74	mdtA_4~~~mdtA_3~~~cusB_1	Multidrug resistance protein MdtA;Cation efflux system protein CusB
74	oprM_3~~~oprM_2	Outer membrane protein OprM
74	group_1290	hypothetical protein
74	group_1137	hypothetical protein
74	group_753	hypothetical protein
74	group_11849	hypothetical protein
74	adh_1~~~adhT~~~adh_3	Alcohol dehydrogenase;hypothetical protein
74	group_10671	hypothetical protein
74	dmlR_2~~~dmlR_4~~~dmlR_9~~~dmlR_8	HTH-type transcriptional regulator DmlR;hypothetical protein
74	baiA~~~baiA_1~~~baiA_3	3-alpha-hydroxycholanate dehydrogenase (NADP(+))
74	group_8119	hypothetical protein
74	group_8093	hypothetical protein
74	group_7608	hypothetical protein
74	group_5963	hypothetical protein
74	group_5926	hypothetical protein
74	adhC2	hypothetical protein;NADP-dependent alcohol dehydrogenase C 2

74	moeZ	putative adenylyltransferase/sulfurtransferase MoeZ
74	mec	CysO-cysteine peptidase
74	cysO_1~~~cysO	Sulfur carrier protein CysO
74	frc_2	Formyl-CoA:oxalate CoA-transferase
74	trsA	Triostin synthetase I
74	bbsG	(R)-benzylsuccinyl-CoA dehydrogenase
74	group_3587	hypothetical protein
74	group_3586	hypothetical protein
74	group_3513	hypothetical protein
74	btuB_4~~~btuB_2	Vitamin B12 transporter BtuB
74	group_2729	hypothetical protein
74	creB~~~rssB_2	Transcriptional regulatory protein CreB;Regulator of RpoS
74	group_1148	hypothetical protein
74	group_873	hypothetical protein
74	group_11995	hypothetical protein
74	mdtA_2~~~mdtA_7	hypothetical protein;Multidrug resistance protein MdtA
74	group_11610	hypothetical protein
74	group_11448	Cobalt-zinc-cadmium resistance protein CzcA
74	group_9467	hypothetical protein
74	cpdA_3~~~cpdA_2~~~cpdA_4~~~cpdA_1	3'5'-cyclic adenosine monophosphate phosphodiesterase CpdA
74	group_1746	hypothetical protein
74	group_1472	hypothetical protein
74	dmlR_9~~~dmlR_19~~~dmlR_14~~~dmlR_4	HTH-type transcriptional regulator DmlR
74	yhcG_1~~~yhcG~~~yhcG_2	Putative nuclease YhcG;hypothetical protein
74	ttuD_2~~~ttuD_1	Putative hydroxypyruvate reductase
74	nemA_2~~~nemA_1	N-ethylmaleimide reductase
74	group_8082	hypothetical protein
74	group_7545	hypothetical protein
74	group_7535	hypothetical protein
74	group_7161	hypothetical protein
74	azoR	FMN-dependent NADH-azoreductase
74	group_6287	hypothetical protein
74	group_6030	hypothetical protein
74	group_5401	hypothetical protein

74	group_3648	hypothetical protein putative methyltransferase
74	ycgJ	YcgJ;hypothetical protein
74	group_11955	hypothetical protein
74	group_10271	hypothetical protein
74	group_6525	hypothetical protein
74	group_11080	hypothetical protein
74	group_9925	hypothetical protein
74	group_9511	hypothetical protein
74	dhmA~~~oleB	Haloalkane dehalogenase;Cis-3-alkyl-4-alkyloxetan-2-one decarboxylase;hypothetical protein
74	yjiB	Putative cytochrome P450 YjiB;hypothetical protein
74	ptlI	Putative cytochrome P450 120;Pentalenene oxygenase;hypothetical protein
74	btuC~~~hmuU~~~btuC_2	Vitamin B12 import system permease protein BtuC;Hemin transport system permease protein HmuU
74	group_9763	hypothetical protein
74	pcpR_6~~~pcpR_5~~~pcpR_7	PCP degradation transcriptional activation protein
74	gltR_2	HTH-type transcriptional regulator GltR
74	group_6206	hypothetical protein
74	pksR~~~pksM~~~pksM_3~~~pksN~~~pksL_1~~~fadA_2~~~pksL~~~fadA_3	hypothetical protein;Polyketide synthase PksR;Polyketide synthase PksM;Polyketide synthase PksN;Polyketide synthase PksL;3-ketoacyl-CoA thiolase
74	acpP_4~~~acpP_3~~~acpP_1~~~pksL_1~~~acpP_2	Acyl carrier protein;hypothetical protein;Polyketide synthase PksL
74	group_513	hypothetical protein
74	group_222	hypothetical protein
74	fadI	3-ketoacyl-CoA thiolase FadI;3-ketoacyl-CoA thiolase
74	ku	hypothetical protein;Non-homologous end joining protein Ku
74	pksL~~~pksL_2~~~pksN~~~pksJ~~~pksR	Polyketide synthase PksL;Polyketide synthase PksN;Polyketide synthase PksJ;Polyketide synthase PksR
74	baeD	hypothetical protein;Polyketide biosynthesis acyltransferase BaeD
74	group_12009	hypothetical protein

74	higB	hypothetical protein;mRNA interferase HigB
74	chrA1_2~~~chrA1_1	Chromate transport protein;hypothetical protein
74	group_6819	hypothetical protein
74	group_5797	hypothetical protein
74	group_1823	hypothetical protein
74	group_10270	hypothetical protein
74	group_8166	hypothetical protein
74	group_7990	hypothetical protein
74	group_7925	hypothetical protein
74	ligD_2~~~ligD_1~~~ligD	Multifunctional non-homologous end joining protein LigD
74	group_3739	hypothetical protein
74	group_2178	hypothetical protein
74	group_1479	hypothetical protein
74	group_11857	hypothetical protein
74	group_10446	hypothetical protein
74	rtcB_1	hypothetical protein;RNA-splicing ligase RtcB
74	group_8397	Cardiolipin synthase;hypothetical protein;Cardiolipin synthase B
74	pqiB_2~~~pqiB_3	Intermembrane transport protein PqiB
74	group_7824	hypothetical protein
74	group_7561	hypothetical protein
74	group_5796	hypothetical protein
74	group_4421	hypothetical protein
74	group_3873	hypothetical protein
74	cmoM~~~cmoM_1	tRNA 5-carboxymethoxyuridine methyltransferase;hypothetical protein
74	COQ5_3~~~COQ5_2~~~COQ5_1	2-methoxy-6-polyprenyl-14-benzoquinol methylase mitochondrial
74	group_2937	hypothetical protein
74	group_2348	hypothetical protein
74	fliY_1~~~fliY_3	L-cystine-binding protein FliY
74	glnM_1~~~glnM_3~~~glnM_2	putative glutamine ABC transporter permease protein GlnM
74	metXA_2~~~metXA_1	Homoserine O-acetyltransferase
74	group_9914	hypothetical protein
74	group_9468	hypothetical protein
74	group_9444	hypothetical protein
74	group_8991	hypothetical protein

74	group_8028	hypothetical protein
74	pntB_2	NAD(P) transhydrogenase subunit beta
74	metR_2~~~argP_2	hypothetical protein;HTH-type transcriptional regulator MetR;HTH-type transcriptional regulator ArgP
74	group_5951	hypothetical protein
74	group_401	hypothetical protein
74	group_47	hypothetical protein
74	group_9244	hypothetical protein
74	group_6309	hypothetical protein
74	group_5891	hypothetical protein
74	group_1628	hypothetical protein
74	group_663	hypothetical protein
74	group_9858	hypothetical protein
74	group_9101	hypothetical protein
74	group_6587	hypothetical protein
74	rhaS_3~~~rhaS_2~~~rhaS_4	HTH-type transcriptional activator RhaS;hypothetical protein
74	group_9878	hypothetical protein
74	group_9528	hypothetical protein
74	group_8100	hypothetical protein
74	group_7789	hypothetical protein
74	group_6902	hypothetical protein
74	group_4957	hypothetical protein
74	group_2557	hypothetical protein;IS5 family transposase ISAzo11
74	licC	hypothetical protein;Lichenan permease IIC component
74	group_10697	hypothetical protein
74	group_3282	hypothetical protein
74	group_3003	hypothetical protein
74	sadH~~~hcaB_2	hypothetical protein;Putative oxidoreductase SadH;3-phenylpropionate-dihydrodiol/cinnamic acid-dihydrodiol dehydrogenase;Baeyer-Villiger monooxygenase
74	group_9830	hypothetical protein
74	group_8343	hypothetical protein
74	group_2835	hypothetical protein
74	sigJ	ECF RNA polymerase sigma factor SigJ
74	group_10972	hypothetical protein
74	group_10847	hypothetical protein

74	group_9416	hypothetical protein
74	group_9393	hypothetical protein
74	group_9387	hypothetical protein
74	group_8730	hypothetical protein
74	group_8167	hypothetical protein
74	group_7689	hypothetical protein
74	group_7268	hypothetical protein
74	group_7262	hypothetical protein
74	decR_4~~~lrp_4	DNA-binding transcriptional activator DecR;Leucine-responsive regulatory protein
74	dhaA~~~dhaA_1	Haloalkane dehalogenase
74	group_6423	hypothetical protein
74	group_6293	hypothetical protein;Leucine efflux protein
74	phnX	Phosphonoacetaldehyde hydrolase
74	group_5814	hypothetical protein
74	hdfR_3~~~hdfR_4	HTH-type transcriptional regulator HdfR
74	group_4494	hypothetical protein
74	group_4115	hypothetical protein
74	group_1727	hypothetical protein
74	group_10341	hypothetical protein
74	group_10299	hypothetical protein
74	group_8420	hypothetical protein
74	group_8189	hypothetical protein
74	group_7278	hypothetical protein;Putative thymidine phosphorylase
74	group_6271	hypothetical protein
74	group_5920	hypothetical protein
74	maoA~~~tynA	Primary amine oxidase
74	nicB	6-hydroxy-3-succinoylpyridine 3-monooxygenase HspA;hypothetical protein
74	group_4465	hypothetical protein
74	group_858	hypothetical protein
74	group_12043	hypothetical protein
74	group_12042	hypothetical protein
74	ligJ_2~~~ligJ_1~~~ligJ	2-keto-4-carboxy-3-hexenedioate hydratase
74	proA_2	4-hydroxy-4-methyl-2-oxoglutarate aldolase/4-carboxy-4-hydroxy-2-oxoadipate aldolase
74	group_11387	hypothetical protein
74	group_10327	hypothetical protein

	yabJ_2~~~yabJ_4~~~yabJ_5~~~yabJ_1~~~	hypothetical protein;2-iminobutanoate/2-iminopropanoate
74	yabJ_3	deaminase
74	dksA_2~~~dksA_1	RNA polymerase-binding transcription factor DksA
74	group_8048	hypothetical protein
74	group_6377	hypothetical protein
74	group_6215	hypothetical protein
74	group_5799	hypothetical protein
74	group_5328	hypothetical protein
74	group_3534	hypothetical protein
74	group_1859	hypothetical protein
74	group_619	hypothetical protein
74	hdfR_4~~~hdfR_3	HTH-type transcriptional regulator HdfR
74	ligU	(4E)-oxalomesaconate Delta-isomerase
74	group_11449	hypothetical protein
74	aqpZ_2~~~aqpZ~~~aqpZ_1	Aquaporin Z
74	group_3121	hypothetical protein
74	group_2659	hypothetical protein
74	group_2604	hypothetical protein
74	group_1640	hypothetical protein
74	mhpB	23-dihydroxyphenylpropionate/23-dihydroxicinnamic acid 12-dioxygenase
74	group_10524	hypothetical protein
74	group_10409	hypothetical protein
74	nodD2_2	Nodulation protein D 2
74	group_8745	hypothetical protein
74	group_8354	hypothetical protein
74	group_8353	hypothetical protein
74	group_8277	hypothetical protein
74	group_7680	hypothetical protein
74	group_6968	hypothetical protein
74	group_6478	hypothetical protein
74	group_6477	hypothetical protein
74	group_6248	hypothetical protein
74	arsC_3~~~arsC_2~~~arsC	Arsenate reductase
74	cadI~~~cadI_1	Cadmium-induced protein CadI
74	group_3677	hypothetical protein
74	group_3125	hypothetical protein
74	group_2738	hypothetical protein

74	group_7	hypothetical protein;Putative antitoxin VapB45
74	group_11876	hypothetical protein
74	uppP_2	Undecaprenyl-diphosphatase
74	group_2853	hypothetical protein
74	group_905	hypothetical protein
74	group_9720	hypothetical protein
74	group_8094	hypothetical protein
74	group_7902	hypothetical protein
74	group_5986	hypothetical protein
74	group_5909	hypothetical protein
74	group_4273	hypothetical protein
74	group_3649	hypothetical protein
74	mdtA_1	Multidrug resistance protein MdtA
74	group_1301	hypothetical protein
74	group_1010	hypothetical protein
74	group_465	hypothetical protein
74	group_425	hypothetical protein
74	dadA_1~~~dadA_4~~~dadA_3	D-amino acid dehydrogenase
74	group_10834	hypothetical protein
74	group_10462	HTH-type transcriptional activator CmpR
74	group_10370	hypothetical protein
74	plcB~~~plcC	Phospholipase C 2;hypothetical protein;Phospholipase C 3
74	group_9847	hypothetical protein
74	group_9567	hypothetical protein
74	group_9392	hypothetical protein
74	group_8812	hypothetical protein
74	desVI	dTDP-3-amino-346-trideoxy-alpha-D-glucopyranose
74	group_7781	hypothetical protein
74	group_7528	hypothetical protein
74	group_7480	hypothetical protein
74	group_7364	hypothetical protein
74	group_7269	hypothetical protein
74	group_7236	hypothetical protein
74	group_7234	hypothetical protein
74	group_7214	hypothetical protein
74	group_6882	hypothetical protein
74	group_6562	hypothetical protein;IS5 family transposase IS1421
74	group_6507	hypothetical protein
74	group_6506	hypothetical protein

74	group_5989	hypothetical protein
74	group_4707	hypothetical protein
74	group_4077	hypothetical protein
74	group_3894	hypothetical protein
74	group_3479	hypothetical protein
74	cdiA	hypothetical protein;Toxin CdiA
74	group_2640	IS3 family transposase ISRso10;IS3 family transposase ISButh1
74	group_2636	hypothetical protein
74	group_2035	hypothetical protein
74	inhA_2	Isonitrile hydratase
74	group_937	hypothetical protein
74	dadA1_2~~~dadA_3~~~thiO	D-amino acid dehydrogenase 1;D-amino acid dehydrogenase;Glycine oxidase
74	yabJ_5~~~yabJ_4	2-iminobutanoate/2-iminopropanoate deaminase;hypothetical protein
74	aam	Acylamidase
74	proP_4~~~ousA~~~proP_5	Proline/betaine transporter;Glycine betaine/proline/ectoine/pipecolic acid transporter OusA
74	group_562	hypothetical protein
74	group_470	hypothetical protein
74	group_6853	hypothetical protein
74	group_6135	hypothetical protein
74	arsC_2~~~arsC_1	Arsenate reductase
74	group_3498	hypothetical protein
74	group_3483	hypothetical protein
74	group_3072	hypothetical protein
74	group_1962	hypothetical protein
74	group_1888	hypothetical protein
74	group_1723	hypothetical protein
74	group_1387	hypothetical protein
74	group_1103	hypothetical protein
74	group_255	hypothetical protein
74	group_11323	hypothetical protein
74	group_10832	hypothetical protein
74	group_10194	hypothetical protein
74	group_10093	hypothetical protein
74	group_9831	hypothetical protein
74	group_9429	hypothetical protein
74	group_9183	Putative universal stress protein
74	group_7991	hypothetical protein

74	dmlA~~~dmlA_1~~~dmlA_3	D-malate dehydrogenase [decarboxylating]
74	group_6825	hypothetical protein
74	group_6824	hypothetical protein
74	group_6802	hypothetical protein
74	group_5935	hypothetical protein
74	group_4444	hypothetical protein
74	group_4344	hypothetical protein
74	aioA	Arsenite oxidase subunit AioA
74	aioB	Arsenite oxidase subunit AioB
74	xerC_7~~~xerC_1~~~xerD_2~~~xerC_5	Tyrosine recombinase XerC;Tyrosine recombinase XerD
74	group_2176	hypothetical protein
74	group_2137	hypothetical protein
74	group_10971	hypothetical protein
74	group_10585	hypothetical protein
74	group_10340	hypothetical protein
74	group_10005	hypothetical protein
74	group_7746	hypothetical protein
74	group_6399	hypothetical protein
74	group_5956	hypothetical protein
74	group_3371	hypothetical protein
74	group_2821	hypothetical protein
74	group_2032	hypothetical protein
74	group_1509	hypothetical protein
74	group_10902	hypothetical protein
74	group_10159	hypothetical protein
74	group_9775	hypothetical protein
74	group_9730	hypothetical protein
74	group_9307	hypothetical protein
74	group_7683	hypothetical protein
74	group_7672	hypothetical protein
74	group_7408	hypothetical protein
74	group_7267	hypothetical protein
74	group_7228	hypothetical protein
74	group_7227	hypothetical protein
74	group_6781	hypothetical protein
74	group_5127	hypothetical protein
74	group_4948	hypothetical protein
74	group_3855	hypothetical protein
74	group_3647	hypothetical protein
74	group_3300	hypothetical protein
74	group_3292	hypothetical protein

74	group_3234	hypothetical protein
74	group_2746	hypothetical protein
74	group_1994	hypothetical protein
74	group_1481	hypothetical protein
74	group_235	hypothetical protein
74	group_67	hypothetical protein
74	group_45	hypothetical protein
74	group_11261	Putative phosphoribosyl transferase
74	group_10726	hypothetical protein
74	group_10090	hypothetical protein
74	sbnD_1	Staphyloferrin B transporter
74	group_9313	hypothetical protein
74	group_8935	hypothetical protein
74	group_7595	hypothetical protein
74	group_7560	hypothetical protein
74	group_7231	hypothetical protein
74	group_7151	hypothetical protein
74	group_7054	hypothetical protein
74	oprM_7~~~oprM_1	Outer membrane protein OprM
74	group_6164	hypothetical protein
74	group_5850	hypothetical protein
74	ytrE_1	ABC transporter ATP-binding protein YtrE
74	fabD_2	hypothetical protein;Malonyl CoA-acyl carrier protein transacylase
74	madD	Malonyl-S-ACP:biotin-protein carboxyltransferase MADD;hypothetical protein
74	accD_2	hypothetical protein;Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta chloroplastic
74	group_2924	hypothetical protein
74	group_2906	hypothetical protein
74	mdcB~~~citG	2-(5''-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase
74	group_2693	hypothetical protein;Universal stress protein
74	group_2566	hypothetical protein
74	group_1650	hypothetical protein
74	group_11272	hypothetical protein
74	group_8150	hypothetical protein
74	group_8145	hypothetical protein
74	group_6196	hypothetical protein

74	group_5019	hypothetical protein
74	nepI~~~nepI_2	Purine ribonucleoside efflux pump NepI
74	group_3514	hypothetical protein
74	group_3216	hypothetical protein
74	kdpB_1	Potassium-transporting ATPase ATP-binding subunit
74	group_2915	hypothetical protein
74	group_2774	hypothetical protein
74	group_2488	hypothetical protein
74	fatA_1	Ferric-anguibactin receptor FatA
74	group_2312	hypothetical protein
74	group_2173	hypothetical protein
74	group_1971	hypothetical protein
74	group_603	hypothetical protein
74	group_10956	hypothetical protein
74	group_10307	hypothetical protein
74	mltF_2~~~mltF_1	Membrane-bound lytic murein transglycosylase F
74	rarA_2	Replication-associated recombination protein A
74	rspR_1	hypothetical protein;HTH-type transcriptional repressor RspR
74	group_9391	hypothetical protein
74	group_9156	hypothetical protein
74	group_9087	hypothetical protein
74	group_9086	hypothetical protein
74	group_8644	hypothetical protein
74	yecD_2~~~yecD_1	Isochorismatase family protein YecD
74	pks5	Mycocerosic acid synthase-like polyketide synthase;hypothetical protein
74	group_6887	hypothetical protein
74	group_6886	hypothetical protein
74	group_6718	hypothetical protein
74	group_6535	hypothetical protein
74	group_6499	hypothetical protein
74	group_6411	hypothetical protein
74	group_5703	hypothetical protein
74	group_3290	hypothetical protein
74	group_1657	hypothetical protein
74	group_1229	hypothetical protein
74	group_1225	hypothetical protein
74	group_8362	hypothetical protein

74	group_7673	hypothetical protein
74	group_7107	hypothetical protein
74	group_6828	hypothetical protein
74	group_6070	hypothetical protein
74	group_5877	hypothetical protein
74	group_5469	hypothetical protein
74	group_5071	hypothetical protein
74	group_4945	hypothetical protein
74	pinR_2~~~pinR	Serine recombinase PinR
74	group_3115	hypothetical protein
74	group_2528	hypothetical protein
74	group_1334	hypothetical protein
74	group_11427	hypothetical protein
74	group_11424	hypothetical protein
74	hupB_3~~~hup	DNA-binding protein HU-beta;DNA-binding protein HU
74	group_11307	hypothetical protein
74	group_10106	IS5 family transposase ISAzo23
74	group_9311	hypothetical protein
74	group_9105	hypothetical protein
74	group_7690	hypothetical protein
74	group_7309	hypothetical protein
74	group_7308	hypothetical protein
74	group_6904	hypothetical protein
74	group_6903	hypothetical protein
74	group_6832	hypothetical protein
74	group_6004	hypothetical protein
74	cynR_4~~~cynR_3	HTH-type transcriptional regulator CynR;hypothetical protein
74	group_5835	hypothetical protein
74	group_4624	hypothetical protein
74	group_4232	hypothetical protein
74	group_3486	hypothetical protein
74	group_3334	hypothetical protein
74	group_3333	hypothetical protein
74	group_2955	hypothetical protein
74	group_2777	hypothetical protein
74	group_2092	hypothetical protein
74	group_2052	hypothetical protein
74	group_2044	hypothetical protein
74	group_1965	hypothetical protein
74	group_1908	hypothetical protein
74	group_1661	hypothetical protein

74	group_1660	hypothetical protein
74	group_1618	hypothetical protein
74	group_1587	hypothetical protein
74	group_1234	hypothetical protein
74	group_1159	hypothetical protein
74	group_1122	hypothetical protein
74	cbbZC_1~~~gph_1~~~cbbZC_2	Phosphoglycolate phosphatase chromosomal;Phosphoglycolate phosphatase;hypothetical protein
74	group_755	hypothetical protein
74	group_11478	IS3 family transposase ISRso20
74	group_10109	IS5 family transposase ISAzo23
74	group_9729	hypothetical protein
74	group_7614	IS5 family transposase ISAzo23
74	group_904	hypothetical protein
74	group_10091	IS5 family transposase IS1021
74	group_10080	IS5 family transposase IS1021
74	group_10079	IS5 family transposase IS1021
74	group_9844	hypothetical protein
74	group_6748	hypothetical protein;IS3 family transposase ISRso20
74	group_6198	hypothetical protein
74	group_5453	hypothetical protein
74	group_3174	hypothetical protein
74	group_1533	IS5 family transposase IS1021
74	mbtH~~~mbtH_1	Protein MbtH
74	mdtK_1~~~mdtK_2	Multidrug resistance protein MdtK
74	group_10108	IS5 family transposase ISAzo23
74	group_10076	IS5 family transposase IS1021
74	group_9310	hypothetical protein
74	egl_2~~~egl	Endoglucanase
74	group_4301	hypothetical protein
74	group_2925	IS3 family transposase ISRso10
73	group_1607	hypothetical protein
73	group_1096	hypothetical protein
72	group_617	hypothetical protein
72	group_606	hypothetical protein
71	group_3474	hypothetical protein
71	group_2849	hypothetical protein
71	group_2700	hypothetical protein
71	group_2598	hypothetical protein
71	group_2597	hypothetical protein;Type IV secretion system protein virB10
71	group_2517	hypothetical protein

71	group_12031	hypothetical protein
71	group_12028	hypothetical protein
71	traG	Conjugal transfer protein TraG;hypothetical protein
71	group_3601	hypothetical protein
71	group_3600	hypothetical protein
71	virB4~~~virB4_2~~~virB4_1	Type IV secretion system protein virB4
71	hcaR~~~hdfR_3~~~hdfR_2~~~benM_1~~~hdfR_1~~~benM_3	Hca operon transcriptional activator HcaR;HTH-type transcriptional regulator HdfR;HTH-type transcriptional regulator BenM
71	group_2331	hypothetical protein
70	group_1656	hypothetical protein
70	group_1654	hypothetical protein
70	intA_2	Prophage integrase IntA
70	group_2382	hypothetical protein
70	group_1416	hypothetical protein
70	group_7362	hypothetical protein;Actin cross-linking toxin VgrG1
70	group_7004	hypothetical protein
70	group_7248	hypothetical protein
70	group_1345	hypothetical protein
70	group_1673	hypothetical protein
70	group_7709	hypothetical protein
70	group_2916	hypothetical protein
70	group_7249	hypothetical protein
70	group_7197	hypothetical protein
70	group_6854	hypothetical protein
70	group_4314	hypothetical protein
70	group_4117	hypothetical protein
70	group_342	hypothetical protein
69	group_9325	hypothetical protein
69	group_7294	hypothetical protein
68	group_7628	hypothetical protein
68	group_7457	Cyanate hydratase
68	triA_1	Melamine deaminase
68	group_3819	hypothetical protein
68	hisC_4	Histidinol-phosphate aminotransferase
67	group_5967	hypothetical protein
67	group_10768	hypothetical protein
67	group_7849	Tyrosine recombinase XerC
66	group_1560	hypothetical protein

66	group_1134	hypothetical protein
65	group_11259	hypothetical protein
65	intS~~~intS_2~~~intS_1~~~intA_3~~~intA	Prophage integrase IntS;Prophage integrase IntA
65	group_10591	hypothetical protein
65	group_1629	hypothetical protein
64	group_1281	hypothetical protein
64	shlB_1~~~shlB_7~~~shlB_3	Hemolysin transporter protein ShlB;hypothetical protein
63	group_9875	HTH-type transcriptional regulator DmlR
63	novR_1~~~novR_2	Decarboxylase NovR
63	nicT_2	Putative metabolite transport protein NicT
62	group_1883	hypothetical protein
62	group_1881	hypothetical protein
62	group_1719	hypothetical protein
62	group_1389	hypothetical protein
62	group_1360	hypothetical protein
62	group_10417	hypothetical protein
62	group_2089	hypothetical protein
62	group_2060	hypothetical protein
62	group_2023	hypothetical protein
62	group_2022	hypothetical protein
62	group_1979	hypothetical protein
62	group_1977	hypothetical protein
62	group_10864	hypothetical protein
62	group_1970	hypothetical protein
62	group_1969	hypothetical protein
62	group_1866	hypothetical protein
62	group_1694	hypothetical protein
62	group_1323	hypothetical protein
62	group_9879	hypothetical protein
62	group_9802	hypothetical protein
62	group_9353	hypothetical protein
62	group_9176	hypothetical protein
62	group_9038	hypothetical protein
62	group_9033	hypothetical protein
62	group_8345	hypothetical protein
62	group_7850	hypothetical protein
62	group_7387	hypothetical protein
62	group_7257	hypothetical protein
62	group_7028	hypothetical protein
62	group_6803	hypothetical protein

62	group_6789	hypothetical protein
62	group_6241	hypothetical protein
62	group_5440	hypothetical protein
62	group_3981	hypothetical protein
62	group_2572	hypothetical protein
62	group_9298	hypothetical protein
61	group_12049	hypothetical protein
61	group_7610	hypothetical protein
61	group_3670	hypothetical protein
60	group_1769	hypothetical protein
60	group_1635	hypothetical protein
59	group_2283	hypothetical protein
59	group_2355	hypothetical protein
58	group_3476	hypothetical protein
58	group_1293	hypothetical protein
58	group_801	hypothetical protein
58	group_2451	hypothetical protein
58	group_2851	hypothetical protein
58	group_806	hypothetical protein
57	group_3015	hypothetical protein;IS3 family transposase ISAisp2
57	group_1615	IS3 family transposase ISAisp2
56	group_11761	hypothetical protein
56	group_11431	hypothetical protein
56	group_11426	hypothetical protein
56	group_10904	hypothetical protein
55	prsA_2	putative parvulin-type peptidyl-prolyl cis-trans isomerase;hypothetical protein;Foldase protein PrsA
55	group_1315	hypothetical protein
54	group_2857	hypothetical protein
54	group_1151	hypothetical protein
54	group_7463	hypothetical protein
53	group_2291	hypothetical protein
53	group_2958	hypothetical protein
53	group_2957	hypothetical protein
53	group_2304	hypothetical protein
52	group_3426	hypothetical protein
52	group_3422	hypothetical protein
51	group_2895	hypothetical protein
51	group_10066	hypothetical protein
50	group_10001	hypothetical protein;All-trans-zeta-carotene desaturase

50	group_9630	hypothetical protein
50	group_8922	Isatin hydrolase;hypothetical protein
50	styA	Styrene monooxygenase StyA;hypothetical protein
50	rhaS_1~~~pchR~~~feaR	hypothetical protein;HTH-type transcriptional activator RhaS;Regulatory protein PchR;Transcriptional activator FeaR
50	ntaB	FMN reductase (NADH) NtaB
50	group_9629	hypothetical protein
50	group_8088	hypothetical protein
50	yisK	putative protein YisK
50	maiA_1~~~maiA_2~~~maiA	Maleate isomerase
50	sdgC~~~namA_1~~~namA_2	Salicyloyl-CoA 5-hydroxylase;NADPH dehydrogenase
50	group_6637	hypothetical protein
50	group_6636	hypothetical protein
50	nitA	Aliphatic nitrilase;hypothetical protein
50	group_5880	14-dihydroxy-2-naphthoyl-CoA hydrolase
50	sdgD_2~~~sdgD_1	Gentisate 12-dioxygenase
49	group_10861	hypothetical protein
49	group_2839	hypothetical protein
48	group_10420	hypothetical protein
48	group_10400	IS3 family transposase ISAisp2
47	group_3445	hypothetical protein
47	group_1287	hypothetical protein
46	group_2979	hypothetical protein
46	group_9524	Putative defective protein IntQ
45	group_9243	hypothetical protein
45	group_9242	hypothetical protein
45	group_8917	hypothetical protein
45	group_9999	hypothetical protein
44	group_1691	hypothetical protein
44	esiB	hypothetical protein;Secretory immunoglobulin A-binding protein EsiB
43	group_11495	hypothetical protein
43	gpFI~~~gpFI_2~~~gpFI_1~~~gpFI_3	Putative prophage major tail sheath protein
43	group_1487	hypothetical protein
43	group_123	hypothetical protein

42	group_2437	Putative transposase InsK for insertion sequence element IS150;IS3 family transposase ISAisp2
42	group_2392	hypothetical protein;IS3 family transposase ISAisp2
41	xerC_5~~~xerC_3~~~xerC_7	hypothetical protein;Tyrosine recombinase XerC
41	group_5769	Tn3 family transposase ISPa43;hypothetical protein
40	group_3764	hypothetical protein
40	group_3767	hypothetical protein
39	dltA_2~~~dltA_1~~~dltA_3	hypothetical protein;D-alanine--D-alanyl carrier protein ligase
39	group_2425	hypothetical protein
39	rhaR_1~~~rhaR_3	HTH-type transcriptional activator RhaR
39	acpP_1~~~acpP_2~~~ppsE~~~ppsE_1	Acyl carrier protein;hypothetical protein;Phthiocerol synthesis polyketide synthase type I PpsE
39	irtA	putative ABC transporter ATP-binding protein;Iron import ATP-binding/permease protein IrtA
39	btuD_3~~~btuD_1~~~btuD_5	Vitamin B12 import ATP-binding protein BtuD
39	fyuA	Pesticin receptor;hypothetical protein
39	acs~~~eryA~~~dltA_1~~~dltA_2~~~dltA_4~~~menE_1~~~acs_2~~~menE_3	Acetyl-coenzyme A synthetase;4-hydroxyphenylalkanoate adenylyltransferase;6-deoxyerythronolide-B synthase EryA2 modules 3 and 4;D-alanine--D-alanyl carrier protein ligase;Phenolphthiocerol synthesis polyketide synthase type I Pks15/1;hypothetical protein;2-succinylbenzoate--CoA ligase
39	group_3012	hypothetical protein
39	COQ5_2~~~sfp~~~COQ5_1	hypothetical protein;2-methoxy-6-polyprenyl-14-benzoquinol methylase mitochondrial;4'-phosphopantetheinyl transferase Sfp
39	group_11370	N-acetylmuramoyl-L-alanine amidase AmiD;hypothetical protein
38	noc_1~~~noc_4~~~noc_2	hypothetical protein;Nucleoid occlusion protein
38	group_3742	hypothetical protein

37	group_3734	hypothetical protein
37	group_2704	hypothetical protein
37	group_2650	hypothetical protein
37	group_2761	hypothetical protein
37	group_1626	hypothetical protein
36	group_11402	hypothetical protein
36	group_1818	hypothetical protein
36	group_11741	hypothetical protein
35	group_55	hypothetical protein
35	group_22	hypothetical protein
34	group_2827	hypothetical protein
34	group_2830	hypothetical protein
34	group_4835	hypothetical protein
33	group_3230	hypothetical protein
33	group_132	hypothetical protein
33	group_2387	hypothetical protein
33	group_599	IS630 family transposase ISCARN39
33	group_11985	hypothetical protein
33	group_11756	hypothetical protein
33	group_11422	hypothetical protein
33	group_407	hypothetical protein
33	group_10389	hypothetical protein
33	group_4318	hypothetical protein
33	group_3457	hypothetical protein
33	group_10388	D-serine/D-alanine/glycine transporter
33	group_8091	hypothetical protein
33	group_6126	hypothetical protein
33	group_10313	hypothetical protein
32	group_3086	hypothetical protein
32	group_4630	hypothetical protein
32	group_1520	hypothetical protein
32	group_1518	hypothetical protein
32	group_4196	hypothetical protein
31	group_2930	hypothetical protein
31	group_12015	hypothetical protein
31	group_11813	hypothetical protein
31	group_11493	hypothetical protein
31	group_10992	hypothetical protein
31	group_10153	hypothetical protein
31	group_10993	hypothetical protein
31	group_10155	hypothetical protein
31	group_405	hypothetical protein

31	group_2568	hypothetical protein
31	gpFI_1~~~gpFI~~~gpFI_2	Putative prophage major tail sheath protein
30	group_2671	IS3 family transposase ISAtu5;IS3 family transposase ISRso14;hypothetical protein
30	group_7602	IS3 family transposase ISBam2;IS3 family transposase IS407
29	group_1255	hypothetical protein
29	group_10213	hypothetical protein
28	group_10892	hypothetical protein
28	group_10006	hypothetical protein
28	group_1829	hypothetical protein
28	group_9342	hypothetical protein
28	group_6627	hypothetical protein
28	group_1350	hypothetical protein
27	group_152	hypothetical protein
27	group_313	hypothetical protein
26	intA_1~~~intA~~~intA_2	Prophage integrase IntA
26	group_2653	hypothetical protein
26	group_1647	hypothetical protein
26	group_1645	hypothetical protein
25	hin_1~~~hin~~~hin_2	DNA-invertase hin;hypothetical protein
25	group_11821	hypothetical protein
25	group_1447	hypothetical protein
25	group_10227	hypothetical protein
25	group_9261	hypothetical protein
25	group_7173	hypothetical protein
25	group_7172	hypothetical protein
25	group_2264	hypothetical protein
25	group_1243	hypothetical protein
24	group_296	hypothetical protein
24	group_11637	hypothetical protein
23	group_9544	hypothetical protein
23	group_10428	hypothetical protein
22	group_3409	IS3 family transposase IS401;hypothetical protein
22	group_2398	IS3 family transposase IS401;Putative transposase InsK for insertion sequence element IS150
21	group_3215	hypothetical protein
21	noc_2~~~noc_4~~~noc_3	Nucleoid occlusion protein;hypothetical protein

21	noc_1~~~noc_2~~~noc_3	Nucleoid occlusion protein;hypothetical protein
20	group_10995	hypothetical protein
20	group_1271	hypothetical protein
20	group_2964	hypothetical protein
20	group_1086	hypothetical protein
20	group_2963	hypothetical protein
19	group_5966	hypothetical protein
19	group_3397	hypothetical protein
19	group_2480	hypothetical protein
18	group_1009	IS66 family transposase IS1313
18	group_1054	IS66 family transposase ISAeh1;IS66 family transposase ISPa82;hypothetical protein
18	group_800	hypothetical protein
17	group_1762	hypothetical protein
17	group_1526	hypothetical protein
17	group_852	hypothetical protein
17	clsB_3	Cardiolipin synthase B;Cardiolipin synthase;hypothetical protein
17	group_845	hypothetical protein
17	group_2862	hypothetical protein
17	group_1672	hypothetical protein
17	group_1643	hypothetical protein
17	group_1525	hypothetical protein
17	group_1270	hypothetical protein
17	group_9198	hypothetical protein
17	group_10795	hypothetical protein
17	group_10190	hypothetical protein
17	group_10131	hypothetical protein
17	group_9905	hypothetical protein
17	group_9904	hypothetical protein
17	group_9130	hypothetical protein
17	group_8475	hypothetical protein
17	group_3280	hypothetical protein
17	ydiP_2~~~ydiP	hypothetical protein;putative BsuMI modification methylase subunit YdiP
17	group_10985	hypothetical protein
17	group_9199	hypothetical protein
17	group_11797	hypothetical protein
17	group_7072	hypothetical protein
16	group_11070	hypothetical protein
16	group_3166	hypothetical protein
15	group_11812	hypothetical protein

15	group_11826	hypothetical protein
15	group_1418	hypothetical protein
15	group_11550	Modification methylase DpnIIB;hypothetical protein
15	group_7175	hypothetical protein
15	group_1279	hypothetical protein
15	xerC_1~~~xerC_2~~~xerC_5~~~xerC_10~~~xerC_3	Tyrosine recombinase XerC
15	group_787	hypothetical protein
15	group_10880	hypothetical protein
15	group_172	hypothetical protein
15	group_2677	hypothetical protein
15	group_500	hypothetical protein
15	group_460	hypothetical protein
15	group_375	hypothetical protein
15	group_315	hypothetical protein
15	group_155	hypothetical protein
15	group_11530	hypothetical protein
15	group_11055	hypothetical protein
15	group_7174	hypothetical protein
15	group_870	hypothetical protein
15	group_266	hypothetical protein
15	group_9263	hypothetical protein
15	group_9274	hypothetical protein
15	group_10332	hypothetical protein
15	group_8787	hypothetical protein
15	group_7922	hypothetical protein
14	smc_2~~~smc_1	Chromosome partition protein Smc
14	group_10693	putative protein/MSMEI_1241
14	group_9792	hypothetical protein
14	group_8210	hypothetical protein
14	group_1203	hypothetical protein
14	group_923	hypothetical protein
14	group_11314	hypothetical protein
14	group_8232	hypothetical protein
13	group_8279	hypothetical protein
13	group_3461	hypothetical protein
13	group_6106	hypothetical protein
13	group_11340	hypothetical protein
13	group_7413	hypothetical protein
13	group_7116	hypothetical protein
13	group_5897	hypothetical protein
13	group_3398	hypothetical protein

13	group_174	hypothetical protein
13	group_7187	hypothetical protein
13	group_5729	hypothetical protein
13	group_5710	hypothetical protein
13	group_10100	hypothetical protein
13	group_8916	hypothetical protein
13	group_7406	hypothetical protein
13	group_6548	hypothetical protein
13	group_6522	hypothetical protein
13	group_6071	hypothetical protein
13	group_4356	hypothetical protein
13	group_3307	hypothetical protein
13	group_2167	hypothetical protein
13	group_11837	hypothetical protein
13	group_10246	Glycine cleavage system transcriptional activator
13	group_6127	Glycine betaine/proline/ectoine/pipecolic acid transporter OusA
12	group_1741	hypothetical protein
12	group_1743	hypothetical protein
12	group_1444	hypothetical protein
12	group_7723	hypothetical protein
11	group_2694	hypothetical protein
11	teaD	Putative universal stress protein;TRAP-T-associated universal stress protein TeaD
11	group_2763	hypothetical protein
11	group_1929	hypothetical protein
11	group_8230	hypothetical protein
11	ccpA_2~~~ccpA_3~~~treR~~~ccpA_1	Catabolite control protein A;HTH-type transcriptional regulator TreR;hypothetical protein
11	sacX	Negative regulator of SacY activity
11	group_11352	Sucrose-6-phosphate hydrolase
11	group_11165	hypothetical protein
11	nemA_2~~~nemA_1~~~nemA_3	N-ethylmaleimide reductase
11	group_10746	Maltoporin
11	citB~~~citB_1~~~acn	Aconitate/2-methylaconitate hydratase;Aconitate hydratase A
11	group_10472	hypothetical protein
11	hicd	Homoisocitrate dehydrogenase
11	araG_1~~~mglA_3~~~araG~~~araG_2	Arabinose import ATP-binding protein AraG;Galactose/methyl

		galactoside import ATP-binding protein MglA;hypothetical protein
11	group_10058	hypothetical protein
11	ptsI_3~~~ptsI_2	hypothetical protein;Phosphoenolpyruvate-protein phosphotransferase
11	group_9715	hypothetical protein
11	group_9664	hypothetical protein
11	group_9615	hypothetical protein
11	group_9071	hypothetical protein
11	group_9027	hypothetical protein
11	group_9026	hypothetical protein
11	group_8315	hypothetical protein
11	resA_2~~~resA_3~~~resA_4	Thiol-disulfide oxidoreductase ResA
11	group_8215	hypothetical protein
11	acnB~~~leuC_2~~~acnB_2~~~acnB_3~~~acnB_1	Aconitate hydratase B;3-isopropylmalate dehydratase large subunit
11	nanR~~~lutR_2~~~lutR_1	HTH-type transcriptional repressor NanR;HTH-type transcriptional regulator LutR
11	emrB_2~~~emrB_3~~~emrY_2	Multidrug export protein EmrB;hypothetical protein;putative multidrug resistance protein EmrY
11	scrK~~~RBKS	Fructokinase;Ribokinase
11	glnP_1~~~glnP_2	Glutamine transport system permease protein GlnP;hypothetical protein
11	group_6780	hypothetical protein
11	group_6762	hypothetical protein
11	bauC_1~~~iolA_1	Putative 3-oxopropanoate dehydrogenase;Malonate-semialdehyde dehydrogenase
11	group_6133	hypothetical protein
11	group_6052	hypothetical protein
11	dmlR_11~~~dmlR_12	HTH-type transcriptional regulator DmlR
11	glnH_1	ABC transporter glutamine-binding protein GlnH
11	folE	GTP cyclohydrolase 1
11	glnM_1~~~yecS_2	putative glutamine ABC transporter permease protein GlnM;L-cystine transport system permease protein YecS
11	group_2519	Glutamine transport ATP-binding protein GlnQ

11	group_2342	hypothetical protein
11	group_1213	hypothetical protein
11	group_11924	hypothetical protein
11	group_11885	hypothetical protein
11	group_11670	hypothetical protein
11	group_11299	hypothetical protein
11	narH_1~~~narH_2~~~narY~~~narH	Respiratory nitrate reductase 1 beta chain;Respiratory nitrate reductase 2 beta chain
11	iolD~~~iolD_1~~~iolD_2	3D-(35/4)-trihydroxycyclohexane-12-dione hydrolase
11	group_11039	hypothetical protein
11	group_11007	hypothetical protein
11	iolG_1	Inositol 2-dehydrogenase/D-chiro-inositol 3-dehydrogenase
11	group_10685	hypothetical protein
11	iolE_2~~~iolE_1	Inosose dehydratase;hypothetical protein
11	group_10330	hypothetical protein
11	group_10037	2-amino-3-ketobutyrate coenzyme A ligase;hypothetical protein
11	group_10036	hypothetical protein
11	group_9970	hypothetical protein
11	group_9954	hypothetical protein
11	thpA~~~thpA_1	D-threitol-binding protein;hypothetical protein
11	iolG_2~~~iolG_1	hypothetical protein;Inositol 2-dehydrogenase/D-chiro-inositol 3-dehydrogenase
11	narJ	hypothetical protein;Nitrate reductase molybdenum cofactor assembly chaperone NarJ
11	sutR_2~~~sutR_3	HTH-type transcriptional regulator SutR;hypothetical protein
11	cmr	HTH-type transcriptional regulator Cmr;Cyclic AMP receptor protein;hypothetical protein
11	group_8988	putative epimerase/dehydratase;hypothetical protein
11	ytfE	Iron-sulfur cluster repair protein YtfE
11	degU_2~~~cheB_2~~~degU_1	Transcriptional regulatory protein DegU;Protein-glutamate methylesterase/protein-glutamine glutaminase

11	mobB_2~~~mobB_1	Molybdopterin-guanine dinucleotide biosynthesis adapter protein;hypothetical protein
11	dmlR_4~~~dmlR_14~~~dmlR_12~~~dmlR_17	HTH-type transcriptional regulator DmlR;hypothetical protein
11	group_8742	hypothetical protein
11	mglA_1~~~btuD_4~~~mglA_2	Galactose/methyl galactoside import ATP-binding protein MglA;Vitamin B12 import ATP-binding protein BtuD
11	murR_1~~~murR_2	HTH-type transcriptional regulator MurR
11	narI~~~narI_1~~~narI_2	Respiratory nitrate reductase 1 gamma chain;hypothetical protein
11	iolC~~~iolC_1	5-dehydro-2-deoxygluconokinase
11	iolE_1~~~iolE_2	Inosose dehydratase
11	queE_2~~~queE_1	7-carboxy-7-deazaguanine synthase;hypothetical protein
11	group_7386	Anaerobic ribonucleoside-triphosphate reductase;hypothetical protein
11	norR_5~~~norR2	hypothetical protein;Anaerobic nitric oxide reductase transcription regulator NorR;Nitric oxide reductase transcription regulator NorR2
11	group_6652	hypothetical protein
11	rbsC_2~~~rbsC_3	Ribose import permease protein RbsC
11	tolC	hypothetical protein;Outer membrane protein TolC
11	nasA_1~~~nasA_2	hypothetical protein;Nitrate transporter
11	narK_1~~~narK_2~~~narK	hypothetical protein;Nitrate/nitrite transporter NarK
11	group_2727	hypothetical protein
11	drgA	Protein DrgA
11	narZ_1~~~narG_2~~~narZ_2~~~narG_3~~~narG_1~~~narG	Respiratory nitrate reductase 2 alpha chain;Respiratory nitrate reductase 1 alpha chain
11	group_1385	hypothetical protein
11	group_1000	hypothetical protein
11	group_291	hypothetical protein
11	group_202	hypothetical protein
11	group_83	hypothetical protein

11	hdfR~~~hdfR_1~~~hdfR_2	HTH-type transcriptional regulator HdfR;hypothetical protein
11	yihG~~~yihG_1~~~yihG_2	putative acyltransferase YihG
11	group_11403	hypothetical protein
11	group_11177	hypothetical protein
11	apbE_2~~~apbE_1	FAD:protein FMN transferase;hypothetical protein
11	group_10534	hypothetical protein;HTH-type transcriptional regulator CdhR;HTH-type transcriptional regulator
11	group_10358	hypothetical protein;Cytochrome c-555
11	group_10352	hypothetical protein
11	group_10134	hypothetical protein
11	mmgF	putative protein;2-methylisocitrate lyase
11	group_9481	hypothetical protein
11	nosZ	Nitrous-oxide reductase
11	group_9388	hypothetical protein
11	yhbU_1~~~yhbU_2~~~yhbU	putative protease YhbU
11	group_9133	hypothetical protein
11	nosL	Copper-binding lipoprotein NosL
11	yahK_2~~~adhC2~~~adhC	Aldehyde reductase YahK;hypothetical protein;NADP-dependent alcohol dehydrogenase C 2;NADP-dependent alcohol dehydrogenase C
11	nosD~~~nosD_1~~~nosD_2	putative ABC transporter binding protein NosD
11	group_7486	hypothetical protein
11	group_6966	hypothetical protein
11	nosY	hypothetical protein;putative ABC transporter permease protein NosY
11	group_6712	Nitrate/nitrite sensor protein NarX;hypothetical protein
11	group_6341	hypothetical protein
11	pgrR_3~~~pgrR_4	HTH-type transcriptional regulator PgrR
11	nosF	putative ABC transporter ATP-binding protein NosF
11	group_5621	hypothetical protein
11	group_5504	hypothetical protein
11	aiiA	N-acyl homoserine lactonase;hypothetical protein
11	group_2819	hypothetical protein

11	aspC_2~~~aspC_1~~~hisC_6	Aspartate aminotransferase;hypothetical protein;Histidinol-phosphate aminotransferase
11	ucpA_2~~~gdh_1	Oxidoreductase UcpA;Galactitol 2-dehydrogenase
11	group_11599	hypothetical protein
11	TDO2	Tryptophan 23-dioxygenase
11	cefD~~~egtE	Isopenicillin N epimerase;hercynylcysteine sulfoxide lyase
11	group_10963	hypothetical protein
11	group_10737	hypothetical protein
11	group_9299	hypothetical protein
11	group_7900	hypothetical protein
11	group_7256	hypothetical protein
11	group_7254	hypothetical protein
11	group_5024	hypothetical protein
11	cyoD_1~~~cyoD_2	Cytochrome bo(3) ubiquinol oxidase subunit 4;hypothetical protein
11	tar_1	Methyl-accepting chemotaxis protein II;hypothetical protein
11	group_9418	hypothetical protein
11	group_7884	Outer membrane porin protein 32;hypothetical protein
11	cyoC_2~~~cyoC_1	Cytochrome bo(3) ubiquinol oxidase subunit 3;hypothetical protein
11	cyoA~~~cyoA_2~~~cyoA_1	hypothetical protein;Cytochrome bo(3) ubiquinol oxidase subunit 2
11	cyoB_2~~~cyoB_3~~~cyoB_4	Cytochrome bo(3) ubiquinol oxidase subunit 1
11	group_3283	hypothetical protein
11	group_9	hypothetical protein
11	group_11194	hypothetical protein
11	group_11157	hypothetical protein
11	group_10947	Outer membrane porin protein 32;Outer membrane porin protein;hypothetical protein
11	triA~~~atzA~~~dadD_1~~~dadD~~~dadD_2	Melamine deaminase;Atrazine chlorohydrolase;5'-deoxyadenosine deaminase;hypothetical protein
11	iolB~~~iolB_1	5-deoxy-glucuronate isomerase
11	ptsJ~~~lysN_2	Vitamin B6 salvage pathway transcriptional repressor PtsJ;hypothetical protein;2-aminoadipate transaminase

11	purU_2	Formyltetrahydrofolate deformylase;hypothetical protein
11	bcr_3~~~bcr_2~~~mdtL~~~bcr_1	Bicyclomycin resistance protein;Multidrug resistance protein MdtL;hypothetical protein
11	group_8096	hypothetical protein
11	group_7987	hypothetical protein
11	soxG	Sarcosine oxidase subunit gamma;hypothetical protein
11	soxD	Sarcosine oxidase subunit delta
11	cdhR_1~~~cdhR_2~~~cdhR_3	HTH-type transcriptional regulator CdhR;hypothetical protein
11	group_6526	hypothetical protein
11	group_6208	Aldo-keto reductase IolS
11	soxA_1~~~soxA_2~~~soxA_3	Sarcosine oxidase subunit alpha;hypothetical protein
11	soxB_1~~~soxB_2~~~soxB	Sarcosine oxidase subunit beta
11	sdaA_1~~~sdaA_2~~~sdaB	L-serine dehydratase;L-serine dehydratase 2
11	group_10663	hypothetical protein
11	group_10662	hypothetical protein
11	sir_2~~~sir_1	Sulfite reductase [ferredoxin];hypothetical protein
11	group_9618	hypothetical protein
11	group_9265	hypothetical protein
11	group_9264	hypothetical protein
11	group_9042	hypothetical protein
11	group_3374	hypothetical protein
11	group_3173	hypothetical protein
11	group_3078	Ribonuclease
11	group_2940	hypothetical protein
11	group_2911	hypothetical protein
11	group_11774	HTH-type transcriptional regulator TsaR;hypothetical protein
11	group_11703	hypothetical protein
11	group_11441	hypothetical protein
11	group_11353	hypothetical protein
11	group_10748	hypothetical protein
11	group_9864	hypothetical protein
11	group_9852	hypothetical protein
11	group_9652	hypothetical protein
11	group_8318	hypothetical protein
11	group_7957	hypothetical protein
11	group_6597	hypothetical protein
11	group_6210	hypothetical protein

11	dkgA	putative oxidoreductase/MSMEI_2347;25-diketo-D-gluconic acid reductase A
11	group_2171	hypothetical protein
11	cdhR_2~~~cdhR_1	HTH-type transcriptional regulator CdhR;hypothetical protein
11	group_9504	hypothetical protein
11	aatA_1~~~aatB	Aspartate/prephenate aminotransferase;Aspartate aminotransferase
11	murR_3	HTH-type transcriptional regulator MurR
11	group_8125	hypothetical protein
11	group_6607	hypothetical protein
11	group_6401	hypothetical protein
11	group_5955	hypothetical protein
11	group_4818	hypothetical protein
11	group_4757	hypothetical protein
11	group_4236	hypothetical protein
11	group_2733	hypothetical protein
11	group_2610	hypothetical protein
11	group_640	hypothetical protein
11	group_11802	hypothetical protein
11	group_11534	hypothetical protein
11	group_11094	hypothetical protein
11	group_10249	hypothetical protein
11	group_9469	hypothetical protein
11	group_9401	hypothetical protein
11	group_8683	hypothetical protein
11	group_8064	hypothetical protein
11	group_7466	hypothetical protein
11	group_7464	hypothetical protein;Outer membrane porin protein
11	group_7431	hypothetical protein
11	group_7193	hypothetical protein
11	group_7179	hypothetical protein
11	group_6549	hypothetical protein
11	group_6312	hypothetical protein
11	group_6113	hypothetical protein
11	lifO	Lipase chaperone
11	mtlK	Mannitol 2-dehydrogenase
11	sadA~~~sadA_2	hypothetical protein;Autotransporter adhesin SadA

11	group_3984	hypothetical protein
11	group_2996	hypothetical protein
11	pdp	Putative thymidine phosphorylase;Pyrimidine-nucleoside phosphorylase
11	group_2872	hypothetical protein
11	group_1206	hypothetical protein
11	group_945	hypothetical protein
11	group_628	hypothetical protein
11	xylB	Xylulose kinase
11	eryD	Erythritol catabolism regulatory protein EryD
11	group_9610	hypothetical protein
11	group_9533	hypothetical protein;Catalase
11	group_9443	hypothetical protein
11	group_9373	hypothetical protein
11	group_7864	hypothetical protein
11	group_7804	hypothetical protein
11	group_7787	hypothetical protein
11	gph_1~~~gph_2~~~yieH_2	Phosphoglycolate phosphatase;6-phosphogluconate phosphatase
11	group_7642	hypothetical protein
11	group_6980	hypothetical protein
11	group_6820	hypothetical protein
11	group_6056	hypothetical protein
11	kbaZ	D-tagatose-16-bisphosphate aldolase subunit KbaZ
11	group_4931	L-aspartate/glutamate-specific racemase;hypothetical protein
11	sugB	hypothetical protein;Trehalose transport system permease protein SugB
11	group_4810	hypothetical protein
11	group_4782	hypothetical protein
11	pgl_4~~~pgl_1	6-phosphogluconolactonase;hypothetical protein
11	malK	Maltose/maltodextrin import ATP-binding protein MalK;hypothetical protein
11	group_3889	hypothetical protein
11	group_2854	hypothetical protein
11	melD	Melibiose/raffinose/stachyose import permease protein MelD;hypothetical protein

11	group_2121	hypothetical protein
11	polS	Sorbitol dehydrogenase
11	group_339	Tagatose kinase
11	group_11892	hypothetical protein
11	group_10484	hypothetical protein
11	group_9348	hypothetical protein
11	group_8153	hypothetical protein
11	pgrR_7~~~argP~~~argP_1~~~pgrR_2	hypothetical protein;HTH-type transcriptional regulator PgrR;HTH-type transcriptional regulator ArgP
11	group_7770	hypothetical protein
11	dapH_3~~~cysE	2345-tetrahydropyridine-26-dicarboxylate N-acetyltransferase;Serine acetyltransferase;hypothetical protein
11	group_7423	hypothetical protein
11	group_6717	putative HTH-type transcriptional regulator
11	group_6716	hypothetical protein
11	group_6413	hypothetical protein
11	ppsB~~~entF	Plipastatin synthase subunit B;Enterobactin synthase component F
11	group_5113	hypothetical protein
11	group_5026	hypothetical protein;putative protein
11	gltR_2~~~gltR_1	HTH-type transcriptional regulator GltR
11	group_4800	hypothetical protein
11	ectB_1~~~ectB~~~ectB_2	Diaminobutyrate--2-oxoglutarate transaminase
11	pvdA	hypothetical protein;L-ornithine N(5)-monooxygenase
11	group_4648	hypothetical protein
11	nepI	Purine ribonucleoside efflux pump NepI
11	group_2555	hypothetical protein
11	soxA_3~~~soxA_4	Monomeric sarcosine oxidase
11	group_387	hypothetical protein
11	group_11178	hypothetical protein
11	group_9757	hypothetical protein
11	pgi_2~~~pgi_1	Glucose-6-phosphate isomerase
11	group_6599	hypothetical protein
11	group_5681	hypothetical protein

11	yodB	Cytochrome b561
11	group_595	hypothetical protein
11	group_9853	hypothetical protein
11	group_8154	hypothetical protein
11	group_7717	hypothetical protein
11	group_7544	hypothetical protein
11	group_2737	hypothetical protein
11	group_743	hypothetical protein
11	group_9527	hypothetical protein
11	group_2289	hypothetical protein
11	group_1963	hypothetical protein
11	group_7866	hypothetical protein
11	group_3723	hypothetical protein
11	group_2591	IS3 family transposase ISBcen15;hypothetical protein
11	group_11161	hypothetical protein
11	group_10763	hypothetical protein
11	group_10531	hypothetical protein
11	group_5046	hypothetical protein
11	group_8961	hypothetical protein
11	group_8375	hypothetical protein
11	group_7951	hypothetical protein
11	group_7946	hypothetical protein
11	group_352	hypothetical protein
11	iolG_1~~~iolG_3~~~iolG_4	Inositol 2-dehydrogenase/D-chiro-inositol 3-dehydrogenase;hypothetical protein
11	gcvA_7~~~gcvA_5	Glycine cleavage system transcriptional activator
11	group_7446	hypothetical protein
11	ssuD_2	Alkanesulfonate monooxygenase
11	dhpH	hypothetical protein;2,6-dihydroxypyridine 3-monooxygenase
11	dmlR_16	HTH-type transcriptional regulator DmlR
11	group_10487	hypothetical protein
11	group_7345	Cystathionine beta-lyase PatB
11	cyt	Cytochrome c-552;hypothetical protein
11	group_3128	hypothetical protein
11	group_3102	Cytochrome c-552
11	group_2960	hypothetical protein
11	group_10429	hypothetical protein

11	group_5516	hypothetical protein
11	jefA~~~stp~~~mdtD_1	hypothetical protein;Drug efflux pump JefA;Multidrug resistance protein Stp;Putative multidrug resistance protein MdtD
11	group_11168	hypothetical protein
11	group_3868	hypothetical protein
11	group_2769	IS3 family transposase ISRso14;IS3 family transposase ISAtu5;hypothetical protein
11	group_265	hypothetical protein
11	group_9197	hypothetical protein
11	group_7412	hypothetical protein
11	group_7071	hypothetical protein
11	lrpC_3~~~lrpC_4	hypothetical protein;HTH-type transcriptional regulator LrpC
11	group_2447	hypothetical protein
11	group_2275	hypothetical protein
11	group_7863	hypothetical protein
11	group_3808	hypothetical protein
11	group_1614	hypothetical protein
	adhC2~~~yahK_1~~~tdh_2~~~adhA_2~~~	NADP-dependent alcohol dehydrogenase C 2;Aldehyde reductase YahK;L-threonine 3-dehydrogenase;hypothetical protein;putative formaldehyde
11	adhA_3	dehydrogenase AdhA
11	group_2617	hypothetical protein
11	group_7956	hypothetical protein
11	group_7737	hypothetical protein
11	group_7563	hypothetical protein
11	group_4127	hypothetical protein
11	group_2973	Putative universal stress protein;Universal stress protein/MSMEI_3859
11	group_2751	hypothetical protein;IS3 family transposase IS222
11	group_10714	Vitamin B12-dependent ribonucleoside-diphosphate reductase;hypothetical protein
11	group_8236	hypothetical protein
11	group_7050	hypothetical protein
11	group_6923	hypothetical protein
11	group_5856	Alcohol dehydrogenase

		hypothetical protein;putative multidrug resistance protein
	emrY_2~~~emrY_1~~~emrB_4~~~emrB_2	EmrY;Multidrug export protein
11	~~~emrB_1	EmrB
11	iorA_1~~~iorA_3~~~iorA_2	Isoquinoline 1-oxidoreductase subunit alpha
11	group_3751	Isoquinoline 1-oxidoreductase subunit beta;hypothetical protein
11	paoD_2~~~pucA_2~~~paoD~~~pucA	Molybdenum cofactor insertion chaperone PaoD;putative xanthine dehydrogenase subunit A
11	group_3630	hypothetical protein
11	dksA_3~~~dksA_2~~~dksA_1	RNA polymerase-binding transcription factor DksA
11	group_1852	hypothetical protein
11	group_1162	hypothetical protein
11	prs_2~~~prs_1	Ribose-phosphate pyrophosphokinase;Putative ribose-phosphate pyrophosphokinase
11	adh_2~~~adh_1	Alcohol dehydrogenase
11	group_7775	hypothetical protein
11	group_5199	hypothetical protein
11	group_5013	IS3 family transposase ISBmu5;IS3 family transposase ISRme12;hypothetical protein
11	mas~~~tdh	hypothetical protein;Mycocerosic acid synthase;L-threonine 3-dehydrogenase
11	group_4049	hypothetical protein
11	group_10479	hypothetical protein
11	farA	Fatty acid resistance protein FarA
11	oprM_7~~~oprM_1~~~oprM_9	Outer membrane protein OprM;hypothetical protein
11	group_4410	hypothetical protein
11	group_2424	hypothetical protein
11	group_2383	hypothetical protein
11	group_2136	hypothetical protein
11	group_2122	hypothetical protein
11	group_2101	hypothetical protein
11	group_2057	hypothetical protein
11	group_1978	hypothetical protein
11	group_10310	IS3 family transposase ISRso11
11	group_7917	hypothetical protein
11	group_7915	hypothetical protein
11	group_7892	hypothetical protein

11	mrpX	Methionine-rich peptide X;hypothetical protein
11	group_3219	hypothetical protein
11	group_2894	hypothetical protein
11	group_2779	putative multidrug-efflux transporter
11	group_2215	hypothetical protein
11	group_2166	hypothetical protein
11	group_2145	hypothetical protein
11	group_1892	hypothetical protein
11	yedY1	Putative protein-methionine-sulfoxide reductase subunit YedZ1
11	group_9578	hypothetical protein
11	group_9384	hypothetical protein
11	rhaS_2~~~rhaS_3~~~rhaS_4	HTH-type transcriptional activator RhaS
11	group_3753	putative zinc-binding alcohol dehydrogenase;hypothetical protein
11	group_3247	hypothetical protein;IS3 family transposase ISPsy24
11	group_3150	hypothetical protein
11	yecD_1~~~yecD_2~~~yecD~~~sbnD_2~~~sbnD_1~~~mdtG	Isochorismatase family protein YecD;Staphyloferrin B transporter;hypothetical protein;N-carbamoylsarcosine amidase;Multidrug resistance protein MdtG
11	rspR_3~~~rspR_1	hypothetical protein;HTH-type transcriptional repressor RspR
11	yedZ1	Putative protein-methionine-sulfoxide reductase subunit YedZ1
11	group_2005	hypothetical protein
11	group_1922	hypothetical protein
11	group_1921	hypothetical protein
11	group_1862	hypothetical protein
11	group_1803	hypothetical protein
11	birA_1	hypothetical protein;Bifunctional ligase/repressor BirA
11	group_1773	hypothetical protein
11	group_1740	hypothetical protein
11	lrpC_2	hypothetical protein;HTH-type transcriptional regulator LrpC
11	group_1251	hypothetical protein
11	group_3592	hypothetical protein
11	group_3148	hypothetical protein

11	group_2841	hypothetical protein
11	group_2531	hypothetical protein
11	group_1664	hypothetical protein
11	group_7766	Flavin-dependent monooxygenase oxygenase subunit HsaA
11	group_11862	hypothetical protein
11	group_7658	hypothetical protein
11	group_2058	hypothetical protein
11	group_9494	hypothetical protein
11	group_8007	hypothetical protein;ATP-dependent RecD-like DNA helicase
11	group_7869	hypothetical protein
11	group_3144	hypothetical protein
11	group_1454	hypothetical protein
11	yqcF~~~yqcF_2	Antitoxin YqcF;hypothetical protein
11	group_11565	hypothetical protein
11	group_10188	hypothetical protein
11	nuoF_1~~~nuoF_3~~~nuoF_2	NADH-quinone oxidoreductase subunit F
11	aspT	Aspartate/alanine antiporter;hypothetical protein
11	sdhA_1~~~sdhA_2	Succinate dehydrogenase flavoprotein subunit;hypothetical protein
11	rcsC_11~~~rcsC_5~~~lldD~~~dctB_1	Sensor histidine kinase RcsC;L-lactate dehydrogenase;C4-dicarboxylate transport sensor protein DctB
11	group_4961	hypothetical protein
11	group_3223	hypothetical protein
11	maeA~~~maeA_2~~~maeA_1~~~mleS	NAD-dependent malic enzyme;putative NAD-dependent malic enzyme 2;Malolactic enzyme
11	dmlR_19	HTH-type transcriptional regulator DmlR
11	group_7809	hypothetical protein
11	yidZ~~~pcpR_7	hypothetical protein;HTH-type transcriptional regulator YidZ;PCP degradation transcriptional activation protein
11	group_6202	hypothetical protein
11	group_4831	hypothetical protein
11	group_9592	Transcriptional regulator SlyA;hypothetical protein
11	group_9256	hypothetical protein

11	dctD_1~~~dctD_2	C4-dicarboxylate transport transcriptional regulatory protein DctD
11	group_7266	hypothetical protein
11	group_4833	hypothetical protein
11	group_1976	hypothetical protein
11	group_1089	hypothetical protein
11	group_11541	hypothetical protein
11	group_7326	hypothetical protein
11	group_6657	hypothetical protein;IS5 family transposase ISRso1
11	group_5408	hypothetical protein
11	group_5314	hypothetical protein
11	group_2656	hypothetical protein
11	group_11922	hypothetical protein
11	group_2575	IS5 family transposase ISRso1
11	group_12046	hypothetical protein
11	mhpD	2-keto-4-pentenoate hydratase;hypothetical protein
11	dsbA_2~~~dsbA_1	Thiol:disulfide interchange protein DsbA;hypothetical protein
11	amnE	4-oxalocrotonate decarboxylase
11	dmpI	2-hydroxymuconate tautomerase
11	group_9817	hypothetical protein
11	group_9169	RNA 2'3'-cyclic phosphodiesterase
11	amnF	2-oxopent-4-enoate hydratase
11	group_8702	3-hydroxy-2-methylpyridine-45-dicarboxylate 4-decarboxylase;hypothetical protein
11	mhpF~~~dmpF	Acetaldehyde dehydrogenase
11	mhbT_2~~~mhbT_1	3-hydroxybenzoate transporter MhbT
11	xylG_2~~~xylG_1~~~betB_5	2-hydroxymuconic semialdehyde dehydrogenase;NAD/NADP-dependent betaine aldehyde dehydrogenase
11	hsaC	Iron-dependent extradiol dioxygenase
11	mhpE_2~~~mhpE_1~~~mhpE	4-hydroxy-2-oxovalerate aldolase
11	group_2535	hypothetical protein
11	intS_1~~~intS~~~intS_2	Prophage integrase IntS;hypothetical protein
11	group_2008	hypothetical protein
11	group_531	hypothetical protein
11	group_507	hypothetical protein

11	group_356	hypothetical protein
11	group_274	hypothetical protein
11	group_198	hypothetical protein
11	group_12045	hypothetical protein
11	rstA_2	Transcriptional regulatory protein RstA;hypothetical protein
11	group_7921	hypothetical protein
11	group_7885	hypothetical protein
11	group_7883	hypothetical protein
11	group_7662	hypothetical protein
11	group_3482	hypothetical protein
11	amnD~~~rutC_2	2-aminomuconate deaminase;RutC family protein;Putative aminoacrylate peracid reductase RutC
11	group_2190	hypothetical protein
11	group_2150	hypothetical protein
11	group_2149	hypothetical protein
11	group_2108	hypothetical protein
11	group_2078	hypothetical protein
11	group_9341	hypothetical protein
11	group_9329	hypothetical protein
11	group_7740	hypothetical protein
11	kgtP_5~~~nanT_2	hypothetical protein;IS5 family transposase ISRso1;Alpha-ketoglutarate permease;Sialic acid transporter NanT
11	group_2952	hypothetical protein
11	cya_1	hypothetical protein;Bifunctional hemolysin/adenylate cyclase
11	slyA_1~~~slyA_3	Transcriptional regulator SlyA
11	group_6396	hypothetical protein
11	group_3781	hypothetical protein
11	livQ	6'''-hydroxyparomomycin C oxidase
11	group_3775	hypothetical protein
11	group_3023	hypothetical protein
11	group_785	hypothetical protein
11	group_78	hypothetical protein
11	group_10316	hypothetical protein
11	group_7207	hypothetical protein
11	group_1254	hypothetical protein
11	group_1069	hypothetical protein
11	group_575	hypothetical protein
11	group_9396	hypothetical protein

11	group_2075	hypothetical protein
11	group_4096	hypothetical protein
11	group_9150	hypothetical protein
11	group_7633	hypothetical protein
11	group_3697	hypothetical protein
11	group_3447	hypothetical protein
11	group_2551	hypothetical protein
11	group_426	hypothetical protein
11	group_7895	hypothetical protein
11	group_2249	hypothetical protein
11	group_607	hypothetical protein
11	group_11748	hypothetical protein
11	group_9438	hypothetical protein
11	group_9064	hypothetical protein
11	group_7911	hypothetical protein
11	group_6858	hypothetical protein
11	group_2621	IS3 family transposase ISAisp2
11	group_2510	IS110 family transposase ISBcen4
11	group_11923	hypothetical protein
11	gcvA_6	Glycine cleavage system transcriptional activator
11	group_8043	Alcohol dehydrogenase
11	group_7051	hypothetical protein
11	group_6429	hypothetical protein
11	rfnT_2~~~rfnT	Riboflavin transporter RfnT
11	glnQ_4~~~glnQ_5	Glutamine transport ATP-binding protein GlnQ
11	group_4453	hypothetical protein
11	group_2860	hypothetical protein;IS5 family transposase ISRso18;IS5 family transposase IS1405
11	group_925	hypothetical protein
11	group_776	hypothetical protein
11	xerC_2~~~xerC_3~~~xerC_4~~~xerC_1	Tyrosine recombinase XerC;hypothetical protein
11	group_8187	Competence protein ComM
11	group_7910	hypothetical protein
11	group_5039	hypothetical protein
11	lldD_1~~~lldD_2~~~lldD	L-lactate dehydrogenase
11	group_4006	hypothetical protein
11	group_3765	hypothetical protein
11	group_11800	hypothetical protein
11	group_10427	Carboxymethylenebutenolidase
11	group_10273	hypothetical protein

11	group_9328	hypothetical protein
11	group_7927	hypothetical protein
11	group_3712	hypothetical protein
11	group_2635	hypothetical protein
11	group_2223	IS3 family transposase ISBcen15;IS3 family transposase ISRso12
11	group_954	hypothetical protein
11	group_9541	hypothetical protein
11	speE_2~~~speE_3	Polyamine aminopropyltransferase
11	group_1351	hypothetical protein
11	group_10626	hypothetical protein
11	group_9248	hypothetical protein;IS4 family transposase ISCro3
11	group_7320	hypothetical protein
11	group_6965	hypothetical protein
11	group_1840	hypothetical protein
11	group_1731	hypothetical protein
11	group_898	hypothetical protein
11	group_2905	hypothetical protein
11	group_2838	hypothetical protein
11	group_2401	hypothetical protein
11	group_2396	hypothetical protein
11	group_2293	IS3 family transposase ISBcen15;hypothetical protein
11	group_3754	hypothetical protein
11	group_2113	hypothetical protein
11	group_11234	hypothetical protein
11	group_10557	hypothetical protein
11	group_3770	hypothetical protein
11	group_2440	hypothetical protein;IS4 family transposase ISCro3
11	group_2330	IS30 family transposase IS1382;IS30 family transposase IST3091;IS30 family transposase ISRme10
11	group_10279	hypothetical protein
11	group_9326	hypothetical protein
11	group_8323	hypothetical protein
11	group_7721	hypothetical protein
11	group_2397	hypothetical protein
11	group_2274	hypothetical protein
11	group_1248	hypothetical protein
11	group_1044	hypothetical protein
11	group_765	hypothetical protein
11	group_10334	hypothetical protein

11	group_9319	hypothetical protein
11	group_7120	hypothetical protein
11	group_2987	IS66 family transposase ISPpu19;hypothetical protein;IS66 family transposase ISBcen19
11	group_2549	hypothetical protein
11	group_11000	hypothetical protein
11	group_10223	IS4 family transposase ISCro3
11	group_10163	hypothetical protein
11	group_7703	IS5 family transposase IS1405;hypothetical protein
11	group_2832	IS4 family transposase ISCro3;hypothetical protein
11	group_2804	hypothetical protein
11	group_2391	IS3 family transposase ISAisp2;Putative transposase InsK for insertion sequence element IS150
11	group_1328	hypothetical protein
11	group_1098	IS3 family transposase ISRso11
11	group_1050	hypothetical protein
11	group_9437	hypothetical protein
11	group_8921	IS3 family transposase ISAisp2
11	group_7727	hypothetical protein
11	group_6609	hypothetical protein;IS3 family transposase ISAisp2
11	group_5344	hypothetical protein;IS4 family transposase ISCro3
11	xerC_6~~~xerC_5~~~xerD_1~~~xerD_2~~~xerC_4	Tyrosine recombinase XerC;Tyrosine recombinase XerD
11	group_3217	hypothetical protein
11	group_2896	hypothetical protein
11	group_2624	hypothetical protein
11	group_2587	hypothetical protein
11	group_2586	hypothetical protein
11	group_1437	hypothetical protein
11	group_697	hypothetical protein
11	group_10218	IS4 family transposase ISCro3
11	group_9141	hypothetical protein
11	group_7265	hypothetical protein
11	group_4864	IS5 family transposase IS1405
11	group_1663	hypothetical protein
11	group_11447	hypothetical protein
11	group_9949	hypothetical protein
11	group_8824	hypothetical protein

11	group_7140	hypothetical protein
11	group_7073	hypothetical protein
11	group_4583	hypothetical protein
11	group_4047	hypothetical protein
11	group_3686	hypothetical protein
11	group_2699	hypothetical protein
11	group_2388	hypothetical protein
11	group_1380	hypothetical protein
11	group_730	hypothetical protein
11	group_497	hypothetical protein
11	group_62	hypothetical protein
11	group_11226	hypothetical protein
11	group_11149	hypothetical protein
11	group_10222	IS4 family transposase ISCro3
11	group_8180	IS4 family transposase ISPhsp1;hypothetical protein;IS4 family transposase ISCro3
11	group_7415	hypothetical protein
11	group_1870	hypothetical protein
11	group_1634	IS3 family transposase ISBmu5;hypothetical protein
11	group_1621	hypothetical protein
11	group_1066	hypothetical protein
11	mntH_1	Divalent metal cation transporter MntH
11	group_11166	hypothetical protein
11	group_10822	hypothetical protein
11	group_10431	hypothetical protein
11	group_9546	hypothetical protein
11	group_8005	hypothetical protein
11	group_7865	IS5 family transposase ISRso18;IS5 family transposase IS1405
11	group_7729	hypothetical protein
11	group_7728	hypothetical protein
11	group_7330	hypothetical protein
11	group_5659	hypothetical protein
11	group_5545	hypothetical protein
11	group_4874	hypothetical protein;IS110 family transposase ISBcen4
11	group_4375	hypothetical protein
11	group_4374	hypothetical protein
11	group_2358	IS5 family transposase ISRso18;hypothetical protein;IS5 family transposase IS1405

		IS1182 family transposase
		ISBusp4;IS1182 family transposase
11	group_1184	ISPpu16;hypothetical protein
		IS256 family transposase
11	group_767	ISRso7;hypothetical protein
11	group_224	hypothetical protein
11	group_11540	hypothetical protein
11	group_11067	hypothetical protein
11	group_10825	hypothetical protein
11	group_10809	hypothetical protein
11	group_10284	hypothetical protein
11	group_10231	IS4 family transposase ISCro3
11	group_9359	hypothetical protein
11	group_9121	hypothetical protein
11	group_7329	hypothetical protein
		IS3 family transposase
11	group_7322	ISRso16;hypothetical protein
		hypothetical protein;Putative
11	rhsC_3~~~rhsC_1	deoxyribonuclease RhsC
		IS110 family transposase
11	group_2977	ISBma3;hypothetical protein
		hypothetical protein;IS4 family
11	group_2527	transposase ISCro3
11	group_2452	hypothetical protein
11	group_1688	hypothetical protein
11	group_989	hypothetical protein
11	group_454	hypothetical protein
11	group_10807	hypothetical protein
11	group_10614	hypothetical protein
		Branched-chain-amino-acid
11	group_10491	aminotransferase
11	ribZ	Riboflavin transporter RibZ
11	group_5820	hypothetical protein
11	group_5776	hypothetical protein
11	group_4372	hypothetical protein
11	group_3308	hypothetical protein
		HTH-type transcriptional regulator
11	benM_4~~~benM_2	BenM
11	group_1564	hypothetical protein
11	group_1492	hypothetical protein
11	group_10767	hypothetical protein
		IS4 family transposase
11	group_5573	ISCro3;hypothetical protein
11	group_4854	hypothetical protein
11	group_4837	hypothetical protein

11	group_4588	hypothetical protein
11	group_4526	hypothetical protein
11	group_4241	hypothetical protein
11	group_4074	hypothetical protein
11	group_3973	hypothetical protein
11	group_334	hypothetical protein
11	group_11390	hypothetical protein
11	group_10267	hypothetical protein
11	group_7964	hypothetical protein
11	group_7017	IS110 family transposase ISBcen4;IS110 family transposase ISPa49;hypothetical protein
11	group_3064	IS1182 family transposase ISBusp4
11	group_2453	IS5 family transposase IS1405
11	group_6571	hypothetical protein
11	group_4355	IS5 family transposase IS1405
11	group_7961	hypothetical protein;IS4 family transposase ISCro3
11	group_7186	IS5 family transposase ISRso1
11	group_4208	hypothetical protein
11	group_3720	hypothetical protein
11	group_3716	IS5 family transposase IS1405
11	group_2552	IS3 family transposase ISAisp2;hypothetical protein
11	group_728	hypothetical protein
11	group_11389	hypothetical protein
11	group_10747	hypothetical protein
11	group_10627	hypothetical protein
11	group_10376	IS110 family transposase ISMno14
11	group_10242	IS3 family transposase ISBam2
11	group_10240	IS3 family transposase ISRso14
11	group_9865	Sarcosine oxidase subunit gamma;hypothetical protein
11	soxA_2~~~soxA_4	Sarcosine oxidase subunit alpha
11	group_7708	hypothetical protein
11	group_6747	hypothetical protein
11	soxB	Sarcosine oxidase subunit beta
11	sdaB	L-serine dehydratase 2
11	group_6193	hypothetical protein
11	group_4781	hypothetical protein
11	group_4369	hypothetical protein
11	group_2553	IS3 family transposase ISAisp2
11	cdhR_1~~~cdhR_4	HTH-type transcriptional regulator CdhR

11	ydeP_1~~~ydeP_2	Protein YdeP
11	group_11004	hypothetical protein
11	group_10940	hypothetical protein
11	group_10639	IS5 family transposase IS1405
11	group_10633	IS5 family transposase IS1405
11	group_10629	hypothetical protein
11	opuAB	Glycine betaine transport system permease protein OpuAB
11	group_10077	hypothetical protein
11	group_9003	hypothetical protein
11	cdhR_6	HTH-type transcriptional regulator CdhR
11	cdhR_4	HTH-type transcriptional regulator CdhR
11	group_7512	hypothetical protein;IS5 family transposase IS1405
11	flhA_3	S-(hydroxymethyl)glutathione dehydrogenase
11	group_7106	NADPH oxidoreductase
11	group_7105	Outer membrane porin protein
11	group_7023	Glycine betaine/proline betaine transport system ATP-binding protein ProV
11	fixB	Protein FixB
11	fdhA	Glutathione-independent formaldehyde dehydrogenase
11	group_6440	Carnitine monooxygenase oxygenase subunit
11	group_6288	Serine hydroxymethyltransferase 2
11	group_6148	hypothetical protein
11	naiP_3	Putative niacin/nicotinamide transporter NaiP
11	group_4365	hypothetical protein
11	group_4313	hypothetical protein
11	group_4010	hypothetical protein
11	stcD	putative N-methylproline demethylase
11	group_3446	IS110 family transposase ISMno14
11	group_2919	hypothetical protein
11	purU_3	Formyltetrahydrofolate deformylase
11	group_1207	hypothetical protein
11	group_1094	IS5 family transposase IS1405
11	group_1093	hypothetical protein
11	group_1018	hypothetical protein

11	moaA_2	GTP 3'8-cyclase;hypothetical protein
11	group_667	hypothetical protein
11	group_11726	IS5 family transposase IS1405
11	group_10645	IS5 family transposase IS1405
11	group_10634	IS5 family transposase IS1405
11	group_10622	hypothetical protein
11	group_10220	IS4 family transposase ISCro3
11	group_9152	hypothetical protein;IS5 family transposase IS1405
11	group_7726	hypothetical protein
11	group_5393	hypothetical protein
11	group_4366	hypothetical protein
11	group_3603	hypothetical protein
11	group_3594	hypothetical protein
11	group_3527	hypothetical protein
11	group_3522	hypothetical protein
11	group_3365	hypothetical protein
11	group_3235	ISNCY family transposase ISBcen27
11	group_1895	TAL effector protein Rip19
11	group_899	hypothetical protein
11	group_10640	IS5 family transposase IS1405
11	group_10617	hypothetical protein
11	group_10379	IS110 family transposase ISMno14
11	group_883	hypothetical protein;IS5 family transposase IS1405
11	group_10692	hypothetical protein
11	group_10656	IS21 family transposase ISRso6
11	amnD	2-aminomuconate deaminase
11	group_8011	hypothetical protein
11	group_7321	hypothetical protein;IS5 family transposase IS1405
11	group_4361	hypothetical protein
11	group_1788	IS5 family transposase IS1405;IS5 family transposase ISRso18
11	rhtB_3	Homoserine/homoserine lactone efflux protein
11	group_10224	IS4 family transposase ISCro3
11	group_7183	IS3 family transposase ISRso10
11	group_5082	hypothetical protein
11	group_9625	hypothetical protein
10	group_3323	hypothetical protein
10	group_3120	hypothetical protein

10	cpdA_5~~~cpdA_4~~~cpdA_3	3'5'-cyclic adenosine monophosphate phosphodiesterase CpdA
10	group_1713	queuosine precursor transporter;hypothetical protein
10	group_1900	hypothetical protein
10	group_1484	hypothetical protein;putative HTH-type transcriptional regulator
9	group_2543	hypothetical protein;IS66 family transposase ISPa82;IS66 family transposase ISAeh1
9	group_2464	IS66 family transposase IS1313
9	group_10466	hypothetical protein
8	group_12012	IS701 family transposase ISRso17
8	group_5593	IS701 family transposase ISRso17
7	group_3715	hypothetical protein
7	group_11742	hypothetical protein
6	group_2238	hypothetical protein
6	group_2237	Lysozyme RrrD
6	group_2236	hypothetical protein
6	xerC_3~~~xerC_5~~~xerC_2~~~xerC_4	Tyrosine recombinase XerC
5	group_2515	IS5 family transposase ISBmu20;IS5 family transposase IS1021;IS5 family transposase ISRso18;IS5 family transposase IS1405;hypothetical protein
5	group_2724	IS5 family transposase IS1405;hypothetical protein;IS5 family transposase ISRso18;IS5 family transposase IS1021
5	group_1046	hypothetical protein;IS5 family transposase IS1021
5	group_3127	IS5 family transposase ISRso18;IS5 family transposase IS1405;IS5 family transposase IS1021;hypothetical protein
5	group_1202	hypothetical protein;IS5 family transposase ISBmu2;IS5 family transposase IS1021
5	group_2889	IS5 family transposase IS1021;hypothetical protein
5	group_1195	hypothetical protein;IS5 family transposase IS1021
5	group_977	IS5 family transposase IS1021;hypothetical protein
5	group_3187	IS5 family transposase IS1021

		IS5 family transposase
5	group_828	IS1021;hypothetical protein
		IS5 family transposase
5	group_1457	IS1021;hypothetical protein
		IS5 family transposase
5	group_1322	IS1021;hypothetical protein
		IS5 family transposase
5	group_3768	IS1021;hypothetical protein
5	group_3099	IS5 family transposase IS1021
		IS5 family transposase
5	group_2320	IS1021;hypothetical protein
5	group_1030	IS5 family transposase IS1021
		IS5 family transposase
5	group_796	IS1021;hypothetical protein
		IS5 family transposase
5	group_1527	IS1021;hypothetical protein
5	group_1142	IS5 family transposase IS1021
		IS5 family transposase
5	group_3709	IS1021;hypothetical protein
		IS5 family transposase
5	group_2325	IS1021;hypothetical protein
5	group_1128	IS5 family transposase IS1021
		IS5 family transposase
5	group_2698	IS1021;hypothetical protein
4	group_11873	hypothetical protein
4	group_11586	hypothetical protein
4	group_10342	hypothetical protein
4	group_5626	hypothetical protein
		cGAMP-activated
4	capV	phospholipase;hypothetical protein
4	group_4325	hypothetical protein
4	group_4197	hypothetical protein
3	group_11485	hypothetical protein
3	group_10978	hypothetical protein
3	group_10130	hypothetical protein
2	group_3722	IS3 family transposase ISRso11
		IS3 family transposase
2	group_7616	ISBcen10;hypothetical protein
1	group_1670	hypothetical protein
1	group_1319	hypothetical protein

9. REFERENCES

Abedon, S.T., 2009. Chapter 1 Phage Evolution and Ecology, in: Advances in Applied Microbiology. Academic Press, pp. 1–45. https://doi.org/10.1016/S0065-2164(08)01001-0

Addy, H.S., Askora, A., Kawasaki, T., Fujie, M., Yamada, T., 2012. The filamentous phage ϕrSS1 enhances virulence of phytopathogenic Ralstonia solanacearum on tomato. Phytopathology 102, 244–251. https://doi.org/10.1094/PHYTO-10-11-0277

Albuquerque, G.M.R., Santos, L.A., Felix, K.C.S., Rollemberg, C.L., Silva, A.M.F., Souza, E.B., Cellier, G., Prior, P., Mariano, R.L.R., 2014. Moko disease-causing strains of Ralstonia solanacearum from Brazil extend known diversity in paraphyletic phylotype II. Phytopathology 104, 1175–1182. https://doi.org/10.1094/PHYTO-12-13-0334-R

Alderley, C.L., Greenrod, S.T.E., Friman, V.-P., 2022. Plant pathogenic bacterium can rapidly evolve tolerance to an antimicrobial plant allelochemical. Evol. Appl. 15, 735–750. https://doi.org/10.1111/eva.13363

Aldon, D., Brito, B., Boucher, C., Genin, S., 2000. A bacterial sensor of plant cell contact controls the transcriptional induction of Ralstonia solanacearum pathogenicity genes. EMBO J. 19, 2304–2314. https://doi.org/10.1093/emboj/19.10.2304

Alfano, J.R., Collmer, A., 1996. Bacterial Pathogens in Plants: Life up against the Wall. Plant Cell 8, 1683–1698. https://doi.org/10.1105/tpc.8.10.1683

Álvarez, B., López, M.M., Biosca, E.G., 2008. Survival strategies and pathogenicity of Ralstonia solanacearum phylotype II subjected to prolonged starvation in environmental water microcosms. Microbiology 154, 3590–3598. https://doi.org/10.1099/mic.0.2008/019448-0

Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., Liang, Y., Wishart, D.S., 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 44, W16–W21. https://doi.org/10.1093/nar/gkw387

Arnold, B.J., Huang, I.-T., Hanage, W.P., 2022. Horizontal gene transfer and adaptive evolution in bacteria. Nat. Rev. Microbiol. 20, 206–218. https://doi.org/10.1038/s41579-021-00650-4

Azad, R.K., Lawrence, J.G., 2011. Towards more robust methods of alien gene detection. Nucleic Acids Res. 39, e56. https://doi.org/10.1093/nar/gkr059

Azarian, T., Huang, I.-T., Hanage, W.P., 2020. Structure and Dynamics of Bacterial Populations: Pangenome Ecology, in: Tettelin, H., Medini, D. (Eds.), The Pangenome: Diversity, Dynamics and Evolution of Genomes. Springer, Cham (CH).

Bargués-Ribera, M., Gokhale, C.S., 2020. Eco-evolutionary agriculture: Host-pathogen dynamics in crop rotations. PLOS Comput. Biol. 16, e1007546. https://doi.org/10.1371/journal.pcbi.1007546

Baroukh, C., Zemouri, M., Genin, S., 2022. Trophic preferences of the pathogen Ralstonia solanacearum and consequences on its growth in xylem sap. MicrobiologyOpen 11, e1240. https://doi.org/10.1002/mbo3.1240

Bart, R.S., Cohn, M., Kassen, A., McCallum, E., Shybut, M., Petriello, A., Krasileva, K., Dahlbeck, D., Medina, C., Alicai, T., Kumar, L., Moreira, L., Neto, J., Verdier, V., Santana, M., Bernal, A., Staskawicz, B.J., 2012. High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. PNAS. https://doi.org/10.1073/pnas.1208003109

Bharath, P., Gahir, S., Raghavendra, A.S., 2021. Abscisic Acid-Induced Stomatal Closure: An Important Component of Plant Defense Against Abiotic and Biotic Stress. Front. Plant Sci. 12.

Bhatt, G., Denny, T.P., 2004. Ralstonia solanacearum iron scavenging by the siderophore staphyloferrin B is controlled by PhcA, the global virulence regulator. J. Bacteriol. 186, 7896–7904. https://doi.org/10.1128/JB.186.23.7896-7904.2004

Bobay, L.-M., Ochman, H., 2017. Biological Species Are Universal across Life's Domains. Genome Biol. Evol. 9, 491–501. https://doi.org/10.1093/gbe/evx026

Bobay, L.-M., Raymann, K., 2019. Population Genetics of Host-Associated Microbiomes. Curr. Mol. Biol. Rep. 5, 128–139. https://doi.org/10.1007/s40610-019-00122-y

Bocsanczy, A.M., Achenbach, U.C.M., Mangravita-Novo, A., Chow, M., Norman, D.J., 2014. Proteomic comparison of Ralstonia solanacearum strains reveals temperature dependent virulence factors. BMC Genomics 15, 280. https://doi.org/10.1186/1471-2164-15-280

Bocsanczy, A.M., Huguettapia, J.C., Norman, D.J., 2017. Comparative genomics of Ralstonia solanacearum identifies candidate genes associated with cool Virulence. Front. Plant Sci. 8, 1565.

Bové, J.M., 2006. Huanglongbing: a destructive, newly-emerging, century-old disease of citrus. J. Plant Pathol. 88, 7–37.

Bragard, C., Dehnen-Schmutz, K., Di Serio, F., Gonthier, P., Jaques Miret, J.A., Justesen, A.F., MacLeod, A., Magnusson, C.S., Milonas, P., Navas-Cortes, J.A., Parnell, S., Potting, R., Reignault, P.L., Thulke, H., Van der Werf, W., Vicent Civera, A., Yuen, J., Zappalà, L., Van der Wolf, J., Kaluski, T., Pautasso, M., Jacques, M., 2019. Pest categorisation of the Ralstonia solanacearum species complex. EFSA J. 17. https://doi.org/10.2903/j.efsa.2019.5618

Brockhurst, M.A., Harrison, E., Hall, J.P.J., Richards, T., Mcnally, A., Maclean, C., 2019. The Ecology and Evolution of Pangenomes. https://doi.org/10.1016/j.cub.2019.08.012

Buchfink, B., Reuter, K., Drost, H.-G., 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat. Methods 18, 366–368. https://doi.org/10.1038/s41592-021-01101-x

Buckling, A., Craig Maclean, R., Brockhurst, M.A., Colegrave, N., 2009. The Beagle in a bottle. Nature 457, 824–829. https://doi.org/10.1038/nature07892

Buddenhagen, I., 1962. Designation of races in Pseudomonas solanacearum. Phytopathology 52, 726.

Buddenhagen, I., Kelman, A., 1964. Biological and Physiological Aspects of Bacterial Wilt Caused by Pseudomonas Solanacearum. Annu. Rev. Phytopathol. 2, 203–230. https://doi.org/10.1146/annurev.py.02.090164.001223

Cantalapiedra, C.P., Hernández-Plaza, A., Letunic, I., Bork, P., Huerta-Cepas, J., 2021. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. Mol. Biol. Evol. 38, 5825–5829. https://doi.org/10.1093/molbev/msab293

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinforma. Oxf. Engl. 25, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T.,

Zaneveld, J., Knight, R., 2010. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7, 335–336. https://doi.org/10.1038/nmeth.f.303

Caruso, P., Palomo, J.L., Bertolini, E., Álvarez, B., López, M.M., Biosca, E.G., 2005a. Seasonal Variation of Ralstonia solanacearum Biovar 2 Populations in a Spanish River: Recovery of Stressed Cells at Low Temperatures. Appl. Environ. Microbiol. 71, 140–148. https://doi.org/10.1128/AEM.71.1.140-148.2005

Caruso, P., Palomo, J.L., Bertolini, E., Álvarez, B., López, M.M., Biosca, E.G., 2005b. Seasonal Variation of Ralstonia solanacearum Biovar 2 Populations in a Spanish River: Recovery of Stressed Cells at Low Temperatures. Appl. Environ. Microbiol. 71, 140–148. https://doi.org/10.1128/AEM.71.1.140-148.2005

Castillo, J.A., Secaira-Morocho, H., Maldonado, S., Sarmiento, K.N., 2020. Diversity and Evolutionary Dynamics of Antiphage Defense Systems in Ralstonia solanacearum Species Complex. Front. Microbiol. 11. https://doi.org/10.3389/fmicb.2020.00961

Cellier, G, Prior, P., 2010. Deciphering Phenotypic Diversity of Ralstonia solanacearum Strains Pathogenic to Potato. https://doi.org/10.1094/PHYTO-02-10-0059

Cellier, G., Prior, P., 2010. Deciphering phenotypic diversity of Ralstonia solanacearum strains pathogenic to potato. Phytopathology 100, 1250–1261. https://doi.org/10.1094/PHYTO-02-10-0059

Cellier, G., Remenant, B., Chiroleu, F., Lefeuvre, P., Prior, P., 2012. Phylogeny and population structure of brown rot- and Moko disease-causing strains of Ralstonia solanacearum phylotype II. Appl. Environ. Microbiol. 78, 2367–2375. https://doi.org/10.1128/AEM.06123-11

Chakraborty, S., Newton, A.C., 2011. Climate change, plant diseases and food security: an overview. Plant Pathol. 60, 2–14. https://doi.org/10.1111/j.1365-3059.2010.02411.x

Chaloner, T.M., Gurr, S.J., Bebber, D.P., 2021. Plant pathogen infection risk tracks global crop yields under climate change. Nat. Clim. Change 11, 710–715. https://doi.org/10.1038/s41558-021-01104-8

Champoiseau, P.G., Jones, J.B., Allen, C., 2009. Ralstonia solanacearum Race 3 Biovar 2 Causes Tropical Losses and Temperate Anxieties. Plant Health Prog. 10, 35. https://doi.org/10.1094/php-2009-0313-01-rv

Chang, J.H., Desveaux, D., Creason, A.L., 2014. The ABCs and 123s of Bacterial Secretion Systems in Plant Pathogenesis. Annu. Rev. Phytopathol. 52, 317–345. https://doi.org/10.1146/annurev-phyto-011014-015624

Chesneau, T., Maignien, G., Boyer, C., Chéron, J.-J., Roux-Cuvelier, M., Vanhuffel, L., Poussier, S., Prior, P., 2017. Sequevar diversity and virulence of Ralstonia solanacearum Phylotype I on Mayotte island (Indian Ocean). Front. Plant Sci. 8, 2209. https://doi.org/10.3389/fpls.2017.02209

Cianciotto, N.P., White, R.C., 2017. Expanding Role of Type II Secretion in Bacterial Pathogenesis and Beyond. Infect. Immun. 85, e00014-17. https://doi.org/10.1128/IAI.00014-17

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., Ruden, D.M., 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 6, 80–92. https://doi.org/10.4161/fly.19695

Ciufo, S., Kannan, S., Sharma, S., Badretdin, A., Clark, K., Turner, S., Brover, S., Schoch, C.L., Kimchi, A., DiCuccio, M., 2018. Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. Int. J. Syst. Evol. Microbiol. 68, 2386–2392. https://doi.org/10.1099/ijsem.0.002809

Clarke, C.R., Studholme, D.J., Hayes, B., Runde, B., Weisberg, A., Cai, R., Wroblewski, T., Daunay, M.-C., Wicker, E., Castillo, J.A., Vinatzer, B.A., 2015. Genome-Enabled Phylogeographic Investigation of the Quarantine Pathogen Ralstonia solanacearum Race 3 Biovar 2 and Screening for Sources of Resistance Against Its Core Effectors. https://doi.org/10.1094/PHYTO-12-14-0373-R

Cohan, F.M., 2002. What are Bacterial Species? Annu. Rev. Microbiol. 56, 457–487. https://doi.org/10.1146/annurev.micro.56.012302.160634

Cohen, D., Melamed, S., Millman, A., Shulman, G., Oppenheimer-Shaanan, Y., Kacen, A., Doron, S., Amitai, G., Sorek, R., 2019. Cyclic GMP-AMP signalling protects bacteria against viral infection. Nature 574, 691–695. https://doi.org/10.1038/s41586-019-1605-5

Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., Brown, C.T., Porras-Alfaro, A., Kuske, C.R., Tiedje, J.M., 2014. Ribosomal Database Project: data and tools for high

throughput rRNA analysis. Nucleic Acids Res. 42, D633-642. https://doi.org/10.1093/nar/gkt1244

Coll, N.S., Valls, M., 2013. Current knowledge on the Ralstonia solanacearum type III secretion system. Microb. Biotechnol. 6, 614–620. https://doi.org/10.1111/1751-7915.12056

Collins, C., Didelot, X., 2018. A phylogenetic method to perform genome-wide association studies in microbes that accounts for population structure and recombination. PLoS Comput. Biol. 14, e1005958. https://doi.org/10.1371/journal.pcbi.1005958

Collins, R.E., Higgs, P.G., 2012. Testing the Infinitely Many Genes Model for the Evolution of the Bacterial Core Genome and Pangenome. Mol Biol Evol 29, 3413–3425. https://doi.org/10.1093/molbev/mss163

Colombi, E., Straub, C., Künzel, Sven, Templeton, Mathew D., McCann, Honour C., Rainey, Paul B., 2017. Evolution of copper resistance in the kiwifruit pathogen Pseudomonas syringae pv. actinidiae through acquisition of integrative conjugative elements and plasmids. Environ. Microbiol. 19, 819–832. https://doi.org/10.1111/1462-2920.13662

Consuegra, J., Gaffé, J., Lenski, R.E., Hindré, T., Barrick, J.E., Tenaillon, O., Schneider, D., 2021. Insertion-sequence-mediated mutations both promote and constrain evolvability during a long-term experiment with bacteria. Nat. Commun. 12. https://doi.org/10.1038/s41467-021-21210-7

Cornelis, G.R., Van Gijsegem, F., 2000. Assembly and Function of Type III Secretory Systems. Annu. Rev. Microbiol. 54, 735–774. https://doi.org/10.1146/annurev.micro.54.1.735

Corral, J., Sebastià, P., Coll, N.S., Barbé, J., Aranda, J., Valls, M., 2020. Twitching and Swimming Motility Play a Role in Ralstonia solanacearum Pathogenicity. mSphere 5. https://doi.org/10.1128/msphere.00740-19

Croll, D., McDonald, B.A., 2017. The genetic basis of local adaptation for pathogenic fungi in agricultural ecosystems. Mol. Ecol. 26, 2027–2040. https://doi.org/10.1111/mec.13870

Croucher, N.J., Page, A.J., Connor, T.R., Delaney, A.J., Keane, J.A., Bentley, S.D., Parkhill, J., Harris, S.R., 2015. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Res. 43, e15–e15. https://doi.org/10.1093/nar/gku1196

Dallas, T., Huang, S., Nunn, C., Park, A.W., Drake, J.M., 2017. Estimating parasite host range. Proc. R. Soc. B Biol. Sci. 284, 20171250. https://doi.org/10.1098/rspb.2017.1250

Deberdt, P., Guyot, J., Coranson-Beaudu, R., Launay, J., Noreskal, M., Rivière, P., Vigné, F., Laplace, D., Lebreton, L., Wicker, E., 2014. Diversity of Ralstonia solanacearum in French Guiana expands knowledge of the “emerging ecotype.” Phytopathology 104, 586–596. https://doi.org/10.1094/PHYTO-09-13-0264-R

Denny, T., 2006. Plant pathogenic Ralstonia species, in: Plant-Associated Bacteria. Springer Netherlands, pp. 573–644. https://doi.org/10.1007/978-1-4020-4538-7_16

Didelot, X., Croucher, N.J., Bentley, S.D., Harris, S.R., Wilson, D.J., 2018. Bayesian inference of ancestral dates on bacterial phylogenetic trees. Nucleic Acids Res. 46, e134. https://doi.org/10.1093/nar/gky783

Didelot, X., Siveroni, I., Volz, E.M., 2021. Additive Uncorrelated Relaxed Clock Models for the Dating of Genomic Epidemiology Phylogenies. Mol. Biol. Evol. 38, 307–317. https://doi.org/10.1093/molbev/msaa193

Didelot, X., Wilson, D.J., 2015. ClonalFrameML: Efficient Inference of Recombination in Whole Bacterial Genomes. PLoS Comput. Biol. 11, e1004041. https://doi.org/10.1371/journal.pcbi.1004041

Dietzgen, R.G., Bejerman, N.E., Goodin, M.M., Higgins, C.M., Huot, O.B., Kondo, H., Martin, K.M., Whitfield, A.E., 2020. Diversity and epidemiology of plant rhabdoviruses. Virus Res. 281, 197942. https://doi.org/10.1016/j.virusres.2020.197942

Dillon, M.M., Almeida, R.N.D., Laflamme, B., Martel, A., Weir, B.S., Desveaux, D., Guttman, D.S., 2019a. Molecular Evolution of Pseudomonas syringae Type III Secreted Effector Proteins. Front. Plant Sci. 10, 418. https://doi.org/10.3389/fpls.2019.00418

Dillon, M.M., Thakur, S., Almeida, R.N.D., Wang, P.W., Weir, B.S., Guttman, D.S., 2019b. Recombination of ecologically and evolutionarily significant loci maintains genetic cohesion in the Pseudomonas syringae species complex. Genome Biol. 20, 3. https://doi.org/10.1186/s13059-018-1606-y

Dixon, P., 2003. VEGAN, a package of R functions for community ecology. J. Veg. Sci. 14, 927–930. https://doi.org/10.1111/j.1654-1103.2003.tb02228.x

Dubois, M., Gilles, K., Hamilton, J.K., Rebers, P.A., Smith, F., 1951. A colorimetric method for the determination of sugars. Nature 168, 167. https://doi.org/10.1038/168167a0

Duchêne, S., Holt, K.E., Weill, F.X., Hello, S.L., Hawkey, J., Edwards, D.J., Fourment, M., Holmes, E.C., 2016. Genome-scale rates of evolutionary change in bacteria. Microb. Genomics 2, e000094. https://doi.org/10.1099/mgen.0.000094

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinforma. Oxf. Engl. 26, 2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Elphinstone, J.G., 2001. Monitoring and control of the potato brown rot bacterium (Ralstonia solanacearum) in the UK: a case study. https://doi.org/10.3929/ETHZ-A-004282945

Elphinstone, J.G. (Fera), Matthews-Berry, S. (Defra), 2017. Plant Pest Factsheet. Monitoring and control of the potato brown rot bacterium in irrigation water (Defra). 1–6.

Elphinstone, J.G., Hennessy, J., Wilson, J.K., Stead, D.E., 1996. Sensitivity of different methods for the detection of Ralstonia solanacearum in potato tuber extracts. EPPO Bull. 26, 663–678. https://doi.org/10.1111/j.1365-2338.1996.tb01511.x

Elphinstone, J.G., Stanford, H., Stead, D.E., 1998a. Survival and transmission of Ralstonia solanacearum in aquatic plants of Solatium dulcamara and associated surface water in England1. EPPO Bull. 28, 93–94. https://doi.org/10.1111/j.1365-2338.1998.tb00709.x

Elphinstone, J.G., Stanford, H.M., Stead, D.E., 1998b. Detection of Ralstonia solanacearum in Potato Tubers, Solanum dulcamara and Associated Irrigation Water, in: Bacterial Wilt Disease. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 133–139. https://doi.org/10.1007/978-3-662-03592-4_19

Elphinstone, M., 2005. The current bacterial wilt situation: A global overview, in: Allen, C., Prior, P., Hayward, A.C. (Eds.), Bacterial Wilt Disease and the Ralstonia Solanacearum Species Complex, APS Press. APS Press, Saint Paul, Minnesota, pp. 9–28.

EPPO, 7/21 (2), 2018. Ralstonia solanacearum, R. pseudosolanacearum and R. syzygii (Ralstonia solanacearum species complex). EPPO Bull. 48, 32–63. https://doi.org/10.1111/epp.12454

Ewels, P., Magnusson, M., Lundin, S., Käller, M., 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinforma. Oxf. Engl. 32, 3047–3048. https://doi.org/10.1093/bioinformatics/btw354

FAO, 2021. World Food and Agriculture – Statistical Yearbook 2021. FAO, Rome. https://doi.org/10.4060/cb4477en

FAO, VMD, 2022. Tackling antimicrobial use and resistance in food-producing animals. FAO; Veterinary Medicines Directorate ;

Fegan, M., Prior, P., 2005. How complex is the Ralstonia solanacearum species complex, in: Allen, C., Prior, Philippe, Hayward, A.C. (Eds.), Bacterial Wilt Disease and the Ralstonia

Solanacearum Species Complex, APS Press. APS Press, Saint Paul, Minnesota, pp. 449–461.

Fegan, Mark, Prior, Ph. (Philippe), 2005. How complex is the Ralstonia solanacearum species complex, in: Allen, C., Prior, P., Hayward, C. (Eds.), Bacterial Wilt Disease and the Ralstonia Solanacearum Species Complex, APS Press. APS Press, Saint Paul, Minnesota, pp. 449–461.

Feng, F., Zhou, J.-M., 2012. Plant–bacterial pathogen interactions mediated by type III effectors. Curr. Opin. Plant Biol., Biotic interactions 15, 469–476. https://doi.org/10.1016/j.pbi.2012.03.004

Flavier, A.B., Clough, S.J., Schell, M.A., Denny, T.P., 1997. Identification of 3-hydroxypalmitic acid methyl ester as a novel autoregulator controlling virulence in Ralstonia solanacearum. Mol. Microbiol. 26, 251–259. https://doi.org/10.1046/j.1365-2958.1997.5661945.x

Frees, D., Brøndsted, L., Ingmer, H., 2013. Bacterial proteases and virulence. Subcell. Biochem. 66, 161–192. https://doi.org/10.1007/978-94-007-5940-4_7

Freschi, L., Vincent, A.T., Jeukens, J., Emond-Rheault, J.-G., Kukavica-Ibrulj, I., Dupont, M.-J., Charette, S.J., Boyle, B., Levesque, R.C., 2019. The *Pseudomonas aeruginosa* Pan-Genome Provides New Insights on Its Population Structure, Horizontal Gene Transfer, and Pathogenicity. Genome Biol. Evol. 11, 109–120. https://doi.org/10.1093/gbe/evy259

Fritz, S.A., Purvis, A., 2010. Selectivity in Mammalian Extinction Risk and Threat Types: a New Measure of Phylogenetic Signal Strength in Binary Traits. Conserv. Biol. 24, 1042–1051. https://doi.org/10.1111/j.1523-1739.2010.01455.x

Gabriel, D.W., Allen, C., Schell, M., Denny, T.P., Greenberg, J.T., Duan, Y.P., Flores-Cruz, Z., Huang, Q., Clifford, J.M., Presting, G., González, E.T., Reddy, J., Elphinstone, J., Swanson, J., Yao, J., Mulholland, V., Liu, L., Farmerie, W., Patnaikuni, M., Balogh, B., Norman, D., Alvarez, A., Castillo, J.A., Jones, J., Saddler, G., Walunas, T., Zhukov, A., Mikhailova, N., 2006. Identification of Open Reading Frames Unique to a Select Agent: Ralstonia solanacearum Race 3 Biovar 2. Mol. Plant-Microbe Interactions® 19, 69–79. https://doi.org/10.1094/MPMI-19-0069

Gabriel, D.W., Kingsley, M.T., Yang, Y., Chen, J., Roberts, P., 1994. Host-specific virulence genes of Xanthomonas, in: Kado, C.I., Crosa, J.H. (Eds.), Molecular Mechanisms of

Bacterial Virulence, Developments in Plant Pathology. Springer Netherlands, Dordrecht, pp. 141–158. https://doi.org/10.1007/978-94-011-0746-4_10

Gao, R., Mack, T.R., Stock, A.M., 2007. Bacterial Response Regulators: Versatile Regulatory Strategies from Common Domains. Trends Biochem. Sci. 32, 225–234. https://doi.org/10.1016/j.tibs.2007.03.002

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

Geng, R., Cheng, L., Cao, C., Liu, Z., Liu, D., Xiao, Z., Wu, X., Huang, Z., Feng, Q., Luo, C., Chen, Z., Zhang, Z., Jiang, C., Ren, M., Yang, A., 2022. Comprehensive Analysis Reveals the Genetic and Pathogenic Diversity of Ralstonia solanacearum Species Complex and Benefits Its Taxonomic Classification. Front. Microbiol. 13.

Genin, S., 2010. Molecular traits controlling host range and adaptation to plants in Ralstonia solanacearum. New Phytol. 187, 920–928. https://doi.org/10.1111/j.1469-8137.2010.03397.x

Genin, S., Boucher, C., 2004. Lessons learned from the genome analysis of Ralstonia solanacearum. Annu Rev Phytopathol 42, 107–141. https://doi.org/10.1146/annurev.phyto.42.011204.104301

Genin, S., Denny, T.P., 2012a. Pathogenomics of the Ralstonia solanacearum Species Complex. https://doi.org/10.1146/annurev-phyto-081211-173000

Genin, S., Denny, T.P., 2012b. Pathogenomics of the Ralstonia solanacearum species complex. Annu. Rev. Phytopathol. 50, 67–89. https://doi.org/10.1146/annurev-phyto-081211-173000

Golicz, A.A., Bayer, P.E., Bhalla, P.L., Batley, J., Edwards, D., 2020. Pangenomics Comes of Age: From Bacteria to Plant and Animal Applications. Trends Genet. TIG 36, 132–145. https://doi.org/10.1016/j.tig.2019.11.006

Gomez-Valero, L., Rusniok, C., Carson, D., Mondino, S., Pérez-Cobas, A.E., Rolando, M., Pasricha, S., Reuter, S., Demirtas, J., Crumbach, J., Descorps-Declere, S., Hartland, E.L., Jarraud, S., Dougan, G., Schroeder, G.N., Frankel, G., Buchrieser, C., 2019. More than 18,000 effectors in the Legionella genus genome provide multiple, independent combinations for replication in human cells. Proc. Natl. Acad. Sci. 116, 2265–2273. https://doi.org/10.1073/pnas.1808016116

Gonçalves, O.S., de Queiroz, M.V., Santana, M.F., 2020a. Potential evolutionary impact of integrative and conjugative elements (ICEs) and genomic islands in the Ralstonia

solanacearum species complex. Sci. Rep. 10, 12498. https://doi.org/10.1038/s41598-020-69490-1

Gonçalves, O.S., Francisco, C.S., Catarine Silva de Assis, J., Suellen Fernandes, A., Santos Souza, T., Guilherme do Carmo Rodrigues, L., Vieira de Queiroz, M., Ferreira Santana, M., 2020b. Transposable elements contribute to the genome plasticity of Ralstonia solanacearum species complex. Microb. Genomics 6. https://doi.org/10.1099/mgen.0.000374

Gonçalves, O.S., Souza, F. de O., Bruckner, F.P., Santana, M.F., Alfenas-Zerbini, P., 2021. Widespread distribution of prophages signaling the potential for adaptability and pathogenicity evolution of Ralstonia solanacearum species complex. Genomics 113, 992–1000. https://doi.org/10.1016/j.ygeno.2021.02.011

González-Tortuero, E., David Sean Sutton, T., Velayudhan, V., Nikolaevich Shkoporov, A., Anne Draper, L., Robert Stockdale, S., Ross, P., Hill, C., 2018. Title: VIGA: a sensitive, precise and automatic de novo VIral Genome Annotator. 1. https://doi.org/10.1101/277509

Gophna, U., Ron, E.Z., Graur, D., 2003. Bacterial type III secretion systems are ancient and evolved by multiple horizontal-transfer events. Gene 312, 151–163. https://doi.org/10.1016/S0378-1119(03)00612-7

Gori, A., Harrison, O.B., Mlia, E., Nishihara, Y., Chan, J.M., Msefula, J., Mallewa, M., Dube, Q., Swarthout, T.D., Nobbs, A.H., Maiden, M.C.J., French, N., Heyderman, R.S., 2020. Pan-GWAS of Streptococcus agalactiae Highlights Lineage-Specific Genes Associated with Virulence and Niche Adaptation. mBio 11, e00728-20. https://doi.org/10.1128/mBio.00728-20

Granada, G.A., Sequeira, L., 1983. Survival of Pseudomonas solanacearum in soil, rhizosphere, and plant roots. Can. J. Microbiol. 29, 433–440. https://doi.org/10.1139/m83-070

Greenrod, S.T.E., Stoycheva, M., Elphinstone, J., Friman, V.-P., 2022a. Global diversity and distribution of prophages are lineage-specific within the Ralstonia solanacearum species complex. BMC Genomics 23, 689. https://doi.org/10.1186/s12864-022-08909-7

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

Grolemund, H.W. and G., 2017. R for Data Science.

Grover, A., Azmi, W., Gadewar, A.V., Pattanayak, D., Naik, P.S., Shekhawat, G.S., Chakrabarti, S.K., 2006. Genotypic diversity in a localized population of Ralstonia solanacearum as revealed by random amplified polymorphic DNA markers. J. Appl. Microbiol. 101, 798–806. https://doi.org/10.1111/j.1365-2672.2006.02974.x

Gu, S., Wei, Z., Shao, Z., Friman, V.P., Cao, K., Yang, T., Kramer, J., Wang, X., Li, M., Mei, X., Xu, Y., Shen, Q., Kümmerli, R., Jousset, A., 2020a. Competition for iron drives phytopathogen control by natural rhizosphere microbiomes. Nat. Microbiol. 5, 1002–1010. https://doi.org/10.1038/s41564-020-0719-8

Gu, S., Wei, Z., Shao, Z., Friman, V.P., Cao, K., Yang, T., Kramer, J., Wang, X., Li, M., Mei, X., Xu, Y., Shen, Q., Kümmerli, R., Jousset, A., 2020b. Competition for iron drives phytopathogen control by natural rhizosphere microbiomes. Nat. Microbiol. 5, 1002–1010. https://doi.org/10.1038/s41564-020-0719-8

Guarischi-Sousa, R., Puigvert, M., Coll, N.S., Siri, M.I., Pianzzola, M.J., Valls, M., Setubal, J.C., 2016. Complete genome sequence of the potato pathogen Ralstonia solanacearum UY031. Stand. Genomic Sci. 11, 7. https://doi.org/10.1186/s40793-016-0131-4

Guidot, A., Coupat, B., Fall, S., Prior, P., Bertolla, F., 2009. Horizontal gene transfer between ralstonia solanacearum strains detected by comparative genomic hybridization on microarrays. ISME J. 3, 549–562. https://doi.org/10.1038/ismej.2009.14

Hall, R.J., Whelan, F.J., Cummins, E.A., Connor, C., McNally, A., McInerney, J.O.Y. 2021, 2021. Gene-gene relationships in an Escherichia coli accessory genome are linked to function and mobility. Microb. Genomics 7, 000650. https://doi.org/10.1099/mgen.0.000650

Hanson, C.A., Fuhrman, J.A., Claire Horner-Devine, M., H Martiny, J.B., 2012. Beyond biogeographic patterns: processes shaping the microbial landscape. https://doi.org/10.1038/nrmicro2795

Hawkey, J., Hamidian, M., Wick, R.R., Edwards, D.J., Billman-Jacobe, H., Hall, R.M., Holt, K.E., 2015. ISMapper: identifying transposase insertion sites in bacterial genomes from short read sequence data. BMC Genomics 16, 667. https://doi.org/10.1186/s12864-015-1860-2

Hawkey, J., Monk, J.M., Billman-Jacobe, H., Palsson, B., Holt, K.E., 2020. Impact of insertion sequences on convergent evolution of Shigella species. PLOS Genet. 16, e1008931. https://doi.org/10.1371/journal.pgen.1008931

Hayward, A.C., 1991a. BIOLOGY AND EPIDEMIOLOGY OF BACTERIAL WILT CAUSED BY PSEUDOMONAS SOLANACEARUM. Annu Rev PhYlOpathol 29, 65–87. https://doi.org/10.1146/annurev.py.29.090191.000433

Hayward, A.C., 1991b. Biology and epidemiology of bacterial wilt caused by pseudomonas solanacearum. Annu. Rev. Phytopathol. 29, 65–87. https://doi.org/10.1146/annurev.py.29.090191.000433

Hayward, A.C., 1964. Characteristics of Pseudomonas solanacearum. J. Appl. Bacteriol. 27, 265–277. https://doi.org/10.1111/j.1365-2672.1964.tb04912.x

Hayward, A.C., 1960. A Method for Characterizing Pseudomonas solanacearum. Nature 186, 405–406. https://doi.org/10.1038/186405a0

He, S.Y., Jin, Q., 2003. The Hrp pilus: learning from flagella. Curr. Opin. Microbiol. 6, 15–19. https://doi.org/10.1016/s1369-5274(02)00007-3

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. Mol. Biol. Evol. 35, 518–522. https://doi.org/10.1093/molbev/msx281

Hotchkiss, R.D., Dubos, R.J., 1940. FRACTIONATION OF THE BACTERICIDAL AGENT FROM CULTURES OF A SOIL BACILLUS. J. Biol. Chem. 132, 791–792. https://doi.org/10.1016/S0021-9258(19)56231-7

Hu, J., Wei, Z., Weidner, S., Friman, V.-P., Xu, Y.-C., Shen, Q.-R., Jousset, A., 2017. Probiotic Pseudomonas communities enhance plant growth and nutrient assimilation via diversity-mediated ecosystem functioning. Soil Biol. Biochem. 113, 122–129. https://doi.org/10.1016/j.soilbio.2017.05.029

Huerta, A.I., Milling, A., Allen, C., 2015. Tropical strains of Ralstonia solanacearum outcompete race 3 biovar 2 strains at lowland tropical temperatures. Appl. Environ. Microbiol. 81, 3542–3551. https://doi.org/10.1128/AEM.04123-14

Hulin, M.T., Armitage, A.D., Vicente, J.G., Holub, E.B., Baxter, L., Bates, H.J., Mansfield, J.W., Jackson, R.W., Harrison, R.J., 2018. Comparative genomics of Pseudomonas syringae reveals convergent gene gain and loss associated with specialization onto cherry (Prunus avium). New Phytol. 219, 672–696. https://doi.org/10.1111/nph.15182

Hwang, M.S.H., Morgan, R.L., Sarkar, S.F., Wang, P.W., Guttman, D.S., 2005. Phylogenetic Characterization of Virulence and Resistance Phenotypes of Pseudomonas syringae.

Appl. Environ. Microbiol. 71, 5182–5191. https://doi.org/10.1128/AEM.71.9.5182-5191.2005

Iranzo, J., Wolf, Y.I., Koonin, E.V., Sela, I., 2019. Gene gain and loss push prokaryotes beyond the homologous recombination barrier and accelerate genome sequence divergence. Nat. Commun. 10. https://doi.org/10.1038/s41467-019-13429-2

Irons, J.L., Hodge-Hanson, K., Downs, D.M., 2020. RidA Proteins Protect against Metabolic Damage by Reactive Intermediates. Microbiol. Mol. Biol. Rev. MMBR 84, e00024-20. https://doi.org/10.1128/MMBR.00024-20

Jaillard, M., Lima, L., Tournoud, M., Mahé, P., Belkum, A. van, Lacroix, V., Jacob, L., 2018. A fast and agnostic method for bacterial genome-wide association studies: Bridging the gap between k-mers and genetic events. PLOS Genet. 14, e1007758. https://doi.org/10.1371/journal.pgen.1007758

Janse, J.D., van den Beld, H.E., Elphinstone, J., Simpkins, S., Tjou-Tam-Sin, N.N.A., van Vaerenbergh, J., 2004. Introduction to Europe of Ralstonia Solanacearum Biovar 2, Race 3 in Pelargonium Zonale Cuttings. J. Plant Pathol. 86, 147–155.

Jensen, L.J., Julien, P., Kuhn, M., von Mering, C., Muller, J., Doerks, T., Bork, P., 2008. eggNOG: Automated construction and annotation of orthologous groups of genes. Nucleic Acids Res. 36, 250–254. https://doi.org/10.1093/nar/gkm796

Jeong, E.L., Timmis, J.N., 2000. Novel insertion sequence elements associated with genetic heterogeneity and phenotype conversion in Ralstonia solanacearum. J. Bacteriol. 182, 4673–4676. https://doi.org/10.1128/JB.182.16.4673-4676.2000

Jiang, G., Wang, N., Zhang, Yaoyu, Zhang, Yuling, Yu, J., Zhang, Yong, Wei, Z., Xu, Y., Geisen, S., Friman, V.-P., Shen, Q., 2021a. The relative importance of soil moisture in predicting bacterial wilt disease occurrence. Soil Ecol. Lett. 3, 356–366. https://doi.org/10.1007/s42832-021-0086-2

Jiang, G., Wang, N., Zhang, Yaoyu, Zhang, Yuling, Yu, J., Zhang, Yong, Wei, Z., Xu, Y., Geisen, S., Friman, V.-P., Shen, Q., 2021b. The relative importance of soil moisture in predicting bacterial wilt disease occurrence. Soil Ecol. Lett. 3, 356–366. https://doi.org/10.1007/s42832-021-0086-2

Jiang, G., Wei, Z., Xu, J., Chen, H., Zhang, Y., She, X., Macho, A.P., Ding, W., Liao, B., 2017. Bacterial wilt in China: history, current status, and future perspectives. Front. Plant Sci. 8, 1549. https://doi.org/10.3389/fpls.2017.01549

Jones, J.D.G., Dangl, J.L., 2006. The plant immune system. Nature 444, 323–329. https://doi.org/10.1038/nature05286

Joung, Y.S., Ge, Z., Buie, C.R., 2017. Bioaerosol generation by raindrops on soil. Nat. Commun. 8, 14668. https://doi.org/10.1038/ncomms14668

Kadam, R.V., Jagtap, G.P., 2018. Effect of different soil types on the growth of R. solanacearum. Int. J. Curr. Microbiol. Appl. Sci. 7, 583–586. https://doi.org/10.20546/ijcmas.2018.711.070

Kahm, M., Hasenbrink, G., Lichtenberg-Fraté, H., Ludwig, J., Kschischo, M., 2010. grofit: Fitting Biological Growth Curves with R. J. Stat. Softw. 33, 1–21. https://doi.org/10.18637/jss.v033.i07

Kai, K., Ohnishi, H., Shimatani, M., Ishikawa, S., Mori, Y., Kiba, A., Ohnishi, K., Tabuchi, M., Hikichi, Y., 2015. Methyl 3-Hydroxymyristate, a Diffusible Signal Mediating phc Quorum Sensing in Ralstonia solanacearum. ChemBioChem 16, 2309–2318. https://doi.org/10.1002/cbic.201500456

Kannan, V., Bastas, K., Antony, R., 2015. Plant Pathogenic Bacteria: An Overview, in: Kannan, V, Bastas, K. (Eds.), Sustainable Approaches to Controlling Plant Pathogenic Bacteria. CRC Press, pp. 1–16. https://doi.org/10.1201/b18892-2

Katz, L., Griswold, T., Morrison, S., Caravas, J., Zhang, S., Bakker, H., Deng, X., Carleton, H., 2019. Mashtree: a rapid comparison of whole genome sequence files. J. Open Source Softw. 4, 1762. https://doi.org/10.21105/joss.01762

Kent, A.G., Dupont, C.L., Yooseph, S., Martiny, A.C., 2016. Global biogeography of Prochlorococcus genome diversity in the surface ocean. ISME J. 10, 1856–1865. https://doi.org/10.1038/ismej.2015.265

Kessler, N., Schuhmann, H., Morneweg, S., Linne, U., Marahiel, M.A., 2004. The Linear Pentadecapeptide Gramicidin Is Assembled by Four Multimodular Nonribosomal Peptide Synthetases That Comprise 16 Modules with 56 Catalytic Domains*. J. Biol. Chem. 279, 7413–7419. https://doi.org/10.1074/jbc.M309658200

Kim, S.H., Olson, T.N., Schaad, N.W., Moorman, G.W., 2003. Ralstonia solanacearum Race 3, Biovar 2, the Causal Agent of Brown Rot of Potato, Identified in Geraniums in Pennsylvania, Delaware, and Connecticut. Plant Dis. 87, 450–450. https://doi.org/10.1094/PDIS.2003.87.4.450C

Kimura, M., 1968. Evolutionary rate at the molecular level. Nature 217, 624–626. https://doi.org/10.1038/217624a0

Konstantinidis, K.T., Ramette, A., Tiedje, J.M., 2006. The bacterial species definition in the genomic era. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 361, 1929–1940. https://doi.org/10.1098/rstb.2006.1920

Koutsoudis, M.D., Tsaltas, D., Minogue, T.D., von Bodman, S.B., 2006. Quorum-sensing regulation governs bacterial adhesion, biofilm development, and host colonization in Pantoea stewartii subspecies stewartii. Proc. Natl. Acad. Sci. U. S. A. 103, 5983–5988. https://doi.org/10.1073/pnas.0509860103

Kuo, C.-H., Moran, N.A., Ochman, H., 2009. The consequences of genetic drift for bacterial genome complexity. Genome Res. 19, 1450–1454. https://doi.org/10.1101/gr.091785.109

Kurm, V., Houwers, I., Coipan, C.E., Bonants, P., Waalwijk, C., van der Lee, T., Brankovics, B., van der Wolf, J., 2021. Whole genome characterization of strains belonging to the Ralstonia solanacearum species complex and in silico analysis of TaqMan assays for detection in this heterogenous species complex. Eur. J. Plant Pathol. 159, 593–613. https://doi.org/10.1007/s10658-020-02190-8

Kutsukake, K., Nakashima, H., Tominaga, A., Abo, T., 2006. Two DNA Invertases Contribute to Flagellar Phase Variation in Salmonella enterica Serovar Typhimurium Strain LT2. J. Bacteriol. 188, 950–957. https://doi.org/10.1128/JB.188.3.950-957.2006

Langille, M.G.I., Hsiao, W.W.L., Brinkman, F.S.L., 2010. Detecting genomic islands using bioinformatics approaches. Nat. Rev. Microbiol. 8, 373–382. https://doi.org/10.1038/nrmicro2350

Langner, T., Harant, A., Gomez-Luciano, L.B., Shrestha, R.K., Malmgren, A., Latorre, S.M., Burbano, H.A., Win, J., Kamoun, S., 2021. Genomic rearrangements generate hypervariable mini-chromosomes in host-specific isolates of the blast fungus. PLoS Genet. 17, e1009386. https://doi.org/10.1371/journal.pgen.1009386

Lapierre, M., Blin, C., Lambert, A., Achaz, G., Rocha, E.P.C., 2016. The Impact of Selection, Gene Conversion, and Biased Sampling on the Assessment of Microbial Demography. Mol. Biol. Evol. 33, 1711–1725. https://doi.org/10.1093/molbev/msw048

Lees, J.A., Galardini, M., Bentley, S.D., Weiser, J.N., Corander, J., 2018. pyseer: A comprehensive tool for microbial pangenome-wide association studies. Bioinformatics 34, 4310–4312. https://doi.org/10.1093/bioinformatics/bty539

Lees, J.A., Tien Mai, T., Galardini, M., Wheeler, N.E., Corander, J., 2019. Improved inference and prediction of bacterial genotype-phenotype associations using pangenome-spanning regressions. bioRxiv. https://doi.org/10.1101/852426

Lelliott, R.A., Stead, D.E., 1987. Methods for the diagnosis of bacterial diseases of plants. Methods Diagn. Bact. Dis. Plants.

Levin, B.R., Cornejo, O.E., 2009. The Population and Evolutionary Dynamics of Homologous Gene Recombination in Bacteria. PLOS Genet. 5, e1000601. https://doi.org/10.1371/journal.pgen.1000601

Li, M., Pommier, T., Yin, Y., Wang, J., Gu, S., Jousset, A., Keuskamp, J., Wang, H., Wei, Z., Xu, Y., Shen, Q., Kowalchuk, G.A., 2022. Indirect reduction of Ralstonia solanacearum via pathogen helper inhibition. ISME J. 16, 868–875. https://doi.org/10.1038/s41396-021-01126-2

Li, M., Wei, Z., Wang, J., Jousset, A., Friman, V.-P., Xu, Y., Shen, Q., Pommier, T., 2019. Facilitation promotes invasions in plant-associated microbial communities. Ecol. Lett. 22, 149–158. https://doi.org/10.1111/ele.13177

Li, P., Liu, M., Li, G., Liu, K., Liu, T., Wu, M., Saleem, M., Li, Z., 2021. Phosphorus availability increases pathobiome abundance and invasion of rhizosphere microbial networks by Ralstonia. Environ. Microbiol. 23, 5992–6003. https://doi.org/10.1111/1462-2920.15696

Li, S., Liu, Y., Wang, J., Yang, L., Zhang, S., Xu, C., Ding, W., 2017. Soil Acidification Aggravates the Occurrence of Bacterial Wilt in South China. Front. Microbiol. 8.

Lindeberg, M., Cunnac, S., Collmer, A., 2009. The evolution of Pseudomonas syringae host specificity and type III effector repertoires. Mol. Plant Pathol. 10, 767–775. https://doi.org/10.1111/j.1364-3703.2009.00587.x

Lu, H., Lema, S.A., Planas-Marquès, M., Alonso-Díaz, A., Valls, M., Coll, N.S., 2018. Type III secretion-dependent and-independent phenotypes caused by Ralstonia solanacearum in Arabidopsis roots. Mol. Plant. Microbe Interact. 31, 175–184. https://doi.org/10.1094/MPMI-05-17-0109-FI

Lurwanu, Y., Wang, Y.-P., Wu, E.-J., He, D.-C., Waheed, A., Nkurikiyimfura, O., Wang, Z., Shang, L.-P., Yang, L.-N., Zhan, J., 2021. Increasing temperature elevates the variation and spatial differentiation of pesticide tolerance in a plant pathogen. Evol. Appl. 14, 1274–1285. https://doi.org/10.1111/eva.13197

Luvisi, A., Aprile, A., Sabella, E., Vergine, M., Nicolì, F., Nutricati, E., Miceli, A., Negro, C., De BELLIS, L., 2017. Xylella fastidiosa subsp. pauca (CoDiRO strain) infection in four olive (Olea europaea L.) cultivars: profile of phenolic compounds in leaves and progression of leaf scorch symptoms. Phytopathol. Mediterr. 56, 259–273.

Lynch, M., Conery, J.S., 2003. The Origins of Genome Complexity. Science 302, 1401–1404. https://doi.org/10.1126/science.1089370

Machado, D.J., White, R.A., Kofsky, J., Janies, D.A., 2021. Fundamentals of genomic epidemiology, lessons learned from the coronavirus disease 2019 (COVID-19) pandemic, and new directions. Antimicrob. Steward. Healthc. Epidemiol. 1, e60. https://doi.org/10.1017/ash.2021.222

Macho, A.P., Zipfel, C., 2015. Targeting of plant pattern recognition receptor-triggered immunity by bacterial type-III secretion system effectors. Curr. Opin. Microbiol. 23, 14–22. https://doi.org/10.1016/j.mib.2014.10.009

Maculins, T., Fiskin, E., Bhogaraju, S., Dikic, I., 2016. Bacteria-host relationship: ubiquitin ligases as weapons of invasion. Cell Res. 26, 499–510. https://doi.org/10.1038/cr.2016.30

Mak, T.N., Schmid, M., Brzuszkiewicz, E., Zeng, G., Meyer, R., Sfanos, K.S., Brinkmann, V., Meyer, T.F., Brüggemann, H., 2013. Comparative genomics reveals distinct host-interacting traits of three major human-associated propionibacteria. BMC Genomics 14, 640. https://doi.org/10.1186/1471-2164-14-640

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M., Verdier, V., Beer, S.V., Machado, M.A., Toth, I., Salmond, G., Foster, G.D., 2012. Top 10 plant pathogenic bacteria in molecular plant pathology. Mol. Plant Pathol. 13, 614–629. https://doi.org/10.1111/j.1364-3703.2012.00804.x

McInerney, J.O., McNally, A., O'Connell, M.J., 2017. Why prokaryotes have pangenomes. Nat. Microbiol. 2, 1–5. https://doi.org/10.1038/nmicrobiol.2017.40

McManus, P.S., Stockwell, V.O., Sundin, G.W., Jones, A.L., 2002. Antibiotic use in plant agriculture. Annu. Rev. Phytopathol. 40, 443–465. https://doi.org/10.1146/annurev.phyto.40.120301.093927

Meena, M., Swapnil, P., Zehra, A., Aamir, M., Dubey, M.K., Patel, C.B., Upadhyay, R.S., 2019. Virulence Factors and Their Associated Genes in Microbes, in: New and Future Developments in Microbial Biotechnology and Bioengineering. Elsevier, pp. 181–208. https://doi.org/10.1016/B978-0-444-63503-7.00011-5

Messiha, N.A.S., van Bruggen, A.H.C., Franz, E., Janse, J.D., Schoeman-Weerdesteijn, M.E., Termorshuizen, A.J., van Diepeningen, A.D., 2009. Effects of soil type, management type and soil amendments on the survival of the potato brown rot bacterium Ralstonia solanacearum. Appl. Soil Ecol. 43, 206–215. https://doi.org/10.1016/j.apsoil.2009.07.008

Milling, A., Babujee, L., Allen, C., 2011. Ralstonia solanacearum Extracellular Polysaccharide Is a Specific Elicitor of Defense Responses in Wilt-Resistant Tomato Plants. PLOS ONE 6, e15853. https://doi.org/10.1371/journal.pone.0015853

Milling, A., Meng, F., Denny, T.P., Allen, C., 2009. Interactions with hosts at cool temperatures, not cold tolerance, explain the unique epidemiology of Ralstonia solanacearum race 3 biovar 2. Phytopathology 99, 1127–1134. https://doi.org/10.1094/PHYTO-99-10-1127

Mira, A., Martín-Cuadrado, A.B., D'Auria, G., Rodríguez-Valera, F., 2010. The bacterial pan-genome:a new paradigm in microbiology. Int. Microbiol. Off. J. Span. Soc. Microbiol. 13, 45–57. https://doi.org/10.2436/20.1501.01.110

Monteil, C.L., Yahara, K., Studholme, D.J., Mageiros, L., Méric, G., Swingle, B., Morris, C.E., Vinatzer, B.A., Sheppard, S.K.Y. 2016, 2016. Population-genomic insights into emergence, crop adaptation and dissemination of Pseudomonas syringae pathogens. Microb. Genomics 2, e000089. https://doi.org/10.1099/mgen.0.000089

Montie, T.C., Doyle-Huntzinger, D., Craven, R.C., Holder, I.A., 1982. Loss of virulence associated with absence of flagellum in an isogenic mutant of Pseudomonas aeruginosa in the burned-mouse model. Infect. Immun. 38, 1296–1298. https://doi.org/10.1128/iai.38.3.1296-1298.1982

Morelli, M., García-Madero, J.M., Jos, Á., Saldarelli, P., Dongiovanni, C., Kovacova, M., Saponari, M., Baños Arjona, A., Hackl, E., Webb, S., Compant, S., 2021. Xylella

fastidiosa in Olive: A Review of Control Attempts and Current Management. Microorganisms 9, 1771. https://doi.org/10.3390/microorganisms9081771

Morris, C.E., Moury, B., 2019. Revisiting the concept of host range of plant pathogens. Annu. Rev. Phytopathol. 57, 63–90. https://doi.org/10.1146/annurev-phyto-082718-100034

Neale, H.C., Jackson, R.W., Preston, G.M., Arnold, D.L., 2018. Supercoiling of an excised genomic island represses effector gene expression to prevent activation of host resistance. Mol. Microbiol. 110, 444–454. https://doi.org/10.1111/mmi.14111

Nelson, R., Wiesner-Hanks, T., Wisser, R., Balint-Kurti, P., 2018. Navigating complexity to breed disease-resistant crops. Nat. Rev. Genet. 19, 21–33. https://doi.org/10.1038/nrg.2017.82

Neogi, U., Bontell, I., Shet, A., Costa, A.D., Gupta, S., Diwan, V., Laishram, R.S., Wanchu, A., Ranga, U., Banerjea, A.C., Sönnerborg, A., 2012. Molecular Epidemiology of HIV-1 Subtypes in India: Origin and Evolutionary History of the Predominant Subtype C. PLOS ONE 7, e39819. https://doi.org/10.1371/journal.pone.0039819

Ngou, B.P.M., Ding, P., Jones, J.D.G., 2022. Thirty years of resistance: Zig-zag through the plant immune system. Plant Cell 34, 1447–1478. https://doi.org/10.1093/plcell/koac041

Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., Minh, B.Q., 2015a. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274. https://doi.org/10.1093/molbev/msu300

Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., Minh, B.Q., 2015b. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274. https://doi.org/10.1093/molbev/msu300

Nowell, R.W., Laue, B.E., Sharp, P.M., Green, S., 2016. Comparative genomics reveals genes significantly associated with woody hosts in the plant pathogen Pseudomonas syringae. Mol. Plant Pathol. 17, 1409–1424. https://doi.org/10.1111/mpp.12423

Nübel, U., Dordel, J., Kurt, K., Strommenger, B., Westh, H., Shukla, S.K., Žemličková, H., Leblois, R., Wirth, T., Jombart, T., Balloux, F., Witte, W., 2010. A Timescale for Evolution, Population Expansion, and Spatial Spread of an Emerging Clone of Methicillin-Resistant Staphylococcus aureus. PLoS Pathog. 6, e1000855. https://doi.org/10.1371/journal.ppat.1000855

Page, A.J., Cummins, C.A., Hunt, M., Wong, V.K., Reuter, S., Holden, M.T.G., Fookes, M., Falush, D., Keane, J.A., Parkhill, J., 2015. Roary: Rapid large-scale prokaryote pan

genome analysis. Bioinformatics 31, 3691–3693. https://doi.org/10.1093/bioinformatics/btv421

Pandey, A., Moon, H., Choi, S., Yoon, H., Prokchorchik, M., Jayaraman, J., Sujeevan, R., Kang, Y.M., McCann, H.C., Segonzac, C., Kim, C.M., Park, S.J., Sohn, K.H., 2021. Ralstonia solanacearum type III effector RipJ triggers bacterial wilt resistance in Solanum pimpinellifolium. Mol. Plant. Microbe Interact. 34. https://doi.org/10.1094/MPMI-09-20-0256-R

Parales, R.E., Harwood, C.S., 2002. Bacterial chemotaxis to pollutants and plant-derived aromatic molecules. Curr. Opin. Microbiol. 5, 266–273. https://doi.org/10.1016/S1369-5274(02)00320-X

Parkinson, N., Bryant, R., Bew, J., Conyers, C., Stones, R., Alcock, M., Elphinstone, J., 2013. Application of variable-number tandem-repeat typing to discriminate Ralstonia solanacearum strains associated with english watercourses and disease outbreaks. Appl. Environ. Microbiol. 79, 6016–6022. https://doi.org/10.1128/AEM.01219-13

Peeters, N., Carrère, S., Anisimova, M., Plener, L., Cazalé, A.-C., Genin, S., 2013a. Repertoire, unified nomenclature and evolution of the Type III effector gene set in the Ralstonia solanacearum species complex. BMC Genomics 14, 859. https://doi.org/10.1186/1471-2164-14-859

Peeters, N., Guidot, A., Vailleau, F., Valls, M., 2013b. *Ralstonia solanacearum*, a widespread bacterial plant pathogen in the post-genomic era. Mol. Plant Pathol. 14, 651–662. https://doi.org/10.1111/mpp.12038

Péret, B., Larrieu, A., Bennett, M.J., 2009. Lateral root emergence: a difficult birth. J. Exp. Bot. 60, 3637–3643. https://doi.org/10.1093/jxb/erp232

Perrier, A., Barlet, X., Peyraud, R., Rengel, D., Guidot, A., Genin, S., 2018. Comparative transcriptomic studies identify specific expression patterns of virulence factors under the control of the master regulator PhcA in the Ralstonia solanacearum species complex. Microb. Pathog. 116, 273–278. https://doi.org/10.1016/j.micpath.2018.01.028

Peynard, R., Cottret, L., Marmiesse, L., Gouzy, J., Genin, S., 2016. A Resource Allocation Trade-Off between Virulence and Proliferation Drives Metabolic Versatility in the Plant Pathogen Ralstonia solanacearum - file. PLOS Pathog. 12.

Peyraud, R., Cottret, L., Marmiesse, L., Gouzy, J., Genin, S., 2016. A Resource Allocation Trade-Off between Virulence and Proliferation Drives Metabolic Versatility in the Plant Pathogen Ralstonia solanacearum. PLoS Pathog. 12, 1–25. https://doi.org/10.1371/journal.ppat.1005939

Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S.E., Lercher, M.J., 2014. PopGenome: An efficient swiss army knife for population genomic analyses in R. Mol. Biol. Evol. 31, 1929–1936. https://doi.org/10.1093/molbev/msu136

Pfennig, K.S., 2001. Evolution of pathogen virulence: the role of variation in host phenotype. Proc. Biol. Sci. 268, 755–760. https://doi.org/10.1098/rspb.2000.1582

Polz, M.F., Cordero, O.X., 2016. Bacterial evolution: Genomics of metabolic trade-offs. Nat. Microbiol. 1. https://doi.org/10.1038/nmicrobiol.2016.181

Poussier, S., Thoquet, P., Trigalet-Demery, D., Barthet, S., Meyer, D., Arlat, M., Trigalet, A., 2003. Host plant-dependent phenotypic reversion of Ralstonia solanacearum from non-pathogenic to pathogenic forms via alterations in the phcA gene. Mol. Microbiol. 49, 991–1003. https://doi.org/10.1046/j.1365-2958.2003.03605.x

Prade, R.A., Zhan, D., Ayoubi, P., Mort, A.J., 1999. Pectins, pectinases and plant-microbe interactions. Biotechnol. Genet. Eng. Rev. 16, 361–391. https://doi.org/10.1080/02648725.1999.10647984

Price, M.N., Arkin, A.P., 2015. Weakly Deleterious Mutations and Low Rates of Recombination Limit the Impact of Natural Selection on Bacterial Genomes. mBio 6, e01302-01315. https://doi.org/10.1128/mBio.01302-15

Prior, Ph. (Philippe), Fegan, M., 2005. Bacterial wilt disease and the Ralstonia solanacearum species complex. American Phytopathological Society.

R Core Team, 2017. R: A language and environment for statistical computing.

R Studio Team, 2020. RStudio: Integrated Development for R.

Ramírez, M., Moncada, R.N., Villegas-Escobar, V., Jackson, R.W., Ramírez, C.A., 2020. Phylogenetic and pathogenic variability of strains of Ralstonia solanacearum causing moko disease in Colombia. Plant Pathol. 69, 360–369. https://doi.org/10.1111/ppa.13121

Rancati, G., Moffat, J., Typas, A., Pavelka, N., 2018. Emerging and evolving concepts in gene essentiality. Nat. Rev. Genet. 19, 34–49. https://doi.org/10.1038/nrg.2017.74

Rasko, D.A., Rosovitz, M.J., Myers, G.S.A., Mongodin, E.F., Fricke, W.F., Gajer, P., Crabtree, J., Sebaihia, M., Thomson, N.R., Chaudhuri, R., Henderson, I.R., Sperandio, V., Ravel, J., 2008. The pangenome structure of Escherichia coli: comparative genomic analysis of E. coli commensal and pathogenic isolates. J. Bacteriol. 190, 6881–6893. https://doi.org/10.1128/JB.00619-08

Raza, W., Ling, N., Liu, D., Wei, Z., Huang, Q., Shen, Q., 2016. Volatile organic compounds produced by Pseudomonas fluorescens WR-1 restrict the growth and virulence traits of Ralstonia solanacearum. Microbiol. Res. 192, 103–113. https://doi.org/10.1016/j.micres.2016.05.014

Remenant, B., Coupat-Goutaland, B., Guidot, A., Cellier, G., Wicker, E., Allen, C., Fegan, M., Pruvost, O., Elbaz, M., Calteau, A., Salvignol, G., Mornico, D., Mangenot, S., Barbe, V., Médigue, C., Prior, P., 2010. Genomes of three tomato pathogens within the Ralstonia solanacearum species complex reveal significant evolutionary divergence. BMC Genomics 11. https://doi.org/10.1186/1471-2164-11-379

Remenant, B., de Cambiaire, J.-C., Cellier, G., Jacobs, J.M., Mangenot, S., Barbe, V., Lajus, A., Vallenet, D., Medigue, C., Fegan, M., Allen, C., Prior, P., 2011. Ralstonia syzygii, the Blood Disease Bacterium and Some Asian R. solanacearum Strains Form a Single Genomic Species Despite Divergent Lifestyles. PLoS ONE 6, e24356. https://doi.org/10.1371/journal.pone.0024356

Robbins, J.C., Oxender, D.L., 1973. Transport Systems for Alanine, Serine, and Glycine in Escherichia coli K-12. J. Bacteriol. 116, 12–18.

Roberts, D.P., Denny, T.P., Schell, M.A., 1988. Cloning of the egl gene of Pseudomonas solanacearum and analysis of its role in phytopathogenicity. J. Bacteriol. 170, 1445–1451. https://doi.org/10.1128/jb.170.4.1445-1451.1988

Sabbagh, R.C.R., Carrere, S., Lonjon, F., Vailleau, F., Macho, A.P., Genin, S., Peeters, N., 2019. Pangenomic type III effector database of the plant pathogenic Ralstonia spp. PeerJ. https://doi.org/10.7717/peerj.7346

Safni, I., Cleenwerck, I., De Vos, P., Fegan, M., Sly, L., Kappler, U., 2014a. Polyphasic taxonomic revision of the Ralstonia solanacearum species complex: proposal to emend the descriptions of Ralstonia solanacearum and Ralstonia syzygii and reclassify current R. syzygii strains as Ralstonia syzygii subsp. syzygii subsp. nov., R. solanacearum phylotype IV strains as Ralstonia syzygii subsp. indonesiensis subsp. nov., banana

blood disease bacterium strains as Ralstonia syzygii subsp. celebesensis subsp. nov. and R. solanacearum phylotype I and III strains as Ralstonia pseudosolanacearum sp. nov. Int. J. Syst. Evol. Microbiol. 64, 3087–3103. https://doi.org/10.1099/ijs.0.066712-0

Safni, I., Cleenwerck, I., De Vos, P., Fegan, M., Sly, L., Kappler, U., 2014b. Polyphasic taxonomic revision of the Ralstonia solanacearum species complex: proposal to emend the descriptions of Ralstonia solanacearum and Ralstonia syzygii and reclassify current R. syzygii strains as Ralstonia syzygii subsp. syzygii subsp. nov., R. solanacearum phylotype IV strains as Ralstonia syzygii subsp. indonesiensis subsp. nov., banana blood disease bacterium strains as Ralstonia syzygii subsp. celebesensis subsp. nov. and R. solanacearum phylotype I and III strains as Ralstonia pseudosolanacearum sp. nov. Int. J. Syst. Evol. Microbiol. 64, 3087–3103. https://doi.org/10.1099/ijs.0.066712-0

Safni, I., Subandiyah, S., Fegan, M., 2018. Ecology, Epidemiology and Disease Management of Ralstonia syzygii in Indonesia. Front. Microbiol. 9, 419. https://doi.org/10.3389/fmicb.2018.00419

Sakata, N., Ishiga, T., Masuo, S., Hashimoto, Y., Ishiga, Y., 2021. Coronatine Contributes to *Pseudomonas cannabina* pv. *alisalensis* Virulence by Overcoming Both Stomatal and Apoplastic Defenses in Dicot and Monocot Plants. Mol. Plant-Microbe Interactions® 34, 746–757. https://doi.org/10.1094/MPMI-09-20-0261-R

Salanoubat, M., Genin, S., Artiguenave, F., Gouzy, J., Mangenot, S., Arlat, M., Billault, A., Brottier, P., Camus, J.C., Cattolico, L., Chandler, M., Choisne, N., Claudel-Renard, C., Cunnac, S., Demange, N., Gaspin, C., Lavie, M., Moisan, A., Robert, C., Saurin, W., Schiex, T., Siguier, P., Thébault, P., Whalen, M., Wincker, P., Levy, M., Weissenbach, J., Boucher, C.A., 2002. Genome sequence of the plant pathogen Ralstonia solanacearum. Nature 415, 497–502. https://doi.org/10.1038/415497a

Sanchez-Garrido, J., Ruano-Gallego, D., Choudhary, J.S., Frankel, G., 2022. The type III secretion system effector network hypothesis. Trends Microbiol. 30, 524–533. https://doi.org/10.1016/j.tim.2021.10.007

Savary, S., Willocquet, L., Pethybridge, S.J., Esker, P., McRoberts, N., Nelson, A., 2019a. The global burden of pathogens and pests on major food crops. Nat. Ecol. Evol. 3, 430–439. https://doi.org/10.1038/s41559-018-0793-y

Savary, S., Willocquet, L., Pethybridge, S.J., Esker, P., McRoberts, N., Nelson, A., 2019b. The global burden of pathogens and pests on major food crops. Nat. Ecol. Evol. 3, 430–439. https://doi.org/10.1038/s41559-018-0793-y

Sawinski, K., Mersmann, S., Robatzek, S., Böhmer, M., 2013. Guarding the Green: Pathways to Stomatal Immunity. Mol. Plant-Microbe Interactions® 26, 626–632. https://doi.org/10.1094/MPMI-12-12-0288-CR

Schandry, N., 2017. A Practical Guide to Visualization and Statistical Analysis of R. solanacearum Infection Data Using R. Front. Plant Sci. 8, 623. https://doi.org/10.3389/fpls.2017.00623

Scherf, J.M., Milling, A., Allen, C., 2010. Moderate temperature fluctuations rapidly reduce the viability of Ralstonia solanacearum race 3, biovar 2, in infected geranium, tomato, and potato plants. Appl. Environ. Microbiol. 76, 7061–7067. https://doi.org/10.1128/AEM.01580-10

Schönfeld, J., Heuer, H., Van Elsas, J.D., Smalla, K., 2003. Specific and sensitive detection of Ralstonia solanacearum in soil on the basis of PCR amplification of fliC fragments. Appl. Environ. Microbiol. 69, 7248–7256. https://doi.org/10.1128/AEM.69.12.7248-7256.2003

Schwyn, B., Neilands, J.B., 1987. Universal chemical assay for the detection and determination of siderophores. Anal. Biochem. 160, 47–56. https://doi.org/10.1016/0003-2697(87)90612-9

Scrimgeour, C., 2007. Handbook of Soil Analysis. Mineralogical, Organic and Inorganic Methods. By M. Pansu and J. Gautheyrou. Berlin, Heidelberg, New York: Springer (2006), pp. 993, £191.50. ISBN 978-3-540-31210-9. Exp. Agric. 43, 401–401. https://doi.org/10.1017/S0014479707005042

Seemann, T., 2014a. Prokka: Rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Seemann, T., 2014b. Prokka: Rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Shamsuddin, N., Lloyd, A., Graham, J., 1978. Survival of the potato strain of Pseudomonas solanacearum in soil. ET J. 44, 212–215.

Shapiro, B.J., Polz, M.F., 2014. Ordering microbial diversity into ecologically and genetically cohesive units. Trends Microbiol., Special Issue: Omics: Fulfilling the Promise 22, 235–247. https://doi.org/10.1016/j.tim.2014.02.006

Sharma, P., Johnson, M.A., Mazloom, R., Allen, C., Heath, L.S., Lowe-Power, T.M., Vinatzer, B.A., 2022. Meta-analysis of the Ralstonia solanacearum species complex (RSSC) based on comparative evolutionary genomics and reverse ecology. Microb. Genomics 8, 000791. https://doi.org/10.1099/mgen.0.000791

Sharma, P., Johnson, M.A., Mazloom, R., Allen, C., Heath, L.S., Lowe-Power, T.M., Vinatzer, B.A., 2021a. Meta Analysis of the Ralstonia solanacearum species complex (RSSC) based on comparative evolutionary genomics and reverse ecology. biorxiv 12.05.471342.

Sharma, P., Johnson, M.A., Mazloom, R., Allen, C., Heath, L.S., Lowe-Power, T.M., Vinatzer, B.A., 2021b. Meta Analysis of the Ralstonia solanacearum species complex (RSSC) based on comparative evolutionary genomics and reverse ecology. biorxiv 12.05.471342. https://doi.org/10.1101/2021.12.05.471342

Sheppard, S.K., Guttman, D.S., Fitzgerald, J.R., 2018. Population genomics of bacterial host adaptation. Nat. Rev. Genet. 19, 549–565. https://doi.org/10.1038/s41576-018-0032-z

Shipton, P.J., 1977. Monoculture and Soilborne Plant Pathogens. Annu. Rev. Phytopathol. 15, 387–407. https://doi.org/10.1146/annurev.py.15.090177.002131

Siddiqui, Z.A., Nesha, R., Singh, N., Alam, S., 2012. Interactions of Plant-Parasitic Nematodes and Plant-Pathogenic Bacteria, in: Maheshwari, D.K. (Ed.), Bacteria in Agrobiology: Plant Probiotics. Springer, Berlin, Heidelberg, pp. 251–267. https://doi.org/10.1007/978-3-642-27515-9_14

Siguier, P., Filee, J., Chandler, M., 2006. Insertion sequences in prokaryotic genomes. Current Opinion in Microbiology 9, 526–531. https://doi.org/10.1016/j.mib.2006.08.005

Singh, S., Gautam, R.K., Singh, D.R., Sharma, T.V.R.S., Sakthivel, K., Roy, S.D., 2015. Genetic approaches for mitigating losses caused by bacterial wilt of tomato in tropical islands. Eur. J. Plant Pathol. 143, 205–221. https://doi.org/10.1007/s10658-015-0690-z

Slatkin, M., 2008. Linkage disequilibrium — understanding the evolutionary past and mapping the medical future. Nat. Rev. Genet. 9, 477–485. https://doi.org/10.1038/nrg2361

Solé, M., Scheibner, F., Hoffmeister, A.-K., Hartmann, N., Hause, G., Rother, A., Jordan, M., Lautier, M., Arlat, M., Büttner, D., 2015. Xanthomonas campestris pv. vesicatoria Secretes Proteases and Xylanases via the Xps Type II Secretion System and Outer Membrane Vesicles. J. Bacteriol. 197, 2879–2893. https://doi.org/10.1128/JB.00322-15

Sonehara, K., Okada, Y., 2021. Genomics-driven drug discovery based on disease-susceptibility genes. Inflamm. Regen. 41, 1–5. https://doi.org/10.1186/s41232-021-00158-7

Sotiropoulos, A.G., Arango-Isaza, E., Ban, T., Barbieri, C., Bourras, S., Cowger, C., Czembor, P.C., Ben-David, R., Dinoor, A., Ellwood, S.R., Graf, J., Hatta, K., Helguera, M., Sánchez-Martín, J., McDonald, B.A., Morgounov, A.I., Müller, M.C., Shamanin, V., Shimizu, K.K., Yoshihira, T., Zbinden, H., Keller, B., Wicker, T., 2022. Global genomic analyses of wheat powdery mildew reveal association of pathogen spread with historical human migration and trade. Nat. Commun. 13, 4315. https://doi.org/10.1038/s41467-022-31975-0

Souza, D.O., Bruckner, F.P., Gonçalves, O.S., Santana, F., Alfenas-zerbini, P., 2021. Genomics Widespread distribution of prophages signaling the potential for adaptability and pathogenicity evolution of Ralstonia solanacearum species complex 113, 992–1000. https://doi.org/10.1016/j.ygeno.2021.02.011

Stackebrandt, E., Frederiksen, W., Garrity, G.M., Grimont, P.A.D., Kämpfer, P., Maiden, M.C.J., Nesme, X., Rosselló-Mora, R., Swings, J., Trüper, H.G., Vauterin, L., Ward, A.C., Whitman, W.B., 2002. Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. Int. J. Syst. Evol. Microbiol. 52, 1043–1047. https://doi.org/10.1099/00207713-52-3-1043

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stevens, P., van Elsas, J.D., 2010. Genetic and phenotypic diversity of Ralstonia solanacearum biovar 2 strains obtained from Dutch waterways. Antonie Van Leeuwenhoek 97, 171–188. https://doi.org/10.1007/s10482-009-9400-1

Straub, C., Colombi, E., McCann, H.C., 2021a. Population genomics of bacterial plant pathogens. Phytopathology 111, 23–31. https://doi.org/10.1094/PHYTO-09-20-0412-RVW

Straub, C., Colombi, E., McCann, H.C., 2021b. Population Genomics of Bacterial Plant Pathogens. Phytopathology® 111, 23–31. https://doi.org/10.1094/PHYTO-09-20-0412-RVW

Su, Y., Xu, Y., Li, Q., Yuan, G., Zheng, D., 2020. The essential genome of Ralstonia solanacearum. Microbiol. Res. 238, 126500. https://doi.org/10.1016/j.micres.2020.126500

Sui, S.J.H., Fedynak, A., Hsiao, W.W.L., Langille, M.G.I., Brinkman, F.S.L., 2009. The Association of Virulence Factors with Genomic Islands. PLOS ONE 4, e8094. https://doi.org/10.1371/journal.pone.0008094

Sun, Y., Wang, K., Caceres-Moreno, C., Jia, W., Chen, A., Zhang, H., Liu, R., Macho, A.P., 2017. Genome sequencing and analysis of Ralstonia solanacearum phylotype I strains FJAT-91, FJAT-452 and FJAT-462 isolated from tomato, eggplant, and chili pepper in China. Stand. Genomic Sci. 12. https://doi.org/10.1186/s40793-017-0241-7

Sundin, George W, Wang, N., 2018. Antibiotic resistance in plant-pathogenic bacteria. Annu. Rev. Phytopathol. 56, 161–180. https://doi.org/10.1146/annurev-phyto-080417-045946

Sundin, George W., Wang, N., 2018. Antibiotic resistance in plant-pathogenic bacteria. Annu. Rev. Phytopathol. 56, 161–180. https://doi.org/10.1146/annurev-phyto-080417-045946

Tans-Kersten, J., Huang, H., Allen, C., 2001. Ralstonia solanacearum Needs Motility for Invasive Virulence on Tomato. J. Bacteriol. 183, 3597–3605. https://doi.org/10.1128/JB.183.12.3597-3605.2001

Tantoso, E., Eisenhaber, B., Kirsch, M., Shitov, V., Zhao, Z., Eisenhaber, F., 2022. To kill or to be killed: pangenome analysis of Escherichia coli strains reveals a tailocin specific for pandemic ST131. BMC Biol. 20, 146. https://doi.org/10.1186/s12915-022-01347-7

Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E.P., Zaslavsky, L., Lomsadze, A., Pruitt, K.D., Borodovsky, M., Ostell, J., 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 44, 6614–6624. https://doi.org/10.1093/nar/gkw569

Tettelin, H., Masignani, V., Cieslewicz, M.J., Donati, C., Medini, D., Ward, N.L., Angiuoli, S.V., Crabtree, J., Jones, A.L., Durkin, A.S., DeBoy, R.T., Davidsen, T.M., Mora, M., Scarselli, M., Margarit y Ros, I., Peterson, J.D., Hauser, C.R., Sundaram, J.P., Nelson, W.C., Madupu, R., Brinkac, L.M., Dodson, R.J., Rosovitz, M.J., Sullivan, S.A., Daugherty, S.C., Haft, D.H., Selengut, J., Gwinn, M.L., Zhou, L., Zafar, N., Khouri, H., Radune, D., Dimitrov, G., Watkins, K., O'Connor, K.J.B., Smith, S., Utterback, T.R., White, O., Rubens, C.E., Grandi, G., Madoff, L.C., Kasper, D.L., Telford, J.L., Wessels, M.R., Rappuoli, R., Fraser, C.M., 2005. Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: Implications for the microbial "pan-genome." Proc. Natl. Acad. Sci. 102, 13950–13955. https://doi.org/10.1073/pnas.0506758102

Tettelin, H., Medini, D. (Eds.), 2020. The Pangenome. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-38281-0

Tomlinson, D.L., Elphinstone, J.G., Soliman, M.Y., Hanafy, M.S., Shoala, T.M., El-Fatah, H.A., Agag, S.H., Kamal, M., El-Aliem, M.M.A., Fawzi, F.G., Stead, D.E., Janse, J.D., 2009. Recovery of Ralstonia solanacearum from canal water in traditional potato-growing areas of Egypt but not from designated Pest-Free Areas (PFAs). Eur. J. Plant Pathol. 125, 589–601. https://doi.org/10.1007/s10658-009-9508-1

Tonkin-Hill, G., MacAlasdair, N., Ruis, C., Weimann, A., Horesh, G., Lees, J.A., Gladstone, R.A., Lo, S., Beaudoin, C., Floto, R.A., Frost, S.D.W., Corander, J., Bentley, S.D., Parkhill, J., 2020. Producing polished prokaryotic pangenomes with the Panaroo pipeline. Genome Biol. 21, 180. https://doi.org/10.1186/s13059-020-02090-4

Tsumori, C., Matsuo, S., Murai, Y., Kai, K., 2022. Quorum sensing-dependent invasion of *Ralstonia solanacearum* into *Fusarium oxysporum* chlamydospores (preprint). Microbiology. https://doi.org/10.1101/2022.11.03.515128

Ujita, Y., Sakata, M., Yoshihara, A., Hikichi, Y., Kai, K., 2019. Signal Production and Response Specificity in the phc Quorum Sensing Systems of Ralstonia solanacearum Species Complex. ACS Chem. Biol. 14, 2243–2251. https://doi.org/10.1021/acschembio.9b00553

Um, H.Y., Kong, H.G., Lee, H.J., Choi, H.K., Park, E.J., Kim, S.T., Murugiyan, S., Chung, E., Kang, K.Y., Lee, S.-W., 2013. Altered gene expression and intracellular changes of the viable but nonculturable state in Ralstonia solanacearum by copper treatment. Plant Pathol. J. 29, 374–385. https://doi.org/10.5423/PPJ.OA.07.2013.0067

Van Elsas, J.D., Kastelein, P., De Vries, P.M., Van Overbeek, L.S., 2001. Effects of ecological factors on the survival and physiology of Ralstonia solanacearum bv. 2 in irrigation water. Can. J. Microbiol. 47, 842–854. https://doi.org/10.1139/w01-084

Van Elsas, J.D., Kastelein, P., Van Bekkum, P., Van der Wolf, J.M., De Vries, P.M., Overbeek, L.S. van, 2000. Survival of Ralstonia solanacearum biovar 2, the causative agent of potato brown rot, in field and microcosm soils in temperate climates. Phytopathology 90, 1358–1366. https://doi.org/10.1094/phyto.2000.90.12.1358

Vasse, J., Frey, P., Trigalet, A., 1995. Microscopic studies of intercellular infection and protoxylem invasion of tomato roots by Pseudomonas solanacearum. https://doi.org/10.1094/MPMI-8-0241

Velásquez, A.C., Castroverde, C.D.M., He, S.Y., 2018. Plant-pathogen warfare under changing climate conditions. Curr. Biol. CB 28, R619–R634. https://doi.org/10.1016/j.cub.2018.03.054

Von Bodman, S.B., Bauer, W.D., Coplin, D.L., 2003. Quorum sensing in plant-pathogenic bacteria. Annu. Rev. Phytopathol. 41, 455–482. https://doi.org/10.1146/annurev.phyto.41.052002.095652

Vos, M., 2009. Why do bacteria engage in homologous recombination? Trends Microbiol. 17, 226–232. https://doi.org/10.1016/j.tim.2009.03.001

Waszczak, C., Carmody, M., Kangasjärvi, J., 2018. Reactive Oxygen Species in Plant Signaling. Annu. Rev. Plant Biol. 69, 209–236. https://doi.org/10.1146/annurev-arplant-042817-040322

Wei, Z., Gu, Y., Friman, V.-P., Kowalchuk, G.A., Xu, Y., Shen, Q., Jousset, A., 2019. Initial soil microbiome composition and functioning predetermine future plant health. Sci. Adv. 5, eaaw0759. https://doi.org/10.1126/sciadv.aaw0759

Wei, Z., Yang, T., Friman, V.-P., Xu, Y., Shen, Q., Jousset, A., 2015a. Trophic network architecture of root-associated bacterial communities determines pathogen invasion and plant health. Nat. Commun. 6, 8413. https://doi.org/10.1038/ncomms9413

Wei, Z., Yang, T., Friman, V.-P., Xu, Y., Shen, Q., Jousset, A., 2015b. Trophic network architecture of root-associated bacterial communities determines pathogen invasion and plant health. Nat. Commun. 6, 8413. https://doi.org/10.1038/ncomms9413

Wei, Z., Yang, X., Yin, S., Shen, Q., Ran, W., Xu, Y., 2011. Efficacy of Bacillus-fortified organic fertiliser in controlling bacterial wilt of tomato in the field. Appl. Soil Ecol. 48, 152–159. https://doi.org/10.1016/j.apsoil.2011.03.013

Weller, S.A., Elphinstone, J.G., Smith, N.C., Boonham, N., Stead, D.E., 2000. Detection of Ralstonia solanacearum Strains with a Quantitative, Multiplex, Real-Time, Fluorogenic PCR (TaqMan) Assay, APPLIED AND ENVIRONMENTAL MICROBIOLOGY.

Whelan, F.J., Rusilowicz, M., McInerney, J.O., 2020. Coinfinder: detecting significant associations and dissociations in pangenomes. Microb. Genomics 6. https://doi.org/10.1099/mgen.0.000338

Wick, R.R., Judd, L.M., Gorrie, C.L., Holt, K.E., 2017a. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput. Biol. 13, 1–22. https://doi.org/10.1371/journal.pcbi.1005595

Wick, R.R., Judd, L.M., Gorrie, C.L., Holt, K.E., 2017b. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput. Biol. 13, 1–22. https://doi.org/10.1371/journal.pcbi.1005595

Wicker, E., Grassart, L., Coranson-Beaudu, R., Mian, D., Guilbaud, C., Fegan, M., Prior, P., 2009. Ralstonia solanacearum strains from Martinique (French West Indies) exhibiting a new pathogenic potential. Appl. Environ. Microbiol. 75, 558. https://doi.org/10.1128/AEM.02635-08

Wicker, E., Lefeuvre, P., de Cambiaire, J.-C., Lemaire, C., Poussier, S., Prior, P., 2012a. Contrasting recombination patterns and demographic histories of the plant pathogen Ralstonia solanacearum inferred from MLSA. ISME J. 6, 961–974. https://doi.org/10.1038/ismej.2011.160

Wicker, E., Lefeuvre, P., de Cambiaire, J.-C., Lemaire, C., Poussier, S., Prior, P., 2012b. Contrasting recombination patterns and demographic histories of the plant pathogen Ralstonia solanacearum inferred from MLSA. ISME J. 6, 961–974. https://doi.org/10.1038/ismej.2011.160

Wickham, R., 2009. ggplot2: Elegant Graphics for Data Analysis.

Wilson, D.B., 2011. Microbial diversity of cellulose hydrolysis. Curr. Opin. Microbiol., Ecology and industrial microbiology / Special section: Archaea 14, 259–263. https://doi.org/10.1016/j.mib.2011.04.004

Wood, D.E., Lu, J., Langmead, B., 2019. Improved metagenomic analysis with Kraken 2. Genome Biol. 20, 257. https://doi.org/10.1186/s13059-019-1891-0

Worning, P., Jensen, L.J., Nelson, K.E., Brunak, S., Ussery, D.W., 2000. Structural analysis of DNA sequence: evidence for lateral gene transfer in Thermotoga maritima. Nucleic Acids Res. 28, 706–709.

Wu, D., Ding, W., Zhang, Y., Liu, X., Yang, L., 2015. Oleanolic Acid Induces the Type III Secretion System of Ralstonia solanacearum. Front. Microbiol. 6.

Xie, Z., Tang, H., 2017. ISEScan: automated identification of insertion sequence elements in prokaryotic genomes. Bioinformatics 33, 3340–3347. https://doi.org/10.1093/bioinformatics/btx433

Xin, X.-F., Kvitko, B., He, S.Y., 2018. Pseudomonas syringae: what it takes to be a pathogen. Nat. Rev. Microbiol. 16, 316–328. https://doi.org/10.1038/nrmicro.2018.17

Xue, Q.Y., Yin, Y.N., Yang, W., Heuer, H., Prior, P., Guo, J.H., Smalla, K., 2011. Genetic diversity of Ralstonia solanacearum strains from China assessed by PCR-based fingerprints to unravel host plant- and site-dependent distribution patterns. FEMS Microbiol. Ecol. 75, 507–519. https://doi.org/10.1111/j.1574-6941.2010.01026.x

Yamada, T., 2016. Bacteriophages of Ralstonia solanacearum: Their Diversity and Utilization as Biocontrol Agents in Agriculture 13.

Yamada, T., 2013. Filamentous phages of Ralstonia solanacearum: double-edged swords for pathogenic bacteria. Front. Microbiol. 4.

Yang, C., Dong, Y., Friman, V.P., Jousset, A., Wei, Z., Xu, Y., Shen, Q., 2019. Carbon resource richness shapes bacterial competitive interactions by alleviating growth-antibiosis trade-off. Funct. Ecol. 33, 868–875. https://doi.org/10.1111/1365-2435.13292

Yang, T., Han, G., Yang, Q., Friman, V.P., Gu, S., Wei, Z., Kowalchuk, G.A., Xu, Y., Shen, Q., Jousset, A., 2018. Resource stoichiometry shapes community invasion resistance via productivity-mediated species identity effects. Proc. Biol. Sci. 285, 20182035. https://doi.org/10.1098/rspb.2018.2035

Yao, J., Allen, C., 2007. The plant pathogen Ralstonia solanacearum needs aerotaxis for normal biofilm formation and interactions with its tomato host. J. Bacteriol. 189, 6415–6424. https://doi.org/10.1128/JB.00398-07

Yoshimochi, T., Hikichi, Y., Kiba, A., Ohnishi, K., 2009. The Global Virulence Regulator PhcA Negatively Controls the Ralstonia solanacearum hrp Regulatory Cascade by Repressing

Expression of the PrhIR Signaling Proteins. J. Bacteriol. 191, 3424–3428. https://doi.org/10.1128/JB.01113-08

Yu, G., Lam, T.T.Y., Zhu, H., Guan, Y., 2018. Two methods for mapping and visualizing associated data on phylogeny using GGTree. Mol. Biol. Evol. 35, 3041–3043. https://doi.org/10.1093/molbev/msy194

Yuliar, Nion, Y.A., Toyota, K., 2015. Recent trends in control methods for bacterial wilt diseases caused by Ralstonia solanacearum. Microbes Environ. 30, 1–11. https://doi.org/10.1264/jsme2.ME14144

Zhang, J., Shao, F., Li, Y., Cui, H., Chen, L., Li, H., Zou, Y., Long, C., Lan, L., Chai, J., Chen, S., Tang, X., Zhou, J.-M., 2007. A Pseudomonas syringae Effector Inactivates MAPKs to Suppress PAMP-Induced Immunity in Plants. Cell Host Microbe 1, 175–185. https://doi.org/10.1016/j.chom.2007.03.006

Zhang, W., He, S.Y., Assmann, S.M., 2008. The plant innate immunity response in stomatal guard cells invokes G-protein-dependent ion channel regulation. Plant J. Cell Mol. Biol. 56, 984–996. https://doi.org/10.1111/j.1365-313X.2008.03657.x

Zheng, X., Zhu, Y., Liu, B., Zhou, Y., Che, J., Lin, N., 2014. Relationship Between Ralstonia solanacearum Diversity and Severity of Bacterial Wilt Disease in Tomato Fields in China. J. Phytopathol. 162, 607–616. https://doi.org/10.1111/jph.12234

Zulperi, D., Sijam, K., Mior Ahmad, Z.A., Awang, Y., Sulaiman Rashid, T., 2014. Occurrence of Ralstonia solanacearum Race 2 Biovar 1 Associated with Moko Disease of Banana (Musa paradisiaca cv. Nipah) in Malaysia. J. Phytopathol. 162, 697–702. https://doi.org/10.1111/jph.12233