Genetic basis of the interaction between *Stenotrophomonas maltophilia* and *Caenorhabditis elegans* from both host and pathogen perspectives

by

Leah Jean Radeke

B.S., University of Wisconsin-Eau Claire, 2015

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Approved by:

Major Professor
Dr. Michael Herman

Copyright

Abstract

Stenotrophomonas maltophilia is an opportunistic bacterial pathogen found ubiquitously in the environment. Although *S. maltophilia* is an emerging pathogen associated with hospital-acquired infections in patients with respiratory diseases, particularly cystic fibrosis, very little is known about its mechanism of pathogenesis in any system. In addition, *S. maltophilia* isolates vary in pathogenicity to several hosts and are genetically diverse, including variation in virulence factors. In this thesis, I address the genetic basis of *S. maltophilia* pathogenesis from both host and bacterial perspectives. Our lab has previously developed *Caenorhabditis elegans* as a model for *S. maltophilia* infection. *Stenotrophomonas* is found in relatively high abundance in the microbiome of *C. elegans*, making it a suitable platform for studying *S. maltophilia*-host interactions. I performed a transcriptomic analysis to determine *C. elegans* responses to several *S. maltophilia* strains of varying pathogenicity. Treatments included K279a, an avirulent clinical isolate, JCMS, a virulent environmental strain isolated in association with nematodes near Manhattan, KS, and JV3, an even more virulent environmental isolate. Overall, I found that most genes (89%) that are differentially expressed in response to pathogenic *S. maltophilia* strains are upregulated, with many even further upregulated in response to the more virulent strain, JV3. Using information from a variety of transcriptomic datasets, I found that most of these genes are also commonly differentially expressed in *C. elegans* in response to other pathogens. Many more genes were differentially expressed specifically in response to JV3 when compared to all other strains (221 genes) than JCMS as compared to all other strains (14 genes), suggesting JV3 has unique virulence mechanisms that could explain its observed increased virulence. Candidate genes were chosen from the above differentially expressed gene sets (differentially expressed in response to both pathogenic *S. maltophilia* strains or in a strain-specific manner) for functional

analysis. Mutational analysis of these candidate genes revealed that several mutants caused increased susceptibility of *C. elegans* to pathogenic *S. maltophilia*, regardless of the strain(s) that caused differential expression of that gene. Furthermore, many of these mutants also caused increased susceptibility to K279a, suggesting that K279a may also employ virulence mechanisms that *wild-type C. elegans* are able to defend against. To address the pathogen side of the interaction, we analyzed draft assemblies of the *S. maltophilia* strains, with the addition of another slightly pathogenic environmental strain, R551-3. We hypothesized that differences in observed pathogenicity and host responses to strains of *S. maltophilia* could be explained by differences in their genomes. When comparing draft assemblies to their respective reference genomes, few differences were observed. However, several genomic features were present in some strains and absent in others, including components of the CmeABC efflux pump and the Type IV secretion system, that might play a role in different virulence mechanisms. Genome-wide comparison of shared and unique genetic features across many *S. maltophilia* strains revealed that most *S. maltophilia* genes are strain-specific, suggesting that many potential virulence factors are unique and have yet to be functionally analyzed. Overall, variation in observed pathogenicity, differences in host transcriptional responses, and comparative genomics of *S. maltophilia* strains reveal that strain-specific mechanisms play important roles in *S. maltophilia* pathogenesis.

Table of Contents

List of Figures

List of Tables

Acknowledgements

My graduate research was funded by the National Science Foundation (NSF) and the Kansas State University Division of Biology. The RNA-sequencing was funded by a COBRE grant (P20-GM103638) and performed at the KU genome sequencing center (GSC). I would like to thank Michael Herman and past and present members of the Herman lab for physical and intellectual input in this project. I would also like to thank Wendy Reeves and Brad Olson for their help with RNA-sequencing analysis, and Dave Wheeler for his help with genome assembly and analysis.

Chapter 1 - Literature Review

Nematodes are the most abundant multicellular organism on Earth and consist of over 25,000 species, both free-living and parasitic (Zhang, 2013). *Caenorhabditis elegans* are free-living nematodes that can be found in the natural environment world-wide (Barriere & Felix, 2007; Barrière & Félix, 2005; Dirksen et al., 2016; Félix & Duveau, 2012; Haber et al., 2018). They feed on a variety of bacteria found in rotting organic matter and are most commonly isolated from rotting fruits, flowers, and stems (reviewed in Schulenburg & Félix, 2017). Therefore, in the natural environment, *C. elegans* are in constant contact with many organisms, including other small invertebrates, bacteria, and fungi. *C. elegans* travel between locations via vectors, such as isopods and snails (Barrière & Félix, 2005; Félix & Duveau, 2012). Nematodes are also prey to a variety of insects and fungi, in addition to playing host to a number of pathogenic and symbiotic bacteria (Dirksen et al., 2016; reviewed in Schulenburg & Félix, 2017) (Figure 1.1).

C. elegans is an excellent genetic model organism for studying many fields of biology, including development, cell biology, innate immunity, and neurobiology. In the laboratory, *C. elegans* are grown on plates seeded with *Escherichia coli* OP50. This artificial lab setting differs significantly from the natural environment and lacks ecologically important biotic components. In addition, many recent studies have determined that bacteria, both in the surrounding environment and the microbiome of many organisms, influence behavior, aging, and overall health (reviewed in Chilton et al., 2015; Dirksen et al., 2016; Ikeda et al., 2007; Rae et al., 2008; Samuel, et al., 2016). These observations led to studies investigating the effects of diverse microorganisms, including bacteria, fungi, and viruses, on *C. elegans*.

C. elegans are advantageous for use in studying animal-microbe interactions because they have many progeny (up to 300), a short generation time (3.5 days), and small size (1 mm). They are also transparent, which allows for easy visualization of organism development, bacterial colonization in the pharynx and intestine, and intracellular properties such as protein localization and expression. In addition, there are a variety of tools for determining the health of *C. elegans*, including fecundity, lifespan, and stress response assays. *C. elegans* are also amenable to genetic manipulation, allowing for an in depth understanding of the genetics underlying host-microbe interactions.

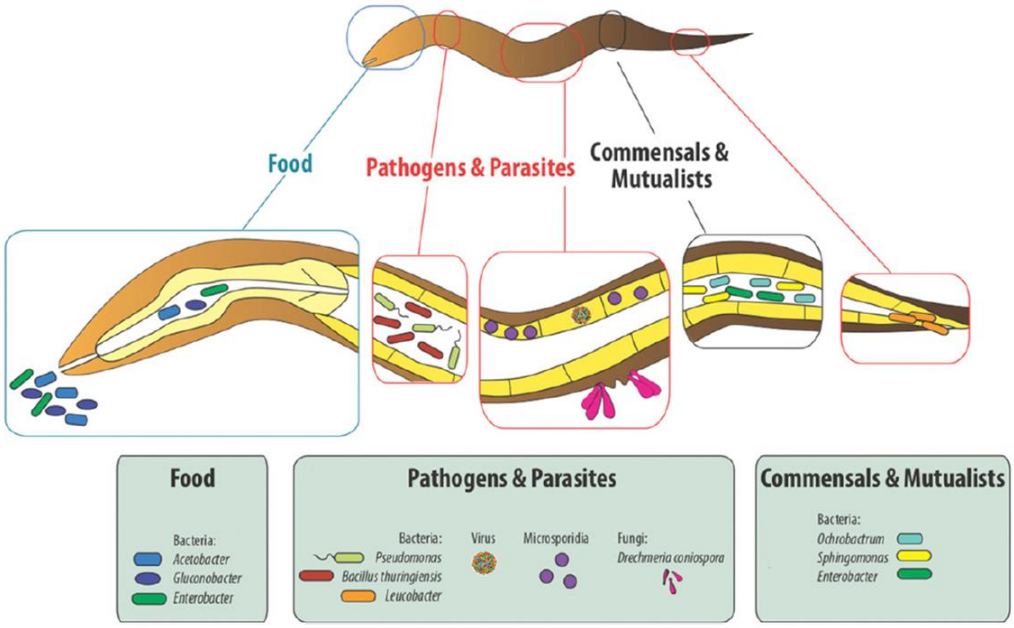

Figure 1.1 Overview of native microorganism interactions with *C. elegans*

Adapted from Schulenburg & Felix, 2017. Highlights interactions such as bacterial food (*Acetobacter*, *Gluconobacter*, and *Enterobacter*), pathogens/parasites (Orsay virus, microsporidia, the fungus *D. coniospora*, and bacteria *Pseudomonas*, *B. thuringiensis*, and *Leucobacter*), and commensals/mutualists (the bacterial genera *Ochrobactrum*, *Spingomonas*, and *Enterobacter*).

C. elegans-bacterial interactions

Bacteria that interact with *C. elegans* can be classified as prey, being consumed by *C. elegans*, mutualists, providing nutrients through metabolism, or pathogens. These interactions are dynamic, and a single species can fit into more than one category. This compound relationship causes varying responses to bacteria that lead to appropriate actions, including a neuronal response, a nutrition response, and a pathogen response (Kim, 2013) (Figure 1.2). The neuronal response is the detection of bacteria as either food or pathogen and leads to further neuronal and endocrine signaling that can affect behavior and longevity (Kim, 2013). For example, *C. elegans* preferentially chose food that promotes growth, and this behavior is dependent on the amphid AIY neurons (Shtonda & Avery, 2006). There is also evidence that neuroendocrine signaling is involved in recognition of pathogens. For example, *C. elegans* can learn to avoid pathogenic bacteria such as *Pseudomonas aeruginosa* and *Serratia marcescens* through a serotonin signaling pathway (Zhang, 2008). Furthermore, signaling between neurons and the intestine plays a role in the immune response to some pathogenic bacteria (Kawli & Tan, 2008; Styer et al., 2008).

The nutrition response is based on the nutritional value of bacteria based on metabolites that the bacteria produces. For example, exposure of *C. elegans* to *E. coli* mutants that produce less folate caused decreased lifespan (Virk et al., 2012). Although the nutrition response is independent of induction of stress or defense responses caused by exposure to bacteria, unravelling the nutritional value of a particular bacterium from its pathogenic potential has proven to be difficult (Kim, 2013). For example, further characterization of the *E. coli* folate synthesis mutants determined that mutations in these genes also decreased the virulence of pathogenic bacteria (Virk et al., 2016). It was also discovered that *E. coli* OP50 is mildly pathogenic, as killing the bacterial cells with UV or kanamycin inhibits bacterial accumulation in

the intestine and increases longevity (Garigan et al., 2002). Therefore, folate abundance produced by bacteria may play a role in pathogen response in *C. elegans* rather than simply changing the nutritional value of the bacteria (Virk et al., 2016).

The pathogen response depends on the activation of several innate immune, defense, and stress pathways that are discussed in detail in the next section. A shift from predator-prey to pathogen-host often occurs as *C. elegans* age and correlates with accumulation of bacteria within the intestine (Garigan et al., 2002). This shift involves three stages: 1) predation, which involves mastication, or break down, of bacteria in the pharyngeal grinder followed by uptake of nutrients from bacterial cell material in the intestine, 2) symbiosis, in which live bacteria that are able to survive pharyngeal grinding inhabit the intestine and provide nutrients to the nematode through metabolism, and 3) dysbiosis, where bacteria accumulate in the intestine and cause damage to tissues (Cabreiro & Gems, 2013) (Figure 1.2). This change is dependent on several factors, including efficiency of the pharynx, the ability of bacteria to proliferate in the intestine, and the capability of the host to reduce bacterial accumulation through defense responses.

The interplay of these responses leads to a multi-faceted relationship between *C. elegans* and bacteria. In the natural environment, for instance, these responses occur simultaneously with a multitude of bacteria. Unraveling these responses involves simplifying interactions encountered in the natural environment. The most common approach is to study the response of *C. elegans* to individual bacterial species. In particular, identifying bacteria that are detrimental to *C. elegans* and determining how *C. elegans* attempts to counteract this damage will provide insight into the underlying mechanisms of the pathogen response.

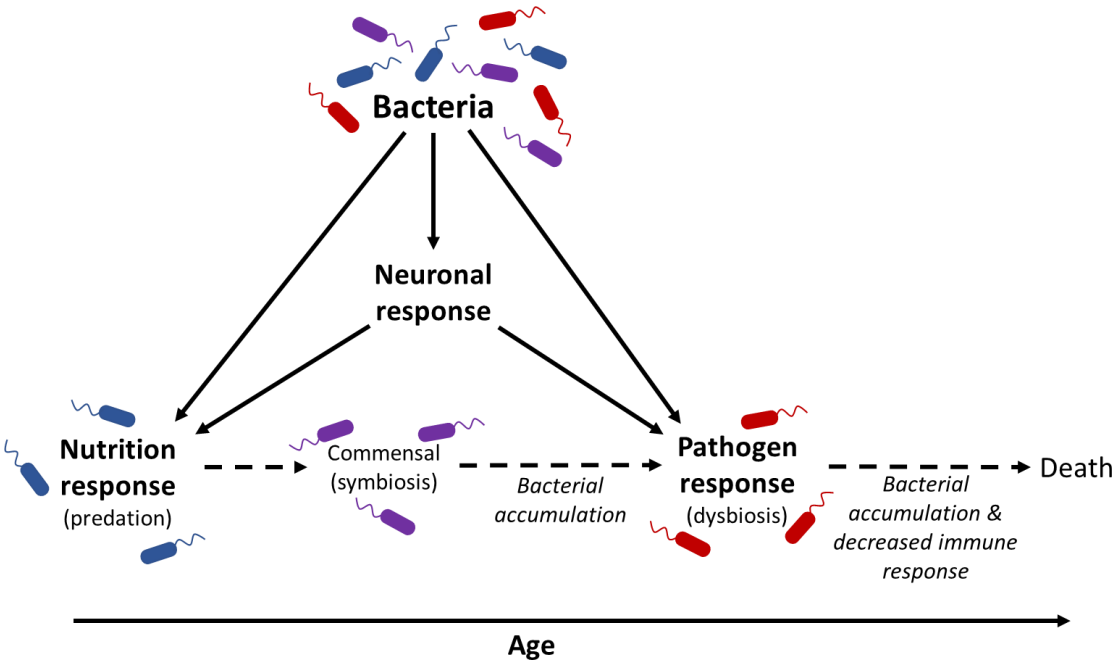

Figure 1.2 The dynamic interactions between *C. elegans* and bacteria

This figure depicts the response types and interactions of responses by *C. elegans* as they encounter bacteria, including the neuronal response, nutrition response, and pathogen response. It also illustrates the shift from predator-prey to pathogen-host as *C. elegans* age.

C. elegans as an innate immunity model

The study of *C. elegans* genes and pathways involved in pathogen response is medically relevant, as many of these genes and pathways are conserved. Since approximately 40% of genes found in *C. elegans* have orthologs in humans, many processes in *C. elegans* are conserved in mammals (Shaye & Greenwald, 2011). In fact, conservation of innate immune genes between nematodes, insects, and mammals has revealed important immune factors in *C. elegans,* indicating similarities between innate immunity in *C. elegans* and other metazoa (reviewed in Dierking, Yang, & Schulenburg, 2016; Garsin et al., 2003; Mallo et al., 2002; Shivers et al., 2010). However, *C. elegans* do not have dedicated innate immune cells as found in vertebrates. Therefore, their immune responses usually occur at physical barriers where pathogenesis begins, such as the cuticle and intestine (Kim & Ewbank, 2015).

More recently, innate immune responses in *C. elegans* have been studied in ecologically-relevant contexts (e.g. Boehnisch et al., 2011; O'Rourke et al., 2006; White et al., 2016). Studying interactions with species encountered in the natural environment of *C. elegans* ensures realistic responses and provides insight into the complex microbiome of *C. elegans*. Some pathogenic species that have been studied because of their ecological importance include *Stenotrophomonas maltophilia, Pseudomonas aeruginosa,* and *Bacillus thuringiensis* (White et al., 2016; Kirienko et al., 2014; Tan, Mahajan-Miklos, & Ausubel, 1999; Boehnisch et al., 2011; Feinbaum et al., 2012; Huffman et al., 2004). The identification of natural bacteria of the microbiome of *C. elegans* and their virulence is reviewed in more detail in the next section. Genes and pathways involved in these defense responses to pathogens are discovered by forward genetic screens, analysis of differentially expressed genes after pathogen exposure, and identification of orthologs between organisms (e.g. Kim et al., 2002; Shivers et al., 2010;

(Shivers et al., 2010) Troemel et al 2006; Murphy et al., 2003; Mallo et al., 2002). These techniques have led to the identification of many conserved and novel mechanisms employed by *C. elegans* to defend against pathogens.

In general, the innate immune response involves three steps, each of which is carried out by different classes of proteins (Kim & Ewbank, 2015). The first step is the recognition of the pathogen. This step can be species specific, by proteins that recognize particular toxins or bacterial proteins, or more general, responding to pathogen-induced damage. The second step involves signaling pathways which activate downstream proteins and eventually transcription factors (Figure 1.3). The final step involves the genes that are regulated downstream of signaling pathways that serve as effector molecules, including anti-microbial peptides (AMPs) (reviewed in Kim & Ewbank, 2015). These steps are detailed below, with an emphasis on responses to bacterial pathogenesis in the intestine.

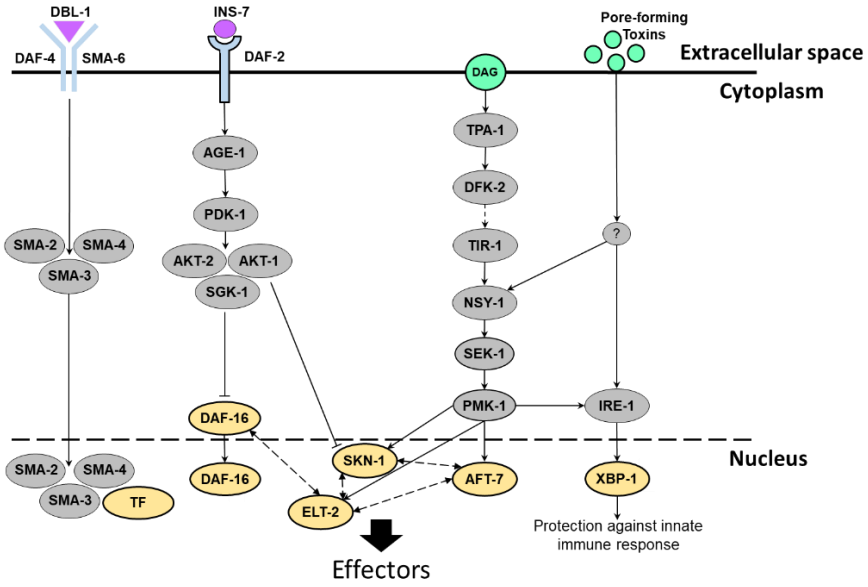

Figure 1.3 Overview of innate immune and defense pathways

Includes TGFβ pathway, DAF-2/16 insulin-like pathway, p38 MAPK pathway, and unfolded protein response (UPR) pathway. Purple indicates ligands, blue indicates receptors, and gray indicates downstream signaling components, and gold indicates transcription factors. DAG= diacylglycerol, TF= transcription factor.

Pathogen recognition proteins

Pathogen recognition can result from direct recognition of structural components or secreted proteins of the pathogen, termed microbe-associated molecular patterns (MAMPs), or indirectly via perturbations induced by infection, termed damage-associated molecular patterns (DAMPs). Interestingly, many conserved receptors involved in MAMP recognition, including peptidoglycan recognition proteins, Gram-negative binding proteins, or nucleotide-binding oligomerization domain (NOD)-like receptors, are not found in *C. elegans* (Kim & Ewbank, 2015). Toll-like pathway receptors play a role in immunity in insects and higher-order metazoa (Lindsay & Wasserman, 2014; Liu & Zhao, 2007). The sole Toll-like receptor (TLR) in *C. elegans*, TOL-1, is vital for proper development and function of sensory neurons (Brandt & Ringstad, 2015). Although the p38 mitogen activated protein kinase (MAPK) pathway downstream of the TLR is important for response to many pathogens in *C. elegans*, TOL-1 does not appear to play a role in recognition of or response to *Staphylococcus aureus* or *Pseudomonas aeruginosa* (Kim et al., 2002; Irazoqui et al., 2010; Pujol et al., 2001). The lack of traditional MAMP recognition mechanisms suggests that pathways involved in innate immunity in *C. elegans* may not be responding directly to the pathogen, but instead to cell damage or other stressors that are a consequence of infection.

In other animals, C-type lectin domain (CTLD) containing proteins recognize and bind bacterial cell walls via pathogen recognition receptors (reviewed in van den Berg, Gringhuis, & Geijtenbeek, 2012). In *C. elegans*, two C-type lectin domain proteins, CLEC-39 and CLEC-49, can directly bind *S. marcescens*, and mutations that inactivate these genes cause increased susceptibility to *S. marcescens* infection (Miltsch, Seeberger, & Lepenies, 2014). However, whether CTLD containing proteins function in pathogen recognition and activation of innate

immune pathways is unknown. These genes are also differentially expressed in response to several pathogens (Irazoqui et al., 2010a; Yang et al., 2015) and are regulated by innate immune pathways (Alper et al., 2007; Troemel et al., 2006). Therefore, it is unclear if these proteins play a role in recognizing pathogens or function as downstream antimicrobial peptides.

The lack of conserved receptors suggests that there may be non-canonical mechanisms involved in pathogen recognition. Recent evidence has shown that the nervous system may be involved in upstream signaling that leads to pathogen responses in the intestine. INS-7, an insulin-like ligand that binds to and activates the insulin-like receptor DAF-2 (discussed below), is expressed mainly in neuronal cells and at low levels in the intestine (Murphy, Lee, & Kenyon, 2007). *ins-7* expression is increased upon exposure to *P. aeruginosa* via excretion from dense core vesicles, leading to activation of the DAF-2 pathway and suppression of the DAF-16 transcription factor (Evans, Kawli, & Tan, 2008; Kawli & Tan, 2008). Mutations in *ins-7* cause resistance to *P. aeruginosa*, and transgenic expression of INS-7 in neuronal cells alone is able to suppress *ins-7* mutant resistance (Evans, Kawli, & Tan, 2008; Kawli & Tan, 2008). This suggests that not only do neuronal cells play a role in innate immunity, but INS-7 alone is able to activate the DAF-2 pathway and decrease resistance to pathogens. The mechanism of how INS-7 activates DAF-2 remains unclear, but it may be that INS-7 is able to recognize pathogens or pathogen-induced damage.

Pathogen and defense response signaling pathways

p38 MAPK pathway

Mitogen-activated protein kinase (MAPK) pathways play significant roles in a variety of cellular responses such as development, differentiation, stress, and apoptosis (Cargnello & Roux, 2011). The p38 MAPK pathway in mammals is activated by cytokines and other stressors in

immune cells (Johnson & Lapadat, 2002). Analysis of *C. elegans* mutants that enhanced susceptibility to *P. aeruginosa* PA14 led to the identification of conserved components of the p38 MAPK signaling pathway including SEK-1, NSY-1, and PMK-1 (Kim et al., 2002). NSY-1, a MAPK kinase kinase (MAPKKK), phosphorylates and activates SEK-1, a MAPKK, which signals via the MAPK PMK-1 (Kim et al., 2002) (Figure 1.3). This pathway acts cell autonomously in the intestine in response to bacterial pathogens, and in the epidermis in response to fungal pathogens and wounding (Pujol et al., 2008; Shivers et al., 2009). In mammalian studies, a variety of transcription factors were identified as direct targets of PMK-1, many of which are conserved in *C. elegans* (Akira, Uematsu, & Takeuchi, 2006; Karin, 1995). A forward genetic screen identified AFT-7, ortholog of human AFT2, as being an important p38 MAPK transcription factor regulating transcription of innate immune genes (Shivers et al., 2010). AFT-7 functions as a transcriptional repressor until it is phosphorylated by PMK-1, then becoming a transcriptional activator of innate immune genes (Shivers et al., 2010) (Figure 1.3).

In *Drosophila*, Toll-like receptors and Toll-Interluekin-1 Receptor (TIR) domain adaptor proteins function upstream of p38 MAPK cascades (Lindsay & Wasserman, 2014). TIR domain adapter proteins specifically bridge the gap between TLR and MAPK signaling, initiating p38 MAPK pathways. In *C. elegans*, mutations in components of the TLR pathway, including *tol-1*, *trf-1*, and *ikb-1*, cause increased accumulation of *Salmonella enterica* in the pharynx and increased susceptibility to *S. enterica* (Tenor et al., 2008). However, the TLR pathway does not appear to play a role in intestinal pathogenesis (Irazoqui et al., 2010; Pujol et al., 2001). The sole TIR domain protein in *C. elegans*, TIR-1, activates MAPK signaling upstream of PMK-1 (Liberati et al., 2004). The mechanism of TIR-1 activation remains unclear but may be related to diacylglycerol (DAG). DAG generation is catalyzed by phospholipase C and causes activation of

protein kinase C (PKC) (Ren et al., 2009). TPA-1, a PKC in *C. elegans*, is directly phosphorylated by DAG and activates DKF-2, a protein kinase D (Irazoqui et al., 2010b; Ren et al., 2009) (Figure 1.3). Mutations in *dkf-2* result in increased susceptibility to *P. aeruginosa* and *Enterococcus faecalis* (Ren et al., 2009). DKF-2 is required for immune response via p38 MAPK signaling, as overexpression of DKF-2 causes increase in phosphorylated PMK-1 (Ren et al., 2009). The mechanism of activation between DKF-2 and MAPK signaling remains unknown, but DKF-2 may directly phosphorylate TIR-1 (Irazoqui et al., 2010b; Kim & Ewbank, 2015).

Pore-forming toxins (PFTs), produced by many human bacterial pathogens, have also been shown to activate p38 MAPK signaling in several organisms, including *C. elegans*, insects, and mammalian cells (reviewed in Porta et al., 2011). In *C. elegans*, mutations in *pmk-1* and *sek-1* caused increased susceptibility to PFTs produced by *B. thuringiensis* (Huffman et al., 2004).

DAF-2/16 insulin-like pathway

The insulin-like signaling pathway was originally identified in *C. elegans* for its role in lifespan, reproduction, and regulating dauer entry, an alternative life stage that occurs under strenuous environmental conditions (Kenyon et al., 1993; Kimura et al., 1997). Mutations in the sole insulin/IGF-1-like receptor *daf-2* in *C. elegans* leads to an almost doubling of lifespan when exposed to many bacteria, including pathogens such as *P. aeruginosa, E. faecalis*, *Staphylococcus aureus* (Garsin et al., 2003). This effect is dependent on the Forkhead transcription factor DAF-16 (Garsin et al., 2003). Other components of this pathway include the phospho-inositide 3-kinase, AGE-1, which is phosphorylated by DAF-2 resulting in conversion of phosphatidylinositol (4,5)-trisphosphate (PIP_2) to phosphatidylinositol (3,4,5)-trisphosphate (PIP_3), which then recruits kinases PDK-1, AKT-1, AKT-2, and SGK-1. PDK-1 phosphorylates AKT-1, AKT-2, and SGK-1, which then form a complex that phosphorylates DAF-16 (Ewbank,

2006). Phosphorylation of DAF-16 results in its localization to the cytoplasm, which prevents it from entering the nucleus and regulating gene expression (reviewed in Landis & Murphy, 2010) (Figure 1.3). When the pathway is inactivated (e.g. in *daf-2* mutants), DAF-16 localizes to the nucleus which results in upregulation of genes involved in longevity and stress resistance.

Because the DAF-2/16 pathway is involved in both longevity and defense responses, this couples the pathway with a likely role in innate immunity. In fact, many of the genes targeted by DAF-16 have antimicrobial activities, such as lysozymes and detoxification enzymes (McElwee, Bubb, & Thomas, 2003; Murphy et al., 2003). In addition, *daf-2* mutant worms display a decrease in bacterial packing, suggesting that regulation of genes by DAF-16 defends against accumulation of bacteria in the intestine (Murphy et al., 2003). Studies examining the effect of pathogen exposure have intriguingly found that *P. aeruginosa* infection suppresses the activity DAF-16, rather than inducing these defense responses (Evans, Kawli, & Tan, 2008). This suggests that although constitutive expression of this pathway results in increased longevity, it may not play a direct role in pathogen response. In fact, it may even be targeted by pathogens as a virulence mechanism. However, this phenomenon may be pathogen specific, as exposure to *S. marcescens* induces expression of several DAF-16 targets that have putative antimicrobial activity (Mallo et al., 2002).

This pathway is complex in nature with a variety of coregulators and cross-talk between other pathways. For example, the increase in lifespan of *daf-2* mutants is dependent on the p38 MAPK pathway, suggesting it acts in parallel or downstream of DAF-2/16 (Troemel et al., 2006). In addition, many DAF-16 targets contain a GATA motif, termed the DAF-16 associated element (DAE) (Murphy et al., 2003). Two GATA transcription factors, ELT-3, specific to the epidermis, and ELT-2, specific to the intestine, both regulate expression of DAF-16 targets in a

tissue specific manner (Zhang et al., 2013). Mutations in *elt-2* cause increased mortality to *P. aeruginosa*, *S. enterica*, and *E. faecalis* and bacterial distention in the intestine (Kerry, TeKippe, Gaddis, & Aballay, 2006; Shapira et al., 2006). ELT-2 regulates expression of many genes involved in innate immunity and defense response (Block et al., 2015), and seems to be specific to pathogen response, as *elt-2* mutants are not susceptible to oxidative stress, heat stress, or cadmium exposure (Shapira et al., 2006). Conflicting results on whether these ELT-2 or ELT-3 can suppress the longevity of *daf-2* mutants suggests that the interplay of these pathways is complex and condition-specific (Budovskaya et al., 2008; Kerry et al., 2006; Zhang et al., 2013). Lastly, SKN-1, a putative transcription factor involved in stress responses in the intestine, can be phosphorylated by AKT-1, resulting in repression of SKN-1 target gene expression (Tullet et al., 2008) (Figure 1.3). Mutations in *skn-1* in a *daf-2* background suppress the longevity phenotype of *daf-2* mutants, suggesting that SKN-1 contributes to increased lifespan and stress responses (Tullet et al., 2008). To further complicate this response, Block and colleagues suggest a complex interplay between ELT-2, SKN-1, and ATF-7, where a combination of factors is required for expression of particular immune genes (Block et al., 2015) (Figure 1.3). Further analysis of double and triple mutants in response to different pathogens is required to further understand the interaction of these pathways.

Unfolded protein response pathway

The IRE-1-XBP-1 branch of the unfolded protein response (UPR) regulates expression of genes involved in ER homeostasis, leading to defense responses and increased longevity (Zhang & Kaufman, 2004). The UPR^{ER} is conserved in animals as well as some fungi. This pathway involves activation of IRE-1, which leads to the alternative splicing and activation of *xbp-1* mRNA in response to accumulation of unfolded proteins in the endoplasmic reticulum (ER)

(Figure 1.3). In addition to unfolded proteins, activation of IRE-1-XBP-1 occurs in response to pore forming toxins (PFT). Mutants of IRE-1 and XBP-1 lead to hypersensitivity to PFT Cry5 (Bischof et al., 2008). Activation of this pathway in response to PFTs, but not unfolded proteins, is dependent on the p38 MAPK pathway (Bischof et al., 2008). However, the IRE-1-XBP-1 branch of the UPR is also involved in response to bacterial pathogens that do not form PFTs. Intestinal infection of *C. elegans* with *P. aeruginosa* induces expression of the heat shock protein HSP-4, a downstream effector of the IRE-1-XBP-1 pathway (Richardson, Kooistra, & Kim, 2010). Again, this activation is dependent on the p38 MAPK signaling pathway. To understand whether the UPR is involved in counteracting the innate immune response, *pmk-1*; *xbp-1* double mutants were examined. Whereas *xbp-1* single mutants show marked decrease in survival, the double mutant is less detrimental and comparable to the single *pmk-1* mutant (Richardson, Kooistra, & Kim, 2010). This suggests the UPR mitigates cell damage induced by the innate immune response.

In addition to UPR in the ER being involved in innate immunity, another study identified an overlap of upregulated genes in response to mitochondrial stress and infection to *P. aeruginosa* (Pellegrino et al., 2014). The activation of several of these genes was dependent on the mitochondrial UPR transcription factor ATFS-1 (Pellegrino et al., 2014). Unlike the UPR^{ER}, regulation of this pathway is independent of the MAPK pathway (Pellegrino et al., 2014). Therefore, the mitochondrial UPR is also able to protect against pathogens that induce mitochondrial stress by coupling antimicrobial and mitochondria homeostasis gene expression.

TGFβ pathway

The transforming growth factor β (TGFβ) pathway is involved in development and embryogenesis (Roberts et al., 2010; Savage-Dunn, 2005). In mammals, the TGFβ pathway is

required for T cell development and differentiation (reviewed in Letterio, 2005). In *C. elegans*, mutants of *dbl-1*, a TGFB ligand, are prone to infection and are more susceptible to *S. marcescens* (Mallo et al., 2002). In addition, several antimicrobial peptides, including CLEC-85 and LYS-8, are regulated by this pathway (Alper et al., 2007; Roberts et al., 2010). The canonical DBL-1/TGFβ pathway components include SMA-6 and DAF-4, type I and II receptors, respectively, and SMA-2, SMA-3, and SMA-4, Smad signal transducers (Savage-Dunn, 2005) (Figure 1.3). Expression of several AMPs is dependent on SMA-2 (Mochii et al., 1999). Further details on the involvement of other downstream components is largely unknown; however, mutations in *dbl-1, sma-6, sma-2, sma-3,* and *sma-4* in *C. elegans* cause increase susceptibility to *S. maltophilia* (White et al., 2016).

Anti-microbial peptides

Traditionally, anti-microbial peptides have been identified by homology to other AMPs and by expression profiling. Genes that are commonly identified as being differentially expressed in response to bacterial pathogens include caenopores, lysozymes, defensin-like AMPs, and C-type lectin domain proteins (reviewed in Dierking, Yang, & Schulenburg, 2016). The characterization of these proteins is largely based on sequence structure, and further functional characterization of these AMPs is not well studied.

Caenopores, or saposin-like proteins in *C. elegans* share structural similarity with saposin-like proteins (SAPLIPS) in protozoa and mammals (Roeder et al., 2010). Although there are 28 saposin-like protein family (SPP) proteins identified in *C. elegans*, only a few have been identified as immune effectors. *spp-9* and *spp-18* are regulated by DKF-2 and are upregulated by *P. aeruginosa* exposure (Ren et al., 2009). *spp-1* and *spp-12* are regulated by DAF-16, and knockdown of these genes results in decreased lifespan on *E. coli* (Alegado & Tan, 2008).

Further, knockdown of *spp-1* and *spp-5* results in accumulation of bacteria in the intestine (Alegado & Tan, 2008; Roeder et al., 2010). Functional analysis of SPP-5 revealed its ability to form pores in and damage bacterial cell walls (Roeder et al., 2010).

Lysozymes are involved in the hydrolysis of peptidoglycan, a major component of bacterial cell walls. Not surprisingly, lysozymes also play a role in digestion (reviewed in Dierking, Yang, & Schulenburg, 2016). *C. elegans* lysozyme genes are classified into ten protist-type (*lys-1* to *lys-10*), and five invertebrate-type (*ilys-1* to *ilys-5*). This is the largest class of genetically diverse lysozymes found in any organism to date (Schulenburg & Boehnisch, 2008). All lysozymes studied to date are expressed mainly in the intestine (reviewed in Schulenburg & Boehnisch, 2008). Similar to caenopores, many lysozymes are regulated by defense pathways such as TGFβ (*lys-1* and *lys-8*), DAF-2/16 (*lys-7* and *lys-8*), and p38 MAPK (*lys-2*) (Alper et al., 2007; Murphy, McCarroll, et al., 2003; Troemel et al., 2006). Many lysozyme proteins are differentially expressed in response to pathogens (Boehnisch et al., 2011; Dierking, Yang, & Schulenburg 2016; Yang et al., 2015) (Figure 1.4). In fact, *ilys-1* and *lys-9* are the only lysozymes that are not differentially expressed upon bacterial pathogen exposure (Dierking, Yang, & Schulenburg 2016) (Figure 1.4). However, to our knowledge, functional characterization of lysozymes to determine their antimicrobial activities has not been performed.

Defensin-like peptides, termed antibacterial factors ABF-1 to ABF-6 in *C. elegans*, were identified based on sequence homology to proteins of *Ascaris suum*, an intestinal parasitic nematode (Kato et al., 2002). ABF-2 displays in vitro antimicrobial activity and knockdown increases pathogen accumulation (Alegado & Tan, 2008; Kato et al., 2002). Regulation of these genes is not well understood, but the M-box motif-class transcription factor HLH-30 appears to be required for *abf-2* expression in response to *S. aureus* (Visvikis et al., 2014).

C-type lectin domain (CTLD) proteins were originally characterized for their Ca^{+}-dependent carbohydrate binding ability (reviewed in Zelensky & Gready, 2005). However, this superfamily has now grown to include proteins with structural similarity that do not display these functional characteristics. In *C. elegans*, CTLD proteins are the most diverse group of effector molecules, containing 283 members (Pees et al., 2016). Although CLEC-39 and CLEC-49 have been shown to directly bind to *S. marcescens* (Miltsch et al., 2014), CTLD proteins have not been shown to have antimicrobial activities. Evidence of interaction between several CTLD proteins and LYS-7 could suggest downstream signaling or co-regulation of immune partners (Kesika & Balamurugan, 2012). In addition, CTLD proteins exhibit differential expression in response to a variety of pathogens (Alper et al., 2007; Dierking, Yang, & Schulenburg, 2016; Irazoqui et al., 2010a; Troemel et al., 2006; Yang et al., 2015) (Figure 1.4). Mutation or knockdown of several CTLD genes, including *clec-17*, *clec-39*, *clec-49*, *clec-70*, and *clec-86*, results in increased susceptibility to pathogens (Irazoqui et al., 2010a; Miltsch et al., 2014; O'Rourke et al., 2006). Although there is a clear role for CTLD proteins in innate immune response, their functional roles in *C. elegans* are not well known (Pees et al., 2016).

WormExp, a database that contains "-omics" data from many experiments under a variety of conditions, was developed to analyze transcriptional and translational responses across experiments. This database was recently used to compare common effectors in response to a variety of pathogens (Dierking, Yang, & Schulenburg, 2016) (Figure 1.4). Dierking and colleagues found that expression of effectors is taxon-specific, but some members of common classes of AMPs, specifically caenopores and lysozymes, are differentially expressed in response to almost all pathogens. Defensin-like peptides, on the other hand, play a less prominent and

more species-specific role in response to bacterial pathogens (Dierking, Yang, & Schulenburg, 2016).

Therefore, although many common innate immune pathways and effectors have been identified, there are differences in responses to different bacteria and even strains of bacteria. For example, one study comparing responses to the intestinal pathogens *S. marcescens*, *E. faecalis*, and *Photorhabdus luminescens,* found only 11% overlap in differentially expressed genes by RNA sequencing (Engelmann et al., 2011). This phenomenon could be due to species specific responses to different pathogens, or the ability of bacteria to manipulate different host responses. Therefore, it is essential to study a variety of pathogens, their virulence factors, and responses to these pathogens in order to fully understand the complexity of genetic mechanisms underlying pathogen defense.

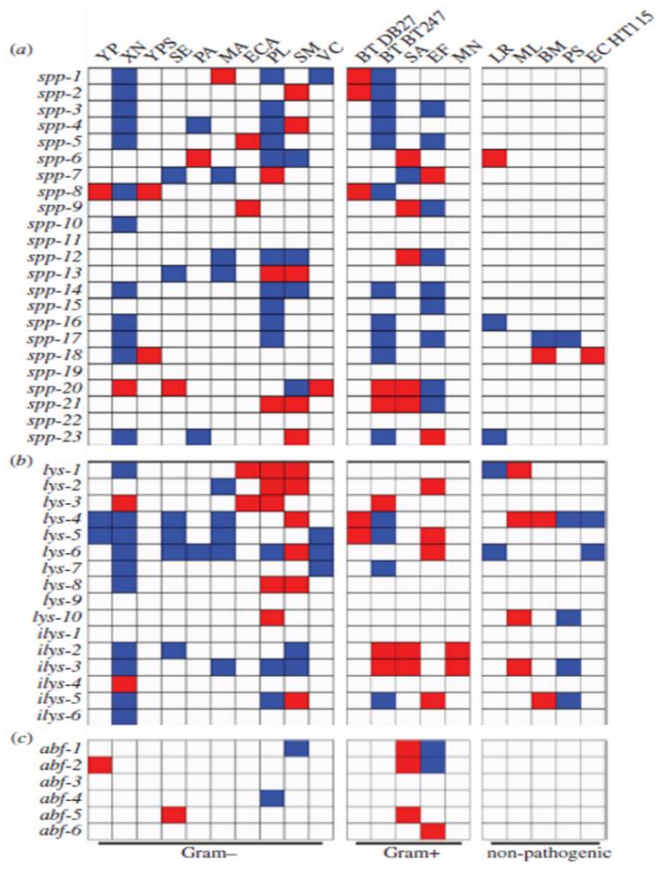

Figure 1.4 Gene expression patterns of caenopores, lysozymes, and defensin-like AMP genes in *C. elegans* upon exposure to bacteria

Red boxes indicate up-regulation and blue boxes represent down-regulation, based on the most responsive gene sets in the WormExp database. YP= *Yersinia pestis*; XN= *Xenorhabdus nematophila*; YPS= *Yersinia pseudotuberculosis*; SE= *Salmonella enterica*; PA= *Pseudomonas aeruginosa*; MA= *Microcystis aeruginosa*; EC= *Erwinia carotovora*; PL= *Photorhabdus luminescens*; SM= *Serratia marcescens*; VC= *Vibrio cholerae*; BT= *Bacillus thuringiensis*; SA= *Staphylococcus aureus*; EF= *Enterococcus faecalis*; MN= *Microbacterium nematophilum*; LR= *Lactobacillus rhamnosus*; ML= *Micrococcus luteus*; BM= *Bacillus megaterium*; PS= *Pseudomonas sp.*; EC= *Escherichia coli*.

Bacteria found in the natural environment of *C. elegans*

Many studies have focused on the interactions between *C. elegans* and bacteria in the context of nutritional value, pathogenesis, and neuronal response (reviewed in Kim, 2013; reviewed in Kim & Ewbank, 2015). However, many of these bacteria, including *S. enterica*, *P. aeruginosa*, *S. aureus*, and *E. faecalis*, were chosen because of their convenience in the lab setting or their implication in human health rather than their native interactions with *C. elegans* in the environment. In order to understand pathogen response and other interactions to individual species, it is necessary to focus efforts on species that realistically interact with *C. elegans*. Previous studies have determined that the microbiome plays a role in lifespan, aging, and disease. In humans, gut microbiome dysbiosis is linked to several diseases, including colorectal cancer, diarrheal diseases, liver diseases, and diabetes (reviewed in Wang et al., 2017). In addition, microbiota of elderly people correlated with health markers such as frailty and nutritional status (Claesson et al., 2012). Some identified commensal species can protect the intestine from pathogenic species by preventing their colonization and reducing inflammatory responses (reviewed in Buffie & Pamer, 2013). In *C. elegans*, worms exposed to lactic acid bacteria, including several species of lactobacilli and bifidobacteria, display increased resistance to *Salmonella enterica,* suggesting a probiotic function of some bacterial species on immunity (Ikeda et al., 2007).

This symbiosis between microbes and animals has led to the hologenome theory, suggesting that microbes and their host act as one unit of selection in evolution (Zilber-Rosenberg & Rosenberg, 2008). This suggests that the microbiome is heritable from generation to generation, but also has the ability to evolve with environmental shifts (Zilber-Rosenberg & Rosenberg, 2008). For example, one study determined that after coevolution of *C. elegans* with a

natural pathogen, *B. thuringiensis*, *C. elegans* became more resistant to killing by *B. thuringiensis* (Schulte et al., 2010). This not only emphasizes the importance of the microbiome, but suggests it is intimately coupled with and causes genetic changes within the host and vice versa.

The emerging evidence of the importance of the microbiome has led to systematic characterization of microbiomes of many organisms (e.g. Cheng et al., 2013; Yun et al, 2014; Ni et al., 2013; Shin et al., 2016). For example, several groups recently determined the microbial repertoire associated with *C. elegans* (Berg et al., 2016; Dirksen et al., 2016; Samuel et al., 2016). All three studies characterized the microbial communities found in substrates, such as rotting plant matter, whereas two of the three studies (Dirksen et al., 2016; Berg et al., 2016) also characterized the internal microbiome of *C. elegans*. In all studies, microbial isolates were obtained from substrates or worms, and the 16S rRNA gene was amplified and sequenced to identify bacterial composition using operational taxonomic units (OTUs). When assessing microbiome communities, worms were washed several times before DNA was isolated to avoid contamination of sequences from substrates or the nematode cuticle. Overall, thousands of OTUs were identified in all substrates and worms, indicating an extremely diverse environment and microbiome.

Microcosm experiment to simulate diverse soil environments

Berg and colleagues created artificial microcosms to simulate natural environments using soil supplemented with various produce, including plants and fruit (Berg et al., 2016). Wild-type worms were subjected to one of three different soil types for three days before sequencing, and DNA was isolated from both worms and soils. Bacterial composition was compared between worms and soils, with all worm-associated bacterial communities, regardless of substrate, more

similar to each other than to their respective substrate (Figure 1.5A). Overall, 4,445 OTUs were identified in the soils, with only 60% (2,656 OTUs) also being represented in the worms. This trend is suggestive of selective pressure, either by competition of bacteria within the intestine or selection of bacteria by the worm. In addition, a core microbiome was identified that contained taxa that were shared among all worm samples. This shared microbiome included the families Burkholderiaceae, Bacillacea, Enterobacteriaceae, Aeromanadaceae, Alcaligeaceae, Pseudomonadaceae, Phizobiaceae, Sphingobacteriaceae, and Xanthomonadaceae. Interestingly, all members of the core microbiome were found in higher relative abundance in the microbiome than within the soils, suggesting preferential inclusion of specific genera within the microbiome (Berg et al., 2016). To further support this, Checkerboard scores, or C-scores, were used to determine whether the assembly of the microbiome was random. C-scores are calculated using an incident matrix based on instances of mutual exclusion of bacterial species. This analysis resulted in a non-normal pattern, indicating that there is some driving force for microbiome assembly in *C. elegans*. Interaction networks between bacterial families revealed a large number of negative interactions between bacteria, indicating that competition may be one such mechanism for microbiome assembly (Berg et al., 2016).

Characterization of bacteria found in the natural habitat of *C. elegans*

Samuel and colleagues identified the bacterial community associated with substrates where *C. elegans* were collected (Samuel et al., 2016). Substrates, including rotting fruit and stems, and snails, a common vector for *C. elegans* dispersal, were sampled in several locations in France and Spain. In addition to the isolation of bacterial DNA from substrates, nematodes and bacteria were also isolated and characterized from each substrate. In total, ~2,400 OTUs were identified from the substrates. Interestingly, this number is lower than total soil diversity and

more similar to specialized niches such as rhizospheres (Lundberg et al., 2012). Relative abundance of genera between substrate types were also generally comparable. Common phyla within substrates include Proteobacteria, Bacteroidetes, Firmicutes, and Actinobacteria.

Individual bacterial species were isolated and further characterized for their effect on the overall health of *C. elegans*. To determine whether individual species are detrimental or beneficial to *C. elegans*, worms were grown on each of 565 bacterial species, and stress and immunity reporters were used to look for induction of pathogen responses. Using these data, each isolate was characterized as either beneficial, intermediate, or detrimental. Of all 565 isolates, approximately 40% were beneficial, 40% were intermediate, and 20% were detrimental. In general, many Proteobacteria, as well as species within the genera *Lactococcus* were beneficial, whereas Bacteriodetes and Gammaproteobacteria tended to be more detrimental. Beneficial and detrimental bacteria were then mixed to determine whether the detrimental isolates inhibit growth and induce stress responses because of lack of nutritional value or because of their antagonistic effect on *C. elegans*. In several cases tested, the addition of *E. coli* OP50 to the detrimental isolate did not neutralize the effect of the detrimental strain, indicating that these strains have pathogenic effects on *C. elegans* rather than being poor nutrition sources. This phenomenon can also be seen in nature, whereby proliferating *C. elegans* are found more commonly on substrates that are enriched in more beneficial Alphaproteobacteria, but lack more detrimental Gammaproteobacteria, suggesting that the inclusion of beneficial bacteria does not outweigh the effects of the more detrimental species (Samuel et al., 2016).

Identification of the native microbiome of *C. elegans*

Dirksen and colleagues combined the approaches of the previous studies by analyzing the microbial communities in the native habitats of *Caenorhabditis* species as well as the

microbiome of sampled nematodes from those habitats (Dirksen et al., 2016). Nematodes were sampled from locations in France and Portugal, and nematodes and respective substrates were collected from several locations in Germany. Nematode species collected include *C. elegans*, *Caenorhabditis remanei*, and *Caenorhabditis briggsae*. Microbial communities were sequenced from all substrates and worms either directly after sampling, or after growing on *E. coli* OP50 for several weeks ("lab enriched" worms). The lab-enriched worm microbiome was utilized to provide insight into bacterial species that persist in the worm and do not simply pass through the intestine. In addition, individual bacteria species were isolated from substrates for further analysis and development of an experimental microbiome. Similar to the microcosm experiments, microbial communities associated with nematodes, regardless of species, were more similar to each other than to microbial communities of substrates (Figure 1.5A). In addition, multivariate analyses resulted in clustering of sample types ("lab enriched" worms vs. natural worms vs. substrates) and nematode species (with *C. elegans* and *C. briggsae* differing from *C. remanei*), suggesting a nematode-specific microbiome as well as a species-specific microbiome. The most abundant OTUs in *C. elegans* microbiome included members of the genera *Pseudomonas*, *Stenotrophomonas*, *Ochrobactrum*, and *Sphingomonas*, as well as unclassified Enterobacteriaceae (Dirksen et al., 2016).

An experimental microbiome consisting of 14 isolates was developed using the commonly identified taxa of the microbiome and isolated bacteria from the substrates (Dirksen et al., 2016). It was determined that the abundance of several of these bacterial species increases within the microbiome as compared to the bacterial lawn, suggesting preference for particular species. In addition, when *C. elegans* were raised on plates of single bacterial isolates and then transferred to plates without bacteria, several species were still detected in the intestine after 24

hours. This finding, along with the identification of a bacterial community in the "lab enriched" worms, suggests that some species are able to persist in, and colonize, the intestine. The experimental microbiome as well as individual isolates were then tested for their effect on population size, a measure of overall health of *C. elegans*. Whereas approximately 70% of individual isolates tested increased population growth, approximately 30% decreased population growth. The species that negatively affected population size belong to Gammaproteobacteria, Actinobacteridae, Bacilli, Flavobacteria, and Sphingobacteria, consistent with the detrimental species identified by Samuel and colleagues (Samuel et al., 2016). The experimental microbiome increased population growth when compared to *E. coli* OP50 under almost all conditions, including changes in plate salt concentration and temperature. This result conflicts with observations from Samuel et al., where many beneficial species were not able to rescue the effects of the detrimental species (Samuel et al., 2016). However, a study that grew *C. elegans* on subsets of natural isolates found that the intrinsic growth rate of *C. elegans* grown on a mixture of species was most similar to the intrinsic growth rate of the most beneficial species when grown on each individually (Darby & Herman, 2014). These conflicting results suggest that this phenomenon is dependent on the species utilized and possibly the method of determining health. Therefore, sometimes beneficial species are able to outweigh the effects of detrimental species and sometimes they are not.

Determination of a common microbiome

A meta-analysis of these studies identified overall patterns when combining data from all experiments (Zhang et al., 2017). This analysis showed again that worm microbiomes, regardless of location or substrate, are different from, and less diverse than, that of the microbial repertoire associated with the substrates (Figure 1.5A). There are several possible explanations for this

phenomenon: some bacterial species are able to pass through the pharynx live and accumulate within the intestine of the worm, resulting in higher abundance, there is competition between species within the intestine, resulting in some species increasing in abundance and some decreasing in abundance, or *C. elegans* are able to recognize and preferentially choose bacteria to consume and avoid, resulting in the ability of only some species to enter the intestine. Likely it is a combination of the above factors, as there is previous evidence for both competition (Berg et al., 2016) and preference for some bacterial species over others (Abada et al., 2009; Zhang, 2008). Further studies, including characterization of the microbiome over time, are needed to completely understand these dynamics.

The robust signature of the microbiomes of *C. elegans* allowed for the identification of common microbes across studies, termed the core microbiome (Figure 1.5B, C). Overall, 260 OTUs were identified in all *C. elegans*' microbiomes, with the most abundant taxa including Enterobacteriaceae, Pseudomonadaceae, and Xanthomonadaceae (Zhang et al., 2017). Other taxa, such as Acetobacteriaceae, Actinobacteria, Moraxellaceae, and Comamonadaceae were found in lower abundance but still common between experiments, suggesting a key role for these species as well. In total, fourteen bacterial families are found in all natural worm microbiotas, with a majority of these, although not all, also being found in microbiomes of the lab enriched and microcosm *C. elegans* (Zhang et al., 2017) (Figure 1.5C). The identification of a common, reproducible microbiome allows for use of this core microbiome for future studies. In addition, focusing efforts on interactions of *C. elegans* with individual species that are encountered in their natural environment ensures that responses to these bacteria are realistic.

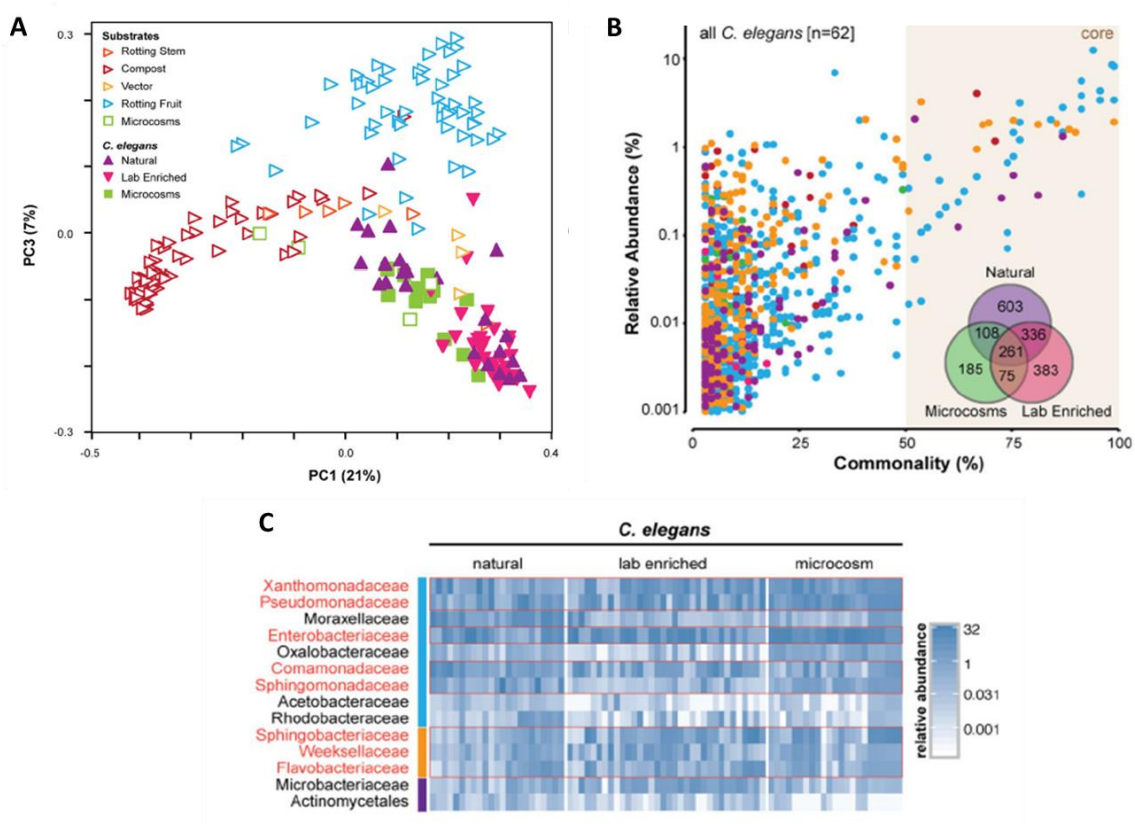

Figure 1.5 Characterizing the core microbiome of *C. elegans*

Adapted from Zhang et al, 2017. **A** Principle coordinate analysis distinguishes bacterial communities found in *C. elegans* (filled) from those found in substrates (open) regardless of study. **B** Scatterplot of OUT-level abundance and commonality across all *C. elegans* microbiomes. Venn diagram compares OTUs from each type of microbiome. **C** Heatmap showing abundance of 14 bacterial families found in all natural worm microbiomes across sample. Red font indicates families that are also present in lab-enriched and microcosm microbiomes. Colored dots **B** and colors to the left of the heatmap **C** represent different bacterial phyla.

Natural *C. elegans* pathogens

In the past, many bacterial pathogens used to study innate immunity in *C. elegans* were chosen because of their medical relevance. However, with the systematic identification of natural pathogens associated with *C. elegans*, we are now able to focus efforts on bacterial pathogens that have medical implications but are also realistic pathogens to *C. elegans*. Many species identified as part of the worms' microbiome have previously been studied in *C. elegans* and do in fact have implications on human health, including species of *Pseudomonas* and *Stenotrophomonas* (Berg et al., 2016; Brooke, 2012; Dirksen et al., 2016; Streit et al., 2004). In addition, some common detrimental taxa found in the core microbiome of *C. elegans* may be of ecological importance because of their increase in abundance in the microbiome as compared to substrates, such as *Ochrobactrum* and *Chryseobacterium* (Dirksen et al., 2016).

Within *Ochrobactrum*, *Ochrobactrum anthropi* is the most widely studied for its occurrence as a human pathogen (reviewed in Ozdemir, Soypacacı, Sahin, Bicik, & Sencan, 2006). Although *Ochrobactrum* infection in humans is rare, this bacterium is multi-drug resistant and able to cause infection in patients via medical devices, suggesting a potential to become an emerging opportunistic pathogen (Ozdemir et al., 2006). Isolation of *O. anthropi* from a patient with Crohn's disease and short-bowel syndrome suggests that this species may be a part of our normal intestinal flora (Alnor et al., 1994). In addition to isolation of *Ochrobactrum* from human patients, it has also been isolated in soils and associated with wheat roots (Lebuhn et al., 2000).

Chryseobacterium indologenes causes rare infections in humans, including bacteremia, wound sepsis, cellulitis, biliary tract infection, and pneumonia (reviewed in Lin et al., 2010). *Chryseobacterium* species can be found in water, soil, and plants (Chaudhari et al., 2009; Gudeta et al., 2016; Nishioka et al., 2016). *Chryseobacterium* is commonly isolated from marine animals

such as trout, shrimp, and seals, posing a threat to consumers of these animals (Maravic et al., 2013).

Interestingly, most species of *Pseudomonas* isolated in association with *C. elegans* were not detrimental to their health (Dirksen et al., 2016; Samuel et al., 2016). However, they are discussed as a pathogen within this section because of the plethora of research using *C. elegans* and *P. aeruginosa* to understand pathogen-host interactions. Two mechanisms of pathogenesis by *P. aeruginosa* PA14 have been identified in *C. elegans*. The first, termed "slow killing," requires live bacterial cells and accumulation of cells in the intestine followed by infection-like symptoms such as distended intestine and reduced mortality (Kirienko et al., 2014; Tan et al., 1999). "Fast killing," on the other hand, occurs on high osmolarity media and does not rely on live bacterial cells but instead on secreted toxins (Kirienko et al., 2014; Tan, Mahajan-Miklos, & Ausubel, 1999). In addition to the detrimental effects of *P. aeruginosa* PA14 exposure to *C. elegans*, this species is also a common cause of nosocomial infections. In fact, *P. aeruginosa* is the most common Gram-negative species to cause nosocomial pneumonias and constituted 12% of all isolates obtained from ICU patients with infection in 2001, ranking second in total abundance after *S. aureus* (Hidron et al., 2008; Streit et al., 2004).

Although not highly virulent, *S. maltophilia* is capable of causing infection in immunocompromised patients, most commonly identified in patients with cystic fibrosis and lung cancer (Brooke, 2012). *Stenotrophomonas* species isolated in association with *C. elegans* show varying pathogenicity to *C. elegans* (Dirksen et al., 2016; Samuel et al., 2016). This is also true of other environmental and clinical isolates of *S. maltophilia*, with only some strains being pathogenic (Pompilio et al., 2011; White et al., 2016; Adamek et al, 2011). In *C. elegans*, bacterial load of virulent strains of *S. maltophilia* is greater than that of *E. coli* OP50, and it

appears that live bacteria are required for virulence, suggesting that toxins are not an employed virulence mechanism (White et al., 2016).

Common characteristics of *Pseudomonas*, *Stenotrophomonas*, *Ochrobactrum*, and *Chryseobacterium* include being Gram-negative, multi-drug resistant (MDR) bacteria that cause infection within tissues of humans. Because of their identification within the microbiome, their mode of pathogenesis occurs mainly within the intestine of *C. elegans*. In general, aging in *C. elegans* correlates with an increase of accumulation of bacteria in the intestine and decrease in tissue integrity (Garigan et al., 2002). As *C. elegans* age, their defense responses, particularly p38 MAPK signaling, decrease, which causes accumulation of live bacteria, resulting in intestine epithelial damage and eventual deterioration (Kim, 2013) (Figure 1.2). Therefore, pathogenic bacteria must be able to avoid pharyngeal grinding and proliferate in the intestine. However, accumulation of *E. coli* alone in the intestine does not shorten lifespan (Virk et al., 2016), so other virulence mechanisms must also be involved in the detrimental nature of intestinal pathogenesis.

Common virulence mechanisms

Horizontal gene transfer and the short generation time of bacteria enable rapid evolution of virulence mechanisms that allow for adaptation to evolving host responses and hosts themselves (Juhas et al., 2009; Smith, 2001). Regions known as genomic islands contain gene-rich regions that are usually attributed with past or present mobility, representing regions of variability between closely related organisms (Juhas et al., 2009). In *S. maltophilia* and *P. aeruginosa*, it appears that most vital virulence factors are contained within the core genome, and non-essential virulence factors are found within strain-specific, mobile elements such as genomic islands and plasmids (Lira, Berg, & Martínez, 2017; Mathee et al., 2008; Adamek et al., 2014).

Identification of virulence mechanisms within many isolates of bacteria will provide insight into both strain specific and common virulence mechanisms, and with the convenience of sequencing technologies, the abundance of bacterial genomic data has begun to provide this information.

Gram-negative bacteria share many common virulence factors, including secreted enzymes, adhesion structures, secretion systems, antibiotic resistance genes, and quorum sensing mechanisms. Table 1.1 summarizes known virulence factors found in each genus introduced above. Structures involved in adhesion include fimbriae, pili, flagella, and capsules. Pili are structures that are used for genetic transfer, and pili, fimbriae, and flagella allow for movement of bacterial cells. All of these structures also aid in biofilm formation, which allows bacteria to build up on medical devices and avoid phagocytosis. Although not typical of this genus, two species of *Chryseobacterium*, *C. bovis* and *C. oranimense*, are capsule-forming and/or contain capsule biosynthesis gene clusters (Laviad-Shitrit et al., 2017; Sharma et al., 2015).

Secretion systems are needed for release of enzymes involved in virulence. Secretion systems in Gram-negative bacteria are classified as Type I secretion systems (T1SS) to Type VI secretion systems (T6SS). Classification is based on proteins that are capable of being transported and how many membranes the system traverses, with some systems being able to transport proteins directly into the host cell (Green & Mecsas, 2016). *P. aeruginosa* contains genes for Type I, II, III, and V, and VI secretion systems (Balasubramanian et al., 2013). However, mutations in genes involved in secretion systems or their effectors in *P. aeruginosa* PA14 did not have a strong phenotypic effect on *C. elegans*, suggesting redundancy between secretion systems and effectors (Feinbaum et al., 2012). In *S. maltophilia*, type I, II, and V secretion systems are common in many strains, whereas type IV, V, and VI secretion systems are strain-specific (Adamek, Linke, & Schwartz, 2014).

Antibiotic resistance genes (ARGs) include efflux pumps and antibiotic-modification enzymes. Efflux pumps enable release of compounds from inside bacterial cells, and several groups of efflux pumps are able to transport antibiotics. These multi-drug resistant (MDR) efflux pumps are classified based on the number of components that the pump has, the energy source needed to activate the pump, and the substrates the pump transports (Piddock, 2006). Common MDR efflux pumps include the ATP-binding cassette (ABC) superfamily, the major facilitator superfamily (MFS), the multidrug and toxic-compound extrusion (MATE) family, and the resistant nodulation division (RND) family (Piddock, 2006). *S. maltophilia* K279a contains nine RND-type efflux pump genes, six of which have been functionally characterized (Crossman et al., 2008; Chang et al., 2015). Specifically, efflux pump SmeDEF is involved in resistance to trimethoprim and sulfamethoxazole (Sánchez & Martínez, 2015). Although resistant to multiple antibiotics, the only reported efflux pump in *Ochrobactrum* species is an arsenic efflux pump in the genome of *O. tritici* (Sousa et al., 2015). β-lactamase is an enzyme that cleaves the β-lactam ring of antibiotics in the penicillin family. Comparison of ARGs in two *Ochrobactrum* strains revealed a role for a β-lactamase gene as well as a florfenicol resistance gene (*FloR*) in antibiotic resistance (Johnning et al., 2013). Lastly, reduced expressed of the outer membrane porin *OprD* in *P. aeruginosa* confers resistance to carbapenems due to decreased ability of these antibiotics to enter bacterial cells (Farra et al., 2008; Pai et al., 2001).

Quorum-sensing allows for signaling between bacterial cells and synchronization of behavior and gene expression (reviewed in Stevens, Schuster, & Rumbaugh, 2012). This leads to effective pathogenesis due to increased expression of secreted proteins and formation of biofilms. Quorum signaling is dependent on the production of autoinducers, which are secreted from cells and then diffuse into neighboring cells to coordinate responses (reviewed in Stevens et

al., 2012). In *S. maltophilia* the autoinducer is Diffusible Signal Factor QS. The synthesis of QS is dependent on the *rpfF* gene, which can be found in two variants in *S. maltophilia* strains (Fouhy et al., 2007; Huedo et al., 2014). *P. aeruginosa* has a quorum-sensing mechanism containing three systems, the LasR/I and RhlR/I systems, which are mediated by N-acyl-homoserine lactones (AHLs), and the PQS system, mediated by 2-alkyl-4 quinolones (AQs) (reviewed in Balasubramanian et al., 2013). These systems function in a hierarchical manner, where the LasR/I system positively regulates the RhlR/I system, and PQS is the downstream signaling mechanism (Dietrich et al., 2006). There is no evidence of quorum sensing signaling in *Ochrobactrum* and *Chryseobacterium*, but strains within these genera have been shown to produce proteins that inactivate AHLs and interfere with quorum sensing of other microorganisms (Mei et al., 2010; Rashid et al., 2011).

Secreted enzymes and secondary metabolites, such as toxins, proteases, DNases, and lipases, function by being released into the extracellular space of the host and damaging functional or structural components of the host. This is a very diverse class of proteins, and every bacterial species contains many secreted proteins. However, whether these proteins play a role in pathogenicity is strain- and condition-specific (Feinbaum et al., 2012; Huang, Lempicki, & Sherman, 2009). In *S. maltophilia*, proteases from two strains harbored nematocidal activity (Huang, Lempicki, et al., 2009; Jankiewicz, Larkowska, & Brzezinska, 2016). Toxins do not seem to play a significant role in *S. maltophilia* pathogenicity, as they are found in a strain-specific manner, and heat-killed *S. maltophilia* are not pathogenic (Adamek et al., 2014; White, 2016). However, most secreted proteins in *P. aeruginosa* that aid in virulence are toxins, such as LasA, LasB, PrpL, ToxA (reviewed in Balasubramanian et al., 2013). Systematic and functional analysis of secreted proteins in *Ochrobactrum* and *Chryseobacterium* is much less studied, but

evidence for 451 potentially secreted proteins, including a Clp protease, were identified in one *O. anthropi* strain (Wang et al., 2016; Wang et al., 2015).

In sum, the variety of virulence mechanisms, the rapid evolution of bacteria, and the specificity of virulence factors between strains complicates the ability to combat bacterial infections. However, gaining insight into genetic and functional mechanisms of virulence factors as well as identifying novel virulence factors provides the possibility of targeting specific bacterial pathogens.

Table 1.1 Overview of virulence factors present in *S. maltophilia, P. aeruginosa, Ochrobactrum,* and *Chryseobacterium* based on functional analyses and sequencing.

Virulence factor	*S. maltophilia*	*P. aeruginosa*	*Ochrobactrum*	*Chryseobacterium*	References
Secretion systems	Type I, II, V- common Type IV, V, VI-strain specific	Type I, II, III V, VI	Type I and IV		Adamek et al., 2014 Balasubramanian et al., 2013 Barquero-Calvo et al., 2009 Chain et al., 2011 Chudasama & Thaker, 2017
Adhesion proteins	Pili Fimbriae Flagella	Capsule Type IV pilus Flagella		Capsule	Adamek et al., 2014 Feinbaum et al. 2012 Sharma et al., 2015 Laviad-Shitrit et al., 2017
Antibiotic Resistant Genes	ABC-type efflux pumps MFS-type efflux pumps RND-type efflux pumps β-lactamase	MFS-type efflux pumps SMR-type efflux pumps MATE-type efflux pumps RND-type efflux pumps ABC-type efflux pumps β-lactamase *OprD*	β-lactamase *FloR*	MFS efflux pump RND-type efflux pump β-lactamase	Crossman, 2008 Gould et al., 2006 Farra et al., 2008 Pai et al, 2001 Wang et al., 2013 He et al., 2004 Johnning et al., 2013 Higgins et al., 2001 Sharma et al., 2015
Quorum-sensing	*rpf*/DSF system	LasR/I, RhlR/I, PQS systems			Lira et al., 2017 Balasubramanian et al., 2013

Conclusions

Using conservation of genes and proteins between organisms, we are able to hypothesize protein function. RNA sequencing is another systematic approach that enables prediction of gene function based on differential expression under specific conditions. Functional validation, a more time- and resource-consuming technique, has confirmed some of these protein functions in both *C. elegans* and bacteria. However, in *C. elegans*, 40% of gene functions are still unknown (Petersen, Dirksen, & Schulenburg, 2015). One reason for this is that the use of artificial lab settings to study *C. elegans* limits responses to environmental stimuli. *C. elegans* and bacteria engage in intimate relationships in their natural environment, so utilizing these interactions to understand gene function in *C. elegans* has already elucidated novel genetic mechanisms (reviewed in Petersen, Dirksen, & Schulenburg, 2015). Because studies of *C. elegans* ecological interactions are only relatively recent, there is still much more to be uncovered.

Identification of organisms in the natural habitat of *C. elegans* through simple isolation has encouraged use of these species in further studies (White et al., 2016; Schulte et al., 2010). Recent high-throughput identification of the microbial communities within *C. elegans* as well as in the surrounding environment has expanded our understanding of natural *C. elegans*-bacterial interactions (Dirksen et al., 2016; Berg et al., 2016; Samuel et al., 2016). Of particular interest are bacterial species that are detrimental to the health of *C. elegans*, not only for their ecological impact, but also because of their medical relevance. Many genera found in the microbiome of *C. elegans* are also human pathogens. The conservation of innate immune and defense pathways and effectors enables hypothesis to be made about how these pathogens cause infection and how these infections are combatted. Studying these interactions with naturally encountered pathogens

will provide insight into not only how *C. elegans* respond to pathogen infection, but also how these bacteria are able to induce infection.

References

Abada, E. A., Sung, H., Dwivedi, M., Park, B., Lee, S., & Ahnn, J. (2009). *C. elegans* behavior of preference choice on bacterial food. *Molecules and Cells, 28*, 209–213. https://doi.org/10.1007/s10059-009-0124-x

Adamek, M., Linke, B., & Schwartz, T. (2014). Virulence genes in clinical and environmental *Stenotrophomonas maltophilia* isolates: A genome sequencing and gene expression approach. *Microbial Pathogenesis, 67–68*, 20–30. https://doi.org/10.1016/j.micpath.2014.02.001

Akira, S., Uematsu, S., & Takeuchi, O. (2006). Pathogen recognition and innate immunity. *Cell, 124*(4), 783–801. https://doi.org/10.1016/j.cell.2006.02.015

Alegado, R. A., & Tan, M. (2008). Resistance to antimicrobial peptides contributes to persistence of Salmonella typhimurium in the *C. elegans* intestine. *Cellular Microbiology, 10*(6), 1259–1273. https://doi.org/10.1111/j.1462-5822.2008.01124.x

Alnor, D., Frimodt-Moller, N., Espersen, F., & Frederiksen, W. (1994). Infections with the unusual human pathogens *Agrobacterium* species and *Ochrobactrum anthropi*. *Clinical Infectious Diseases, 18*(6), 914–920.

Alper, S., Mcbride, S. J., Lackford, B., Freedman, J. H., & Schwartz, D. A. (2007). Specificity and complexity of the *Caenorhabditis elegans* innate immune response. *Molecular and Cellular Biology, 27*(15), 5544–5553. https://doi.org/10.1128/MCB.02070-06

Balasubramanian, D., Schneper, L., Kumari, H., & Mathee, K. (2013). A dynamic and intricate regulatory network determines *Pseudomonas aeruginosa* virulence. *Nucleic Acids Research, 41*(1), 1–20. https://doi.org/10.1093/nar/gks1039

Barquero-Calvo, E., Conde-Alverez, R., Chacon-Diaz, C., Quesada-Lobo, L., Martirosyan, A., Guzman-Verri, C., Iriarte, M., Mancek-Keber, M., Jerala, R., Gorvel, P.J., Moriyon, I., Moreno, E., & Chaves-Olarte, E. (2009). The differential interaction of *Brucella* and *Ochrobactrum* with innate immunity reveals traits related to the evolution of stealthy pathogens. *PLoS ONE, 4*(6), 1–17. https://doi.org/10.1371/journal.pone.0005893

Barrière, A., & Félix, M.-A. (2007). Temporal dynamics and linkage disequilibrium in natural *Caenorhabditis elegans* populations. *Genetics, 176*(2), 999–1011. https://doi.org/10.1534/genetics.106.067223

Barrière, A., & Félix, M.-A. (2005). High local genetic diversity and low outcrossing rate in *Caenorhabditis elegans* natural populations. *Current Biology, 15*, 1176–1184. https://doi.org/10.1016/j.cub.2005.06.022

Berg, M., Stenuit, B., Ho, J., Wang, A., Parke, C., Knight, M., Alvarez-Cohen, L., & Shapira, M. (2016). Assembly of the *Caenorhabditis elegans* gut microbiota from diverse soil

microbial environments. *The ISME Journal, 10*, 1998–2009. https://doi.org/10.1038/ismej.2015.253

Bischof, L. J., Kao, C., Los, F. C. O., Gonzalez, M. R., Shen, Z., Briggs, S. P., van der Goot. F.G., & Aroian, R. V. (2008). Activation of the unfolded protein response is required for defenses against bacterial pore-forming toxin *in vivo. PLoS Pathogens, 4*(10), e1000176. https://doi.org/10.1371/journal.ppat.1000176

Block, D. H. S., Twumasi-boateng, K., Kang, H. S., Carlisle, J. A., Hanganu, A., Lai, T. Y.-J., & Shapira, M. (2015). The developmental intestinal regulator ELT-2 controls p38-dependent immune responses in adult *C. elegans. PLoS Genetics, 11*(5), e1005265. https://doi.org/10.1371/journal.pgen.1005265

Boehnisch, C., Wong, D., Habig, M., Isermann, K., Michiels, N. K., Roeder, T., May, R. C., & Schulenburg, H. (2011). Protist-type lysozymes of the nematode *Caenorhabditis elegans* contribute to resistance against pathogenic *Bacillus thuringiensis. PLoS ONE, 6*(9), e24619. https://doi.org/10.1371/journal.pone.0024619

Brandt, J. P., & Ringstad, N. (2015). Toll-like receptor signaling promotes the development and function of sensory neurons required for a *C. elegans* pathogen-avoidance behavior. *Current Biology, 25*(17), 2228–2237. https://doi.org/10.1016/j.cub.2015.07.037.

Brooke, J. S. (2012). *Stenotrophomonas maltophilia*: An emerging global opportunistic pathogen. *Clinical Microbiology Reviews, 25*(1), 2–41. https://doi.org/10.1128/CMR.00019-11

Budovskaya, Y. V, Wu, K., Southworth, L. K., Jiang, M., Tedesco, P., Johnson, T. E., & Kim, S. K. (2008). An *elt-3/elt-5/elt-6* GATA transcription circuit guides aging in *C. elegans*. *Cell, 134*(2), 291–303. https://doi.org/10.1016/j.cell.2008.05.044

Buffie, C. G., & Pamer, E. G. (2013). Microbiota-mediated colonization resistance against intestinal pathogens. *Nature Reviews Immunology, 13*, 790–801. https://doi.org/10.1038/nri3535

Cabreiro, F., & Gems, D. (2013). Worms need microbes too: microbiota, health and aging in *Caenorhabditis elegans. EMBO Molecular Medicine, 5*, 1300–1310. https://doi.org/10.1002/emmm.201100972

Cargnello, M., & Roux, P. P. (2011). Activation and function of the MAPKs and their Substrates, the MAPK-activated protein kinases. *Microbiology and Molecular Biology Reviews, 75*(1), 50–83. https://doi.org/10.1128/MMBR.00031-10

Chain, P. S. G., Lang, D. M., Comerci, D. J., Malfatti, S. A., Vergez, L. M., Shin, M., Ugalde, R.A., Garcia, E., & Tolmasky, M. E. (2011). Genome of *Ochrobactrum anthropi* ATCC 49188, a versatile opportunistic pathogen and symbiont of several eukaryotic hosts. *Journal of Bacteriology, 193*(16), 4274–4275. https://doi.org/10.1128/JB.05335-11

Chang, Y.-T., Lin, C.-Y., Chen, Y.-H., & Hsueh, P.-R. (2015). Update on infections caused by *Stenotrophomonas maltophilia* with particular attention to resistance mechanisms and therapeutic options. *Frontiers in Microbiology, 6*(893), 1–20. https://doi.org/10.3389/fmicb.2015.00893

Cheng, X-Y., Tian, X-L., Wang, Y-S., Lin, R-M., Mao, Z-C., Chen, N., & Xie, B-Y. (2013). Metagenomic analysis of the pinewood nematode microbiome reveals a symbiotic relationship critical for xenobiotics degradation. *Scientific Reports, 3*(1869). https://doi.org/10.1038/srep01869

Chaudhari, P. N., Wani, K. S., Chaudhari, B. L., & Chincholkar, S. B. (2009). Characteristics of sulfobacin A from a soil isolate *Chryseobacterium gleum*. *Applied Biochemistry and Biotechnology, 158*, 231–241. https://doi.org/10.1007/s12010-008-8417-7

Chilton, S. N., Enos, M. K., Burton, J. P., & Reid, G. (2015). The effects of diet and the microbiome on reproduction and longevity: A comparative review across 5 continents. *Journal of Nutrition & Food Sciences, 5*(3). https://doi.org/10.4172/2155-9600.1000364

Chudasama, K. S., & Thaker, V. S. (2017). Genomics data genome sequence of *Ochrobactrum anthropi* strain SUBG007, a plant pathogen and potential xenobiotic compounds degradation bacterium. *Genomics Data, 11*, 116–117. https://doi.org/10.1016/j.gdata.2017.01.001

Claesson, M. J., Jeffery, I. B., Conde, S., Power, S. E., O'Connor, E. M., Cusack, S., Harris, H. M. B., Coakley, M., Lakshminarayanan, B., O'Sullivan, O., Fitzgerald, G. F., Deane, J., O'Connor, M., Harnedy, N., O'Connor, K., O'Mahony, D., van Sinderen, D., Wallace, M., Brennan, L., Stanton, C., Marchesi, J. R., Fitzgerald, A. P., Shanahan, F., Hill, C., Ross, R. P., & O'Toole, P. W. (2012). Gut microbiota composition correlates with diet and health in the elderly. *Nature, 488*, 178–185. https://doi.org/10.1038/nature11319

Crossman, L. C., Gould, V. C., Dow, J. M., Vernikos, G. S., Okazaki, A., Sebaihia, M., Saunders, D., Arrowsmith, C., Carver, T., Peters, N., Adlem, E., Kerhornou, A., Lord, A., Murphy, L., Seeger, K., Squares, R., Rutter, S., Quail, M. A., Rajandream, A., Harris, D., Churcher, C., Bentley, S. D., Parkhill, J., Thomson, N. R., & Avison, M. B. (2008). The complete genome, comparative and functional analysis of *Stenotrophomonas maltophilia* reveals an organism heavily shielded. *Genome Biology*, *9*(4), R74.1-R74.13. https://doi.org/10.1186/gb-2008-9-4-r74

Darby, B. J., & Herman, M. A. (2014). Effect of prey richness on a consumer's intrinsic growth rate. *Oecologia, 175*, 243–250. https://doi.org/10.1007/s00442-014-2883-0

Dierking, K., Yang, W., & Schulenburg, H. (2016). Antimicrobial effectors in the nematode *Caenorhabditis elegans*: An outgroup to the Arthropoda. *Philosophical Transactions of the Royal Society B, 371*(1695), 20150299. https://doi.org/10.1098/rstb.2015.0299

Dietrich, L. E. P., Price-whelan, A., Petersen, A., Whiteley, M., & Newman, D. K. (2006). The phenazine pyocyanin is a terminal signaling factor in the quorum sensing network of

Pseudomonas aeruginosa. *Molecular Microbiology*, *61*(5), 1308–1321. https://doi.org/10.1111/j.1365-2958.2006.05306.x

Dirksen, P., Marsh, S. A., Braker, I., Heitland, N., Wagner, S., Nakad, R., Mader, S., Petersen, C., Kowallik, Rosenstiel, P., Felix, M.-A., & Schulenburg, H. (2016). The native microbiome of the nematode *Caenorhabditis elegans*: gateway to a new host-microbiome model. *BMC Biology, 14*(38). https://doi.org/10.1186/s12915-016-0258-1

Engelmann, I., Griffon, A., Tichit, L., Montanana-Sanchis, F., Wang, G., Reinke, V., Waterston, R. H., Hillier, L. W., & Ewbank, J. J. (2011). A comprehensive analysis of gene expression changes provoked by bacterial and fungal infection in *C. elegans. PLoS ONE, 6*(5), e19055. https://doi.org/10.1371/journal.pone.0019055

Evans, E. A., Kawli, T., & Tan, M. (2008). *Pseudomonas aeruginosa* suppresses host immunity by activating the DAF-2 insulin-like signaling pathway in *Caenorhabditis elegans. PLoS Pathogens, 4*(10), e1000175. https://doi.org/10.1371/journal.ppat.1000175

Ewbank, J. J. Signaling in the immune response (January 23, 2006), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.83.1

Farra, A., Islam, S., Str, A., Mikael, S., & Wretlind, B. (2008). Role of outer membrane protein OprD and penicillin-binding proteins in resistance of *Pseudomonas aeruginosa* to imipenem and meropenem. *International Journal of Antimicriobial Agents, 31,* 427–433. https://doi.org/10.1016/j.ijantimicag.2007.12.016

Feinbaum, R. L., Urbach, J. M., Liberati, N. T., Djonovic, S., Adonizio, A., Carvunis, A., & Ausubel, F. M. (2012). Genome-wide identification of *Pseudomonas aeruginosa* virulence-related genes using a *Caenorhabditis elegans* infection model. *PLoS Pathogens, 8*(7), e1002813. https://doi.org/10.1371/journal.ppat.1002813

Félix, M., & Duveau, F. (2012). Population dynamics and habitat sharing of natural populations of *Caenorhabditis elegans* and *C. briggsae*. *BMC Biology*, *10*(1), 59. https://doi.org/10.1186/1741-7007-10-59

Fouhy, Y., Scanlon, K., Schouest, K., Spillane, C., Crossman, L., Avison, M. B., Ryan, R. P., & Dow, J. M. (2007). Diffusible signal factor-dependent cell-cell signaling and virulence in the nosocomial pathogen *Stenotrophomonas maltophilia*. *Journal of Bacteriology, 189*(13), 4964–4968. https://doi.org/10.1128/JB.00310-07

Garigan, D., Hsu, A., Fraser, A. G., Kamath, R. S., Ahringer, J., & Kenyon, C. (2002). Genetic analysis of tissue aging in *Caenorhabditis elegans*: A role for heat-shock factor and bacterial proliferation. *Genetics, 161*, 1101–1112.

Garsin, D. A., Villanueva, J. M., Begun, J., Kim, D. H., Costi, D., Calderwood, S. B., Ruvkun, G., & Ausubel, F. M. (2003). Long-lived *C. elegans daf-2* mutants are resistant to bacterial pathogens. *Science, 300*(5627), 1921.

Gould, V. C., Okazaki, A., & Avison, M. B. (2006). Beta-lactam resistance and beta-lactamase expression in clinical *Stenotrophomonas maltophilia* isolates having defined phylogenetic relationships. *Journal of Antimicrobial Chemotherapy, 57*(2), 199–203. https://doi.org/10.1093/jac/dki453

Green, E. R., & Mecsas, J. (2016). Bacterial Secretion Systems – An overview. *Microbiology Spectrum, 4*(1), 1–32. https://doi.org/10.1128/microbiolspec.VMBF-0012-2015.Bacterial

Gudeta, D. D., Pollini, S., Docquier, J., Bortolaia, V., Rossolini, M., & Guardabassi, L. (2016). Biochemical characterization of CPS-1, a subclass B3 metallo-beta-lactamase from a Chryseobacterium piscium soil isolate. *Antimicrobial Agents and Chemotherapy, 60*(3), 1869–1873. https://doi.org/10.1128/AAC.01924-15.Address

Haber, M., Schungel, M., Putz, A., Muller, S., Hasert, B., & Schulenburg, H. (2018). Evolutionary history of *Caenorhabditis elegans* inferred from microsatellites: Evidence for spatial and temporal genetic differentiation and the occurrence of outbreeding. *Molecular Biology and Evolution, 22*(1), 160–173. https://doi.org/10.1093/molbev/msh264

He, G., Kuroda, T., Mima, T., Morita, Y., Mizushima, T., & Tsuchiya, T. (2004). An H+-coupled multidrug efflux pump, PmpM, a member of the MATE family of transporters, from *Pseudomonas aeruginosa*. *Journal of Bacteriology, 186*(1), 262–265. https://doi.org/10.1128/JB.186.1.262

Hidron, A. I., Edwards, J. R., Patel, J., Teresa, C., Sievert, D. M., Pollock, D. A., & Fridkin, S. K. (2008). Antimicrobial-resistant pathogens associated with healthcare-associated infections: Annual summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention. *Infection Control and Hospital Epidemiology, 29*(11), 996–1011. https://doi.org/10.1086/591861

Higgins, C. S., Avison, M. B., Jamieson, L., Simm, A. M., Bennett, P. M., & Walsh, T. R. (2001). Characterization, cloning and sequence analysis of the inducible *Ochrobactrum anthropi* AmpC beta-lactamase. *Journal of Antimicrobial Chemotherapy, 47*(6), 745–754.

Huang, D. W., Lempicki, R. a, & Sherman, B. T. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols, 4*(1), 44–57. https://doi.org/10.1038/nprot.2008.211

Huedo, P., Yero, D., Martínez-Servat, S., Estibariz, I., Planell, R., Martínez, P., Ruyra, A., Roher, N., Roca, I., Vila, J., Caura, X., & Gibert, I. (2014). Two different *rpf* clusters distributed among a population of *Stenotrophomonas maltophilia* clinical strains display differential diffusible signal factor production and virulence regulation. *Journal of Bacteriology, 196*(13), 2431–2442. https://doi.org/10.1128/JB.01540-14

Huffman, D. L., Abrami, L., Sasik, R., Corbeil, J., Goot, F. G. Van Der, & Aroian, R. V. (2004). Mitogen-activated protein kinase pathways defend against bacterial pore-forming toxins. *Proceedings of the National Academy of Sciences, 101*(30), 10995–11000.

Ikeda, T., Yasui, C., Hoshino, K., Arikawa, K., & Nishikawa, Y. (2007). Influence of lactic acid bacteria on longevity of *Caenorhabditis elegans* and host defense against *Salmonella enterica* serovar enteritidis. *Applied and Environmental Microbiology, 73*(20), 6404–6409. https://doi.org/10.1128/AEM.00704-07

Irazoqui, J. E., Troemel, E. R., Feinbaum, R. L., Luhachack, L. G., Cezairliyan, B. O., & Ausubel, F. M. (2010a). Distinct pathogenesis and host responses during infection of *C. elegans* by *P. aeruginosa* and *S. aureus. PLoS Pathogens, 6*(7), e1000982. https://doi.org/10.1371/journal.ppat.1000982

Irazoqui, J. E., Urbach, J. M., & Ausubel, F. M. (2010b). Evolution of host innate defence: insights from *C. elegans* and primitive invertebrates. *Nature Reviews Immunology, 10*(1), 47–58. https://doi.org/10.1038/nri2689.Evolution

Jankiewicz, U., Larkowska, E., & Brzezinska, M. S. (2016). Production, characterization, gene cloning, and nematocidal activity of the extracellular protease from *Stenotrophomonas maltophilia* N4. *Journal of Bioscience and Bioengineering, 121*(6), 614–618. https://doi.org/10.1016/j.jbiosc.2015.11.011

Johnning, A., Moore, E. R. B., Svensson-stadler, L., Shouche, Y. S., Larsson, D. G. J., & Kristiansson, E. (2013). Acquired genetic mechanisms of a multiresistant bacterium isolated from a treatment plant receiving wastewater from antibiotic production. *Applied and Environmental Microbiology, 79*(23), 7256–7263. https://doi.org/10.1128/AEM.02141-13

Johnson, G. L., & Lapadat, R. (2002). Mitogen-activated protein kinase pathways mediated by ERK, JNK, and p38 protein kinases. *Science, 298*(5600), 1911–1912.

Juhas, M., van der Meer, J. R., Gaillard, M., Harding, R. M., Hood, D. W., & Crook, D. W. (2009). Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiology Review, 33*, 376–393. https://doi.org/10.1111/j.1574-6976.2008.00136.x

Karin, M. (1995). The regulation of AP-1 activity by Mitogen-activated protein kinases. *The Journal of Biological Chemistry, 270*(26), 16483–16486.

Kato, Y., Aizawa, T., Hoshino, H., Kawano, K., Nitta, K., & Zhang, H. (2002). *abf-1* and *abf-2*, ASABF-type antimicrobial peptide genes in *Caenorhabditis elegans. Biochemical Journal, 15(361(Pt 2))*, 221–230.

Kawli, T., & Tan, M. (2008). Neuroendocrine signals modulate the innate immunity of *Caenorhabditis elegans* through insulin signaling. *Nature Immunology, 9*(12), 1415–1424.

Kenyon, C., Chang, J., Gensch, E., Rudner, A., & Tabtiang, R. (1993). A *C. elegans* mutant that lives twice as long as wild type. *Nature, 366*, 461–464.

Kerry, S., TeKippe, M., Gaddis, N. C., & Aballay, A. (2006). GATA transcription factor required for immunity to bacterial and fungal pathogens. *PLoS ONE, 1*, e77. https://doi.org/10.1371/journal.pone.0000077

Kesika, P., & Balamurugan, K. (2012). Studies on *Shigella boydii* infection *in Caenorhabditis elegans* and bioinformatics analysis of immune regulatory protein interactions. *Biochimica et Biophysica Acta, 1824*(12), 1449–1456. https://doi.org/10.1016/j.bbapap.2012.07.008

Kim, D. H. (2013). Bacteria and the aging and longevity of *Caenorhabditis elegans*. *Annual Review of Genetics, 47*, 233–246. https://doi.org/10.1146/annurev-genet-111212-133352

Kim, D. H., & Ewbank, J. J. (2015). Signaling in the innate immune response (December 22, 2015), *WormBook*, ed. The *C. elegans* Research Community, WormBook, https://doi.org/10.1111/mec.13526

Kim, D. H., Feinbaum, R., Alloing, G., Emerson, F. E., Garsin, D. A., Inoue, H., Tanaka-Hino, M., Hisamoto, N., Matsumoto, K., Tan, M.-W., & Ausubel, F. M. (2002). A conserved p38 MAP Kinase pathway in *Caenorhabditis elegans* innate immunity. *Science, 297*, 623–626.

Kimura, K. D., Tissenbaum, H. A., Liu, Y., & Ruvkun, G. (1997). *daf-2*, an insulin receptor-like gene that regulates longevity and diapause in *Caenorhabditis elegans. Science, 277*(5328), 942–946.

Kirienko, N. V, Kirienko, D. R., Larkins-Ford, J., Wahlby, C., Ruvkun, G., & Ausubel, F. M. (2014). *Pseudomonas aeruginosa* disrupts *Caenorhabditis elegans* iron homeostasis, causing a hypoxic response and death. *Cell Host Microbe, 13*(4), 406–416. https://doi.org/10.1016/j.chom.2013.03.003

Landis, J. N., & Murphy, C. T. (2010). Integration of diverse inputs in the regulation of *Caenorhabditis elegans* DAF-16/FOXO. *Developmental Dynamics, 239*, 1405–1412. https://doi.org/10.1002/dvdy.22244

Laviad-Shitrit, S., Göker, M., Huntemann, M., Clum, A., Pillay, M., Palaniappan, K., Varghese, N., Mikhailova, N., Stamitis, D., Reddy, T. B. K., Daum, C., Shapiro, N., Markowitz, V., Ivanova, N., Woyke, T., Klenk, H.-P., Kyrpides, N. C., & Halpern, M. (2017). High quality permanent draft genome sequence of *Chryseobacterium bovis* DSM 19482 T, isolated from raw cow milk. *Standards in Genomic Sciences, 12*(31), 1–7. https://doi.org/10.1186/s40793-017-0242-6

Lebuhn, M., Achouak, W., Schloter, M., Berge, O., Meier, H., Barakat, M., Hartmann, A., & Heulin, T. (2000). Taxonomic characterization of *Ochrobactrum* sp. isolates from soil samples and wheat roots, and description of *Ochrobactrum tritici* sp. nov. and *Ochrobactrum grignonense* sp. nov. *Internation Journal of Systematic and Evolutionary Microbiology, 50*, 2207–2223.

Letterio, J. J. (2005). TGF-β signaling in T cells: Roles in lymphoid and epithelial neoplasia. *Oncogene, 24*(37), 5701–5712. https://doi.org/10.1038/sj.onc.1208922

Liberati, N. T., Fitzgerald, K. A., Kim, D. H., Feinbaum, R., Golenbock, D. T., & Ausubel, F. M. (2004). Requirement for a conserved Toll/interleukin-1 resistance domain protein in the *Caenorhabditis elegans* immune response. *Proceedings of the National Academy of Sciences, 101*(17), 6593–6598.

Lin, Y., Jeng, Y., Lin, M., Yu, K., Wang, F., & Liu, C. (2010). Clinical and microbiological characteristics of *Chryseobacterium indologenes* bacteremia. *Journal of Microbiology, Immunology and Infection, 43*(6), 498–505. https://doi.org/10.1016/S1684-1182(10)60077-1

Lindsay, S. A., & Wasserman, S. A. (2014). Conventional and non-conventional Drosophila Toll signaling. *Developmental and Comparative Immunology, 42*(1), 16–24. https://doi.org/10.1016/j.dci.2013.04.011

Lira, F., Berg, G., & Martínez, J. L. (2017). Double-face meets the bacterial world: The opportunistic pathogen *Stenotrophomonas maltophilia. Frontiers in Microbiology, 8*, 2190. https://doi.org/10.3389/fmicb.2017.02190

Liu, G., & Zhao, Y. (2007). Toll-like receptors and immune regulation: their direct and indirect modulation on regulatory CD4+ and CD25+ T cells. *Immunology, 122*(2), 149-156. https://doi.org/10.1111/j.1365-2567.2007.02651.x

Lundberg, D. S., Lebeis, S. L., Paredes, S. H., Yourstone, S., Gehring, J., Malfatti, S., Tremblay, J., Engeibrektson, A., Kunin, V., del Rio, T. G., Edgar, R. C., Eickhorst, T., Ley, R. E., Hugenholtz, P., Tringe, S. G., & Dangl, J. L. (2012). Defining the core *Arabidopsis thaliana* root microbiome. *Nature, 488*, 86–94.

Mallo, G. V., Kurz, C. L., Couillault, C., Pujol, N., Granjeaud, S., Kohara, Y., & Ewbank, J. J. (2002). Inducible antibacterial defense system in *C. elegans. Current Biology, 12*(14), 1209–1214. https://doi.org/10.1016/S0960-9822(02)00928-4

Maravic, A., Skoc, M., Samanic, I., & Puizina, J. (2013). Profile and multidrug resistance determinants of *Chryseobacterium indologenes* from seawater and marine fauna. *World Journal of Microbiology and Biotechnology, 29*, 515–522. https://doi.org/10.1007/s11274-012-1205-0

Mathee, K., Narasimhan, G., Valdes, C., Qiu, X., Matewish, J. M., Koehrsen, M., Rokas, A., Yandava, C. N., Engels, R., Zeng, E., Olavarietta, R., Doud, M., Smith, R. S., Montgomery, P., White, J. R., Godfrey, P. A., Kodira, C., Birren, B., Galagan, J. E., & Lory, S. (2008). Dynamics of *Pseudomonas aeruginosa* genome evolution. *Proceedings of the National Academy of Sciences, 105*(8), 3100–3105.

McElwee, J., Bubb, K., & Thomas, J. H. (2003). Transcriptional outputs of the *Caenorhabditis elegans* forkhead protein DAF-16. *Aging Cell, 2*(2), 111–121. https://doi.org/10.1046/j.1474-9728.2003.00043.x

Mei, G., Yan, X., Turak, A., Luo, Z., & Zhang, L. (2010). AidH, an alpha/beta-hydrolase fold family member from an *Ochrobactrum* sp. strain, is a novel N-acylhomoserine lactonase. *Applied and Environmental Microbiology, 76*(15), 4933–4942. https://doi.org/10.1128/AEM.00477-10

Merkx-jacques, A., Coors, A., Brousseau, R., Masson, L., Mazza, A., Tien, Y., & Topp, E. (2013). Evaluating the pathogenic potential of environmental *Escherichia coli* by using the *Caenorhabditis elegans* infection model. Applied and *Environmental Microbiology, 79*(7), 2435–2445. https://doi.org/10.1128/AEM.03501-12

Miltsch, S. M., Seeberger, P. H., & Lepenies, B. (2014). The C-type lectin-like domain containing proteins Clec-39 and Clec-49 are crucial for *Caenorhabditis elegans* immunity against *Serratia marcescens* infection. *Developmental and Comparative Immunology, 45*(1), 67–73. https://doi.org/10.1016/j.dci.2014.02.002

Mochii, M., Yoshida, S., Morita, K., Kohara, Y., & Ueno, N. (1999). Identification of transforming growth factor-beta-regulated genes in *Caenorhabditis elegans* by differential hybridization of arrayed cDNAs. *Proceedings of the National Academy of Sciences, 96*(26), 15020–15025.

Murphy, C. T., Lee, S., & Kenyon, C. (2007). Tissue entrainment by feedback regulation of insulin gene expression in the endoderm of *Caenorhabditis elegans. Proceedings of the National Academy of Sciences, 104*(48), 19046–19050.

Murphy, C. T., McCarroll, S. A., Bargmann, C. I., Fraser, A., Kamath, R. S., Ahringer, J., Li, H., & Kenyon, C. (2003). Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans. Nature, 424*(6946), 277–283. https://doi.org/10.1038/nature01789

Ni, J., Yan, Q., Yu, Y., & Zhang, T. (2013). Factors influencing the grass carp gut microbiome and its effect on metabolism. *FEMS Microbiology Ecology, 87*(3), 704-714. https://doi.org/10.1111/1574-6941.12256

Nishioka, T., Elsharkawy, M. M., Suga, H., Kageyama, K., Hyakumachi, M., & Shimizu, M. (2016). Development of culture medium for the isolation of flavobacterium and chryseobacterium from rhizosphere soil. *Microbes and Environment, 31*(2), 104–110. https://doi.org/10.1264/jsme2.ME15144

O'Rourke, D., Baban, D., Demidova, M., Mott, R., & Hodgkin, J. (2006). Genomic clusters, putative pathogen recognition molecules, and antimicrobial genes are induced by infection of *C. elegans* with *M. nematophilum. Genome Research, 16*(8), 1005–1016. https://doi.org/10.1101/gr.50823006

Ozdemir, D., Soypacacı, Z., Sahin, I., Bicik, Z., & Sencan, I. (2006). *Ochrobactrum anthropi* endocarditis and septic shock in a patient with no prosthetic valve or rheumatic heart disease: Case report and review of the literature. *Japanese Journal of Infectious Disease, 59*, 264–265.

Pai, H., Kim, J., Kim, J., Lee, J. I. H., Choe, K. W., & Gotoh, N. (2001). Carbapenem resistance mechanisms in *Pseudomonas aeruginosa* clinical isolates. *Antimicrobial Agents and Chemotherapy, 45*(2), 480–484. https://doi.org/10.1128/AAC.45.2.480

Pees, B., Yang, W., Zárate-potes, A., Schulenburg, H., & Dierking, K. (2016). High innate immune specificity through diversified c-type lectin-like domain proteins in invertebrates. *Journal of Innate Immunity, 8*, 129–142. https://doi.org/10.1159/000441475

Pellegrino, M. W., Nargund, A. M., Kirienko, N. V, Gillis, R., Fiorese, C. J., & Haynes, C. M. (2014). Mitochondrial UPR-regulated innate immunity provides resistance to pathogen infection. *Nature, 516*(7531), 414–417. https://doi.org/10.1038/nature13818

Petersen, C., Dirksen, P., & Schulenburg, H. (2015). Why we need more ecology for genetic models such as *C. elegans. Trends in Genetics, 31*(3), 120–127. https://doi.org/10.1016/j.tig.2014.12.001

Piddock, L. J. V. (2006). Multidrug-resistance efflux pumps — not just for resistance. *Nature Review Microbiology, 4*(8), 629–636.

Pompilio, A., Pomponio, S., Crocetta, V., Gherardi, G., Verginelli, F., Fiscarelli, E., Dicuonzo, G., Savini, V., D'Antonio, D., & Bonaventura, G. Di. (2011). Phenotypic and genotypic characterization of *Stenotrophomonas maltophilia* isolates from patients with cystic fibrosis: Genome diversity, biofilm formation, and virulence. *BMC Biology, 11*(159), 1–16.

Porta, H., Cancino-Rodezno, A., Soberón, M., & Bravo, A. (2011). Role of MAPK p38 in the cellular responses to pore-forming toxins. *Peptides, 32*(3), 601–606. https://doi.org/10.1016/j.peptides.2010.06.012

Pujol N, Link, E. M., Liu, L. X., Kurz, C. L., Alloing, G., Tan, M-W., Ray, K. P., Solari, R., Johnson, C. D., Ewbank, J. J. (2001). A reverse genetic analysis of components of the Toll signaling pathway in *Caenorhabditis elegans*. Current Biology, 11(11), 809–21.

Pujol, N., Cypowyj, S., Ziegler, K., Millet, A., Astrain, A., Goncharov, A., Jin, Y., Chisholm, A. D., & Ewbank, J. J. (2008). Distinct innate immune responses to infection and wounding in the *C. elegans* epidermis. *Current Biology, 18*, 481–489. https://doi.org/10.1016/j.cub.2008.02.079

Rae, R., Riebesell, M., Dinkelacker, I., Wang, Q., Herrmann, M., Weller, A. M., Dieterich, C., & Sommer, R. J. (2008). Isolation of naturally associated bacteria of necromenic *Pristionchus* nematodes and fitness consequences. *The Journal of Experimental Biology, 211*, 1927–1936. https://doi.org/10.1242/jeb.014944

Rashid, R., Morohoshi, T., Someya, N., & Ikeda, T. (2011). Degradation of N-acylhomoserine lactone quorum sensing signaling molecules by potato root surface-associated *Chryseobacterium Strains*. *Microbes and Environment, 26*(2), 144–148. https://doi.org/10.1264/jsme2.ME10207

Ren, M., Feng, H., Fu, Y., Land, M., & Rubin, C. S. (2009). Protein kinase D (DKF-2), a diacylglycerol effector, is an essential regulator of *C. elegans* innate immunity. *Immunity, 30*(4), 521–532. https://doi.org/10.1016/j.immuni.2009.03.007.Protein

Richardson, C. E., Kooistra, T., & Kim, D. H. (2010). An essential role for XBP-1 in host protection against immune activation in *C. elegans. Nature, 463*(7284), 1092–1095. https://doi.org/10.1038/nature08762

Roberts, A. F., Gumienny, T. L., Gleason, R. J., Wang, H., & Padgett, R. W. (2010). Regulation of genes affecting body size and innate immunity by the DBL-1/BMP-like pathway in *Caenorhabditis elegans. BMC Developmental Biology, 10*(61).

Roeder, T., Stanisak, M., Gelhaus, C., Bruchhaus, I., Gro, J., & Leippe, M. (2010). Caenopores are antimicrobial peptides in the nematode *Caenorhabditis elegans* instrumental in nutrition and immunity. *Developmental and Comparative Immunology, 34*, 203–209. https://doi.org/10.1016/j.dci.2009.09.010

Samuel, B. S., Rowedder, H., Braendle, C., Félix, M.-A., & Ruvkun, G. (2016). *Caenorhabditis elegans* responses to bacteria from its natural habitats. *Proceedings of the National Academy of Sciences, 113*(27), E3941–E3949. https://doi.org/10.1073/pnas.1607183113

Sánchez, M. B., & Martínez, J. L. (2015). The Efflux Pump SmeDEF contributes to trimethoprim-sulfamethoxazole resistance in *Stenotrophomonas maltophilia. Antimicrobial Agents and Chemotherapy, 59*(7), 4347–4348. https://doi.org/10.1128/AAC.00714-15

Savage-Dunn, C. TGF-β signaling (September 9, 2005), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.22.1

Schulenburg, H., & Boehnisch, C. (2008). Diversification and adaptive sequence evolution of *Caenorhabditis* lysozymes (Nematoda: Rhabditidae). *BMC Evolutionary Biology, 8*(114). https://doi.org/10.1186/1471-2148-8-114

Schulenburg, H., & Félix, M. (2017). The natural biotic environment of *Caenorhabditis elegans*. *Genetics, 206*(1), 55–86.

Schulte, R. D., Makus, C., Hasert, B., Michiels, N. K., & Schulenburg, H. (2010). Multiple reciprocal adaptations and rapid genetic change upon experimental coevolution of an animal host and its microbial parasite. *Proceedings of the National Academy of Sciences, 107*(16), 7359–7364. https://doi.org/10.1073/pnas.1003113107

Shapira, M., Hamlin, B. J., Rong, J., Chen, K., Ronen, M., & Tan, M. (2006). A conserved role for a GATA transcription factor in regulating epithelial innate immune responses. *Proceedings of the National Academy of Sciences, 103*(38), 14086–14091.

Sharma, P., Gupta, S. K., Diene, S. M., & Rolain, J.-M. (2015). Whole-genome sequence of *Chryseobacterium oranimense,* a colistin resistant bacterium isolated from a cystic

fibrosis patient in France. *Antimicrobial Agents and Chemotherapy, 59*(3), 1696-1706. https://doi.org/10.1128/AAC.02417-14

Shaye, D. D., & Greenwald, I. (2011). OrthoList: A compendium of *C. elegans* genes with human orthologs. *PLoS ONE, 6*(5). https://doi.org/10.1371/journal.pone.0020085

Shin, J., Lee, S., Go, M-J., Lee, S. Y., Kim, S. C., Lee, C-H., & Cho, B-K. (2016). Analysis of the mouse gut microbiome using full-length 16S rRNA amplicon sequencing. *Scientific Reports 6*(29681). https://doi.org/10.1038/srep29681

Shivers, R. P., Kooistra, T., Chu, S. W., Pagano, D. J., & Kim, D. H. (2009). Tissue-specific activities of an immune signaling module regulate physiological responses to pathogenic and nutritional bacteria in *C. elegans. Cell Host and Microbe, 6*(4), 321–330. https://doi.org/10.1016/j.chom.2009.09.001

Shivers, R. P., Pagano, D. J., Kooistra, T., Richardson, C. E., Reddy, K. C., Whitney, J. K., Kamanzi, O., Matsumoto, K., Hisamoto, N., & Kim, D. H. (2010). Phosphorylation of the conserved transcription factor ATF-7 by PMK-1 p38 MAPK regulates innate immunity in *Caenorhabditis elegans. PLoS Genetics, 6*(4), e1000892. https://doi.org/10.1371/journal.pgen.1000892

Shtonda, B. B., & Avery, L. (2006). Dietary choice behavior in *Caenorhabditis elegans. The Journal of Experimental Biology, 209*, 89–102. https://doi.org/10.1242/jeb.01955

Smith, J. (2001). The social evolution of bacterial pathogenesis. *Proceedings of the Royal Society of London B, 268*, 61–69. https://doi.org/10.1098/rspb.2000.1330

Sousa, T., Branco, R., Piedade, A. P., & Morais, P. V. (2015). Hyper accumulation of arsenic in mutants of *Ochrobactrum tritici* silenced for arsenite efflux pumps. *PLoS ONE, 10*(7), 1–14. https://doi.org/10.1371/journal.pone.0131317

Stevens, A. M., Schuster, M., & Rumbaugh, K. P. (2012). Working together for the common good: Cell- cell communication in bacteria. *Journal of Bacteriology, 194*(9), 2131–2141. https://doi.org/10.1128/JB.00143-12

Streit, J. M., Jones, R. N., Sader, H. S., & Fritsche, T. R. (2004). Assessment of pathogen occurrences and resistance profiles among infected patients in the intensive care unit: Report from the SENTRY Antimicrobial Surveillance Program (North America, 2001). *International Journal of Antimicriobial Agents, 24*, 111–118. https://doi.org/10.1016/j.ijantimicag.2003.12.019

Styer, K. L., Singh, V., Macosko, E., Steele, S. E., Bargmann, C. I., & Aballay, A. (2008). Innate immunity in *Caenorhabditis elegans* is regulated by neurons expressing NPR-1/GPCR. *Science, 322*(5900), 460–465. https://doi.org/10.1126/science.1163673

Tan, M.-W., Mahajan-Miklos, S., & Ausubel, F. M. (1999). Killing of *Caenorhabditis elegans* by *Pseudomonas aeruginosa* used to model mammalian bacterial pathogenesis. *Proceedings of the National Academy of Sciences, 96*, 715–720.

Tenor, J. L., & Aballay, A. (2008). A conserved Toll-like receptor is required for *Caenorhabditis elegans* innate immunity. *EMBO reports 9*(1), 103-109. https://doi.org/10.1038/sj.embor.7401104

Troemel, E. R., Chu, S. W., Reinke, V., Lee, S. S., Ausubel, F. M., & Kim, D. H. (2006). p38 MAPK regulates expression of immune response genes and contributes to longevity *in C. elegans. PLoS Genetics, 2*(11), e183. https://doi.org/10.1371/journal.pgen.0020183

Tullet, J. M. A., Hertweck, M., An, J. H., Baker, J., Hwang, J. Y., Liu, S., Oliveira, R. P., Baumeister, R., & Blackwell, T. K. (2008). Direct inhibition of the longevity-promoting factor SKN-1 by insulin-like signaling in *C. elegans. Cell, 132*(6), 1025–1038. https://doi.org/10.1016/j.cell.2008.01.030

van den Berg, L. M., Gringhuis, S. I., & Geijtenbeek, T. B. H. (2012). An evolutionary perspective on C-type lectins in infection and immunity. *Annals of the New York Academy of Sciences, 1253*, 149–158. https://doi.org/10.1111/j.1749-6632.2011.06392.x

Virk, B., Correia, G., Dixon, D. P., Feyst, I., Jia, J., Oberleitner, N., … Weinkove, D. (2012). Excessive folate synthesis limits lifespan in the *C. elegans*: *E. coli* aging model. *BMC Biology, 10*(67). https://doi.org/10.1186/1741-7007-10-67

Virk, B., Jia, J., Maynard, C. A., Helliwell, N., Cipinska, M., Weinkove, D., Briggs, Z., Hodge, E., Edwards, R., Ward, J., Gems, D., & Weinkove, D. (2016). Folate acts in *E. coli* to accelerate *C. elegans* aging independently of bacterial biosynthesis. *Cell Reports, 14*(7), 1611–1620. https://doi.org/10.1016/j.celrep.2016.01.051

Visvikis, O., Ihuegbu, N., Labed, S. A., Luhachack, L. G., Alves, A. F., Wollenberg, A. C., Stuart, L. M., Stormo, G. D., & Irazoqui, J. E. (2014). Innate host defense requires TFEB-mediated transcription of cytoprotective and antimicrobial genes. *Immunity, 40*(6), 896–909. https://doi.org/10.1016/j.immuni.2014.05.002

Wang, B., Ya, M., Lv, L., Ling, Z., & Li, L. (2017). The human microbiota in health and disease. *Engineering, 3*(1), 71-82. https://doi.org/10.1016/J.ENG.2017.01.008

Wang, T., Jiang, X., Feng, C., Li, A., Dong, H., Wu, S., & Zheng, B. (2016). Whole genome sequencing uncovers a novel IND-16 metallo-β-lactamase from an extensively drug-resistant *Chryseobacterium indologenes* strain J31. *Gut Pathogens, 8*(47), 1–8. https://doi.org/10.1186/s13099-016-0130-4

Wang, X., Jin, D., Zhou, L., & Zhang, Z. (2015). Draft genome sequence of *Ochrobactrum anthropi* Strain W13P3, a halotolerant polycyclic aromatic hydrocarbon-degrading bacterium. *Genome Announcements, 3*(4), 1–2. https://doi.org/10.1128/genomeA.00867-15.Copyright

White, C. V, Darby, B. J., Bredden, R. J., & Herman, M. A. (2016). A *Stenotrophomonas maltophilia* strain evades a major *Caenorhabditis elegans* defense pathway. *Infection and Immunity, 84*(2), 524–536.

Yang, W., Dierking, K., Esser, D., Tholey, A., Leippe, M., Rosenstiel, P., & Schulenburg, H. (2015). Overlapping and unique signatures in the proteomic and transcriptomic responses of the nematode *Caenorhabditis elegans* toward pathogenic *Bacillus thuringiensis*. *Developmental and Comparative Immunology*, *51*(1), 1–9. https://doi.org/10.1016/j.dci.2015.02.010

Yun, J-H., Roh, S. W., Whon, T. W., Jung, M-J., Kim, M-S., Park, D-S., Yoon, C., Nam, Y-D., Kim, Y-J., Choi, J-H., Kim, J-Y., Shin, N-R., Kim, S-H., Lee, W-J., & Bae, J-W. (2014). Insect gut bacterial diversity determined by environmental habitat, diet, developmental stage, and phylogeny of host. *Applied Environmental Microbiology, 80*(17), 5254-5264. https://doi.org/10.1128/AEM.01226-14

Zelensky, A. N., & Gready, J. E. (2005). The C-type lectin-like domain superfamily. *FEMS Journal, 272*, 6179–6217. https://doi.org/10.1111/j.1742-4658.2005.05031.x

Zhang, F., Berg, M., Dierking, K., Félix, M., Shapira, M., Samuel, B. S., & Schulenburg, H. (2017). *Caenorhabditis elegans* as a model for microbiome research. *Frontiers in Microbiology, 8*(485), 1–10. https://doi.org/10.3389/fmicb.2017.00485

Zhang, K., & Kaufman, R. J. (2004). Signaling the unfolded protein response from the endoplasmic reticulum. *The Journal of Biological Chemistry, 279*(25), 25935–25938. https://doi.org/10.1074/jbc.R400008200

Zhang, P., Judy, M., Lee, S., & Kenyon, C. (2013). Direct and indirect gene regulation by a life-extending FOXO protein in *C. elegans*: Roles for GATA factors and lipid gene regulators. *Cell Metabolism, 17*, 85–100. https://doi.org/10.1016/j.cmet.2012.12.013

Zhang, Y. (2008). Neuronal mechanisms of *Caenorhabditis elegans* and pathogenic bacteria interactions. *Current Opinion in Microbiology*, *11*(3), 257–261. https://doi.org/10.1016/j.mib.2008.04.003

Zhang, Z. (2013). Animal biodiversity: An update of classification and diversity in 2013. *Zootaxa, 3703*(1), 5–11.

Zilber-Rosenberg, I., & Rosenberg, E. (2008). Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiology Review, 32*, 723–735. https://doi.org/10.1111/j.1574-6976.2008.00123.x

Chapter 2 - Identification and characterization of differentially expressed genes in *Caenorhabditis elegans* in response to pathogenic and nonpathogenic *Stenotrophomonas maltophilia*

Introduction

Stenotrophomonas maltophilia is a Gram-negative, nosocomial pathogen that can cause infection in immunocompromised patients. *S. maltophilia* is often found in patients with cystic fibrosis and lung cancer, and is associated with infections such as pneumonia, endocarditis, bacteremia, and meningitis (reviewed in Brooke, 2012). Although not highly virulent, *S. maltophilia* is multi-drug resistant and capable of forming biofilms (Elting & Bodey, 1990; Jägevall, Rabe, & Pedersen, 2011), thus developing treatment methods for this pathogen is becoming an increasing concern. *S. maltophilia* is ubiquitous within the environment, commonly found in aqueous sources, soils, and associated with plant roots, and can also be isolated in hospitals from water sources and medical devices (reviewed in Brooke, 2012; Chang et al., 2015).

Sequencing and functional validation have identified virulence factors such as antibiotic resistant and quorum sensing mechanisms in both clinical and environmental isolates of *S. maltophilia* (Adamek et al., 2011; Berg, Roskot, & Smalla, 1999; Lira, Berg, & Martínez, 2017). In addition, phylogenetic and antibiotic profile analyses of environmental and clinical isolates revealed heterogeneity between strain origins (Berg, Roskot, & Smalla, 1999; Lira, Berg, & Martínez, 2017). This is important because it suggests that environmental *S. maltophilia* isolates have the potential to become infectious to humans without acquiring new virulence factors. However, strains isolated from clinical settings have higher mutation rates than environmental

isolates, suggesting clinical isolates may be better able to adapt to stressors such as the immune response and antibiotics (Turrientes et al., 2010). Sequencing and proteomic analysis have also identified differences in genomic features between strains of *S. maltophilia*, including virulence factors (Adamek, Linke, & Schwartz, 2014; Ferrer-Navarro et al., 2013; Lira, Berg, & Martínez, 2017; Rocco et al., 2009). Therefore, diversity exists between strains that results in different virulence mechanisms (Adamek et al., 2014; Kaiser, Biehler, & Jonas, 2009). Although studies have identified virulence factors within *S. maltophilia* genomes, their functions have rarely been analyzed and mechanisms of host responses are poorly understood. Therefore, we recently established *Caenorhabditis elegans* as a model to study host responses to *S. maltophilia* infection (White et al., 2016).

C. elegans are bacterivores found in decaying fruits and stems where they are in constant contact with many bacterial species. These interactions come in many forms, including symbiotic, predator-prey, and pathogen-host. Recent studies have found that *Stenotrophomonas* is one of the most abundant genera of bacteria found in the native microbiome of *C. elegans* (M. Berg et al., 2016; Dirksen et al., 2016; Zhang et al., 2017). Furthermore, *Stenotrophomonas* is found in higher abundance within the microbiome than in the rotting substrate (Dirksen et al., 2016; Zhang et al., 2017), suggesting that it accumulates within the intestine, a common signature of pathogenesis in *C. elegans* (Garigan et al., 2002; McGhee et al., 2007). In fact, many of these isolates of *Stenotrophomonas* were found to be detrimental to the health of *C. elegans*, while few were found to be beneficial (Dirksen et al., 2016; Samuel et al., 2016). This is consistent with previous observations that strains of *S. maltophilia* show varying pathogenicity to *C. elegans*, amoeba (*Dictyostelium discoideum* and *Acanthamoeba castellanii*), and zebrafish (Adamek et al., 2011; Ferrer-Navarro et al., 2013; White et al., 2016). This suggests that

different *S. maltophilia* strains utilize different virulence mechanisms that result in different host responses.

Many innate immune pathways in *C. elegans* are conserved from invertebrates to mammals, making it a model organism for studying pathogen-host interactions and innate immunity. Details of important immune and defense pathways in *C. elegans* are reviewed in Chapter 1, but briefly, the p38 mitogen-activated protein kinase (MAPK) pathway plays a role in defense against several pathogens, including *S. maltophilia*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, and *Salmonella enterica* (Aballay et al., 2003; Kim et al., 2002; Sifri et al., 2003; White et al., 2016). In addition, activation of the insulin-like signaling pathway receptor DAF-16 decreases bacterial packing, suggesting that regulation of genes by DAF-16 defends against accumulation of bacteria in the intestine (Murphy, Mccarroll, et al., 2003). Analyses of mutations affecting genes in these pathways have identified downstream proteins involved in pathogen defense, such as lysozymes, C-lectins, and CUB-domain containing proteins (Alper et al., 2007; Troemel et al., 2006).

Although many important innate immune pathways and effectors have been identified, there are differences in responses to different bacterial pathogens. For example, one study comparing responses to intestinal pathogens *Serratia marcescens*, *Enterococcus faecalis*, and *Photorhabdus luminescens* found that only 11% of genes in *C. elegans* where commonly differentially expressed in response to all three species (Engelmann et al., 2011). This phenomenon could be due to species-specific responses to different pathogens, or the ability of bacteria to manipulate different host responses. Therefore, it is essential to study a variety of pathogens in order to fully understand the complexity of genetic mechanisms underlying pathogen defense.

Here, we used a transcriptomic approach to identify and characterize the genetic responses of *C. elegans* to several different *S. maltophilia* strains. Specifically, we performed RNA sequencing on *C. elegans* following exposure to either *E. coli* OP50 or one of three *S. maltophilia* strains: two pathogenic environmental isolates, JCMS and JV3, and one nonpathogenic clinical isolate, K279a. Using this experimental set-up, we identified and characterized responses that are common to both pathogenic *S. maltophilia* strains and responses that are strain-specific. In addition, we determined that responses to *S. maltophilia* were overall similar to responses to other pathogens. Finally, we chose candidate genes that represent common, *S. maltophilia*-specific, and strain-specific responses and determined that several candidate genes were important for survival of *C. elegans* upon exposure to *S. maltophilia* as well as other common *C. elegans* pathogens.

Materials and Methods

Nematode and bacteria strains and growth

The following *C. elegans* strains were obtained from the *Caenorhabditis* Genetics Center (CGC): RB1573 *dod-22(ok1918),* VC1749 *F55G11.8(gk3130) ZK185.2(gk828),* VC3059 *ZK6.11(ok3738),* VC2477 *T24B8.5(ok3236),* RB1893 *lys-1(ok2445),* VC2249 *dod-19(ok2679),* RB2095 *clec-67(ok2770),* VC2176 *nhr-110 (gk987),* RB2114 *sodh-1 (ok2799),* VC1011 *acdh-1(ok1489),* RB2473 *cpr-4(ok3413),* LIU1 [*dhs-3p::dhs-3::*GFP + *unc-76(+)*], AU78 [*T24B8.5p::*GFP*::unc-54*-3' UTR + *ttx-3p::*GFP*::unc-54*-3' UTR], CF2124 [*sodh-1p::*RFP(NLS) + *rol-6(su1006)*], VL717 [*acdh-1p::*GFP]. *C. elegans* strains containing the following alleles were obtained from the National BioResource Project (NBRP): *lys-2(tm2398), scl-2(tm2428), dhs-3(tm6151), F13D12.6(tm7051*), *pho-1(tm5302), C55A6.7*(*tm6807*), *dhs-2(tm7516), acox-1.4(tm6415).* All alleles were outcrossed 4 times, with worms screened via PCR

after each outcross to obtain homozygous mutants. Forward and reverse primers used to test for each desired allele can be found on Wormbase (wormbase.org).

C. elegans strains containing the following expression constructs and alleles were generated as described below: *mhEx284*[*F19B2.5p::F19B2.5*::wrmScarlet::*unc-54*-3' UTR] and *mhEx282*[*K08D8.4p::K08D8.4*::wrmScarlet::*unc-54*-3' UTR], *F08G2.5*(*mh86*), *K08D8.4(mh76), B0024.4(mh82),W02A2.8(mh87), K08D8.12(mh101), fbxa-77(mh94), Y82E9BR.5*(*mh93*), *C25F9.11*(*mh97*).

Bristol *N2* strain was obtained from the CGC and used as *wild-type*. All strains were maintained on nematode growth media (NGM) plates seeded with *E. coli* OP50 at 20° C.

Bacterial strains include *E. coli* OP50 from the CGC, *Stenotrophomonas maltophilia* JCMS isolated by our lab in association with nematodes from Konza Prairie near Manhattan, KS (White et al., 2016), *Stenotrophomonas maltophilia* K279a from R. Ryan (University College Cork), *Stenotrophomonas maltophilia* JV3 from J. Tiedje (Michigan State University), *Serratia marcescens* DB10 from CGC, *Staphylococcus aureus* NCTC8325 from J. Irazoqui (UMass Medical Center), *Xenorhabdus nematophila* 1462 from P. Stock (University of Arizona), and *Pseudomonas aeruginosa* PA14 from F. M. Ausubel (Harvard Medical School).

All bacteria strains were frozen at -80° C upon arrival to the lab and thawed frequently for experimentation. The *S. maltophilia* strains and also PA14 and DB10 are naturally Ampicillin resistant, thus were grown on Luria Broth (LB) agar containing 100 µg/mL Ampicillin to selectively isolate and maintain each strain while avoiding contamination. Other bacterial strains were grown on regular LB agar. Plates were incubated at 37° C overnight and kept at 4° C thereafter. *S. maltophilia* strains, PA14, and DB10 were grown in liquid LB containing 100 µg/mL Ampicillin, and other strains were grown in liquid LB and shaken overnight at 37° C

overnight. Liquid cultures were then seeded onto NGM and grown at room temperature overnight before being used for experimentation.

RNA isolation

Wild-type nematodes were synchronized by bleaching, plated on *E. coli* OP50, and maintained at 20° C. Synchronized larval stage 4 (L4) worms were washed several times in M9 buffer and transferred to treatment bacteria or *E. coli* OP50. Treatments included *S. maltophilia* strains K279a, JCMS, and JV3. After 12 hours of exposure to treatment bacteria at 25° C, worms were collected in M9 buffer and lysed in TRIzol® (Life Technologies). 12 hours of exposure to treatments was chosen because at this point bacterial accumulation in the intestine has begun (White et al., 2016), but almost all worms in each treatment were still alive. Only non-contaminated, un-starved populations were used for RNA extraction, and three biological replicates were collected for each treatment. Bulk RNA was extracted from these populations using PureLink RNA Mini Kit (Invitrogen), and DNase treated using On-Column PureLink® DNase Treatment (Invitrogen) following the manufacturer's protocol. RNA quality was checked by determining 260/280 and 260/230 absorbance ratios using a NanoDrop™ 8000 Spectrophotometer and observation of 18S and 28S rRNA bands using gel electrophoresis.

RNA sequencing and analysis

Extracted RNA was sent to the University of Kansas Center for Molecular Analysis of Disease Pathways Genome Sequencing core facility for library preparation and sequencing. Three biological replicates, consisting of pooled bulk nematode RNA, and two technical replicates of each biological replicate were sequenced for each treatment. Libraries were sequenced on Illumina HiSeq 2500 platform resulting in 100 base pair single-end reads. Sequence quality was assessed using FastQC.

Tophat 2, which uses the short-read mapping program Bowtie (Kim et al., 2013), was used to map reads to the *C. elegans* genome. Technical replicates were combined at this step. Transcriptome and genome versions WS235 were used as the reference (wormbase.org). Minimum intron length was set to 15 base pairs (*-i 15*) and the parameter for *-no-novel-juncs* was used. The remainder of settings were set to default. Cuffdiff, a program within Cufflinks, is used to compare expression of transcripts at the isoform-level between treatments, accounting for variability within biological replicates (Trapnell et al., 2013). The parameter *-multi-read-correct* was used to account for reads mapping to multiple locations, with the remainder of settings set to default. Transcripts were considered significantly differentially expressed between treatments if the fold change >2 and the false discovery rate (FDR) <0.05. Heatmap analysis and comparison of differentially expressed genes between different conditions were performed in R (Vienna, Austria: R Foundation for Statistical Computing) using the package gplots.

Gene ontology enrichment analysis

Differentially expressed genes of interest were queried for gene ontology (GO) term enrichment using DAVID Bioinformatics Resources 6.8 (Huang, Lempicki, & Sherman, 2009a; Huang, Lempicki, & Sherman, 2009b) with the background set to the entire *C. elegans* gene list. Each gene is assigned one or more GO terms and categorized into Biological Process, Molecular Function, and Cellular Component. Significant enrichment of GO terms was determined using a Fisher's exact test (as described in Huang, Sherman, & Lempicki, 2009a); this test associates a p-value, or EASE score, to each GO term based on the number of genes associated with that term as compared to background (Huang et al., 2009b). GO terms were considered to be significantly enriched if FDR <0.05.

Mutant generation using CRISPR/Cas9

CRISPR/Cas9 was used to generate mutations, usually deletions, in selected genes. Guide RNA (gRNA) sequences were chosen within the coding sequence of the gene of interest (GOI) using the CRISPRseek package in R to select guides with high efficacy, and CRISPR design (crispr.mit.edu) to identify possible off-target effects. Two to four gRNAs were identified and constructed for each GOI (Table 2.1). Double-stranded gRNA sequences consisted of 20 base pairs prior to the PAM site (NGG) plus overhanging base pairs on each end that overlapped with *Bsa*I-cut pRB1017 plasmid. This overlap allowed for proper ligation of the gRNA sequence into *Bsa*I-cut pRB1017 (Arribere et al., 2014).

A co-CRISPR method, described in Arribere et al., was used to facilitate detection of gene-editing events (Arribere et al., 2014). Briefly, an injection mix of 50 ng/µl *Peft:Cas9* vector (Friedland et al., 2013), 20-25 ng/µl of *dpy-10* gRNA (Arribere et al., 2014), and 20-25 ng/µl of each target gRNA-carrying plasmid were injected into young adult worm gonads (as described in Evans, 2006). F_1 Dpy worms were then moved to new plates and allowed to lay eggs. DNA was then isolated from F_1 Dpy worms and amplified with primers targeting genomic sites flanking the gRNAs of the GOI (Table 2.1). Gene-editing events were identified by differences in amplicon size as compared to *wild-type*, indicating an insertion or deletion in the gene. Worms containing mutant alleles were then sequenced to determine the mutant lesion and outcrossed twice to *wild-type* males eliminate possible off-target mutations. A summary of CRISPR/Cas9 generated alleles is shown in Figure 2.1.

Generation of expression construct strains

NEBuilder HiFi DNA Assembly (New England BioLabs) was used to assemble the vector backbone (pPD95.75), promoter and gene of interest (GOI), and fluorescent tag

(wrmScarlet). This kit assembles DNA vectors by ligating fragments with overlapping sequence using an endonuclease to create single-stranded overhangs within the overlap sequences and ligase to ligate the fragments together. In this case, the three fragments were generated via PCR using high fidelity Phusion DNA Polymerase (Thermo Fisher Scientific). Fragment 1, encoding the fluorescent protein wrmScarlet, was amplified from pSEM89_*egl-23::SL2::wrmScarlet* (Mouridi et al., 2017) using forward primer 5'- ATGGTCAGCAAGGGAGAGGCAG -3' and reverse primer 5'- TTACTTGTAGAGCTCGTCCATTCCTCC -3'. Fragment 2, the plasmid pPD95.75, which contains GFP followed by the *unc-54* 3'UTR, was amplified using forward primer 5'- GACGAGCTCTACAAGTAACATTCGTAGAATTCCAACTGAGCG -3' and reverse primer 5'- TTTTTCTACCGGTACCCTCCAAGGG -3'. This generated a linearized vector backbone that included a majority of the plasmid, excluding the GFP coding sequence. Fragment 3, which contains the GOI driven by its endogenous promoter (either 2 kb upstream of the gene or to the nearest upstream gene) and differed for each gene, was amplified with the following primers: F19B2.5 driven by the F19B2.5 promoter (pF19B2.5::F19B2.5) forward primer 5'- GGAGGGTACCGGTAGAAAAATGATTATTTCCGGCTCGGG - 3' and reverse primer 5'- CTCCCTTGCTGACCATCTGGCTGTCGTCGGCTC - 3', and K08D8.4 driven by the K08D8.4 promoter (pK08D8.4::K08D8.4) forward primer 5'- GAGGGTACCGGTAGAAAAACACCCAAGGATTTGAAG -3' and reverse primer 5'- CTCTCCCTTGCTGACCATGACCAGCATAACAAAACC -3'. The primers used to amplify the vector backbone and the promoter/GOI fragment contain appropriate overlap sequence, resulting in circular assembly of the promoter and GOI fragment ligated to the wrmScarlet fragment ligated to the vector backbone. Fragments were then gel purified using PureLink™

Quick Gel Extraction Kit (Invitrogen), followed by assembly and cloning using NEBuilder HiFi Assembly Master Mix and Cloning Kit following manufacturer's protocol.

Colonies containing possible positive constructs after cloning were tested using PCR to ensure the fragments were assembled correctly. DNA was extracted from confirmed correct colonies. Finally, 20-50 ng/ul of each construct along with 20 ng/ul *dpy-10*(+) plasmid were injected into Dpy worms. F_2 *wild-type* worms were then screened for wrmScarlet expression, and 3 independent transgene-containing lines were obtained for each GOI, with one representative line chosen for further analysis.

Gene expression analysis

Nematodes containing extrachromosomal or integrated alleles for transcriptional or translational fluorescent protein fusions were anesthetized (10mM sodium azide) for observation at 100x and 400x magnification using a Zeiss Axio Imager.ZI microscope equipped with epifluorescence and differential interference contrast (DIC) optics.

C. elegans survival assays

Treatment or control *E. coli* OP50 bacteria were cultured in liquid LB (with Ampicillin for *S. maltophilia* strains, PA14, and DB10) overnight and 100 µl of bacteria were plated onto NGM agar plates the day prior to use. Worms were bleached to synchronize and reared at 20° C on lawns of *E. coli* OP50. For survival assays, 10-12 L4 worms were transferred to each treatment plate, with three replicates of each treatment, and maintained at 25° C. Worms were transferred to new plates every day until they stopped laying eggs to separate them from their progeny. Surviving worms were recorded each day and dead worms were removed from plates, as determined by movement following prodding with a platinum wire pick. Plates that became contaminated or worms that crawled off the agar and died were removed from data analysis.

Statistical analyses were performed in R to determine differences between independent variables, including bacterial treatments and nematode genotype, with the dependent variable being the probability of nematode death on a given day. Survival probability estimates over time were determined using the Kaplan-Meier formula and used to produce survival curve graphs using Microsoft Excel. The Cox proportional hazards model was then used to compare the effects of independent variables using hazard ratios. Treatment nematode strains and bacteria were compared to *wild-type* within the same round. If contamination occurred on plates of *wild-type* nematodes, treatments were compared to all rounds of *wild-type* nematodes. If multiple rounds were performed using the same treatments, the model was modified to account for differences between rounds of experimentation.

Results

Strains of *S. maltophilia* used in this study, including K279a, JCMS, and JV3, display differing levels of pathogenicity to *C. elegans* as determined by the Cox proportional hazards test. This test utilizes the survival of *C. elegans* over time to determine the probability of a nematode dying at a given time, referred to as the hazard. Hazards of different conditions can be compared, resulting in hazard ratios. Hazard ratios greater than one indicate treatments that are more detrimental, or hazardous, to the health of *C. elegans*; whereas hazard ratios less than one indicate more beneficial conditions. *S. maltophilia* K279a, a clinical isolate of *S. maltophilia*, has previously been shown to be pathogenic to *C. elegans* (Fouhy et al., 2007); however, K279a does not appear to be pathogenic in our hands, as worms fed K279a have similar bacterial load and hazard values to worms fed the standard lab food *E. coli* OP50 (White et al., 2016) (Figure 2.2). However, *C. elegans* exposed to JCMS, a strain isolated in association with nematodes from a prairie in Kansas, are approximately five times more likely to die than *C. elegans* exposed to

OP50 (Figure 2.2). *C. elegans* exposed to JV3, another environmental isolate, are approximately 65 times more likely to die than *C. elegans* fed OP50 (Figure 2.2). We utilized these strains of *S. maltophilia* of varying pathogenicity to *C. elegans* in a transcriptomic analysis to provide a more comprehensive understanding of *C. elegans*-pathogen interactions.

To investigate transcriptomic responses to *S. maltophilia*, RNA-sequencing was performed after 12 hours of exposure to pathogenic *S. maltophilia* JCMS or JV3, or nonpathogenic *S. maltophilia* K279a or *E. coli* OP50. The 12-hour time point was chosen based on previous observations that accumulation of bacteria occurs by this time (White et al., 2016) but *S. maltophilia*-induced mortality has not yet begun (Figure 2.2). In addition, other groups have identified transcriptional changes at 4-8 hours of exposure to pathogens, including *S. aureus*, *B. thuringiensis*, and *P. aeruginosa* (Irazoqui, Troemel, et al., 2010; Troemel et al., 2006; Yang et al., 2015). Therefore, at 12 hours, pathogen recognition has begun, but transcriptional changes associated with aging and mortality, which correlate with a decreased immune response (Cabreiro & Gems, 2013), should not complicate interpretation of data.

Comparison of differentially expressed genes in *C. elegans* in response to different *S. maltophilia* strains and between pathogenic and nonpathogenic strains allowed us to address two overarching questions: 1) Are there common and strain specific responses to pathogenic *S. maltophilia* bacteria? and 2) Is the response to pathogenic *S. maltophilia* strains common to other pathogenic bacteria?

C. elegans exhibit common and strain-specific responses to *S. maltophilia*

Overall gene expression patterns were analyzed using a heatmap of genes that were significantly differentially expressed between any two treatments (Figure 2.3). Transcripts were considered differentially expressed if they had a false discovery rate (FDR)-adjusted p-value of

less than 0.05 and an absolute fold change greater than two. Gene expression profiles showed clustering of nonpathogenic (K279a and *E. coli* OP50) and pathogenic (JCMS and JV3) treatments (Figure 2.3). In fact, the expression profiles of the nonpathogenic treatments were more similar than that of the pathogenic treatments (Figure 2.3). Therefore, to identify the common response to pathogenic *S. maltophilia*, we compared differentially expressed genes in *C. elegans* between pathogenic and nonpathogenic treatments (Figure 2.4). In total, 1,296 genes were significantly differentially expressed when comparing worms fed any pathogenic (JV3 and JCMS) to any nonpathogenic (K279a and *E. coli* OP50) strain, with 11% (145) commonly differentially expressed between all pathogenic and nonpathogenic comparisons (Figure 2.4, Table 2.2). These most likely represent a core set of genes that are regulated upon exposure to pathogenic *S. maltophilia* and are therefore referred to as the "common pathogenic *S. maltophilia* response" (CPSR). Because these genes are differentially expressed in response to pathogenic vs nonpathogenic strains of the same species, this should remove general responses to *S. maltophilia* and represent genes specifically involved in pathogen response to *S. maltophilia*. Of the 145 CPSR genes, 89% (129) were upregulated in response to the pathogenic strains as compared to the nonpathogenic strains, whereas 10% (15) were downregulated (Table 2.2). One gene, *lys-10*, is upregulated in response to the pathogenic strains compared to OP50 but downregulated in response to pathogenic strains compared to K279a. Interestingly, most upregulated genes, 90 of 129, were even further upregulated in response to JV3 as compared to JCMS. Because JV3 is more virulent than JCMS, this suggests that the level of virulence influences the expression of *S. maltophilia*-induced genes (Table 2.2).

A gene ontology (GO) enrichment analysis was performed on all CPSR genes using DAVID to identify common structural components, biological processes, and molecular

functions of these genes (Table 2.3). This analysis compares GO terms associated with genes within a specified list to the entire gene list of *C. elegans* and utilizes a Fischer exact test to determine whether each GO term is enriched. From this analysis, the "biological process of innate immune response" (FDR= 4.14E-53), "biological process of defense to Gram-negative bacterium" (FDR= 2.72E-11), "molecular function of carbohydrate binding" (FDR= 1.28E-4), and "cellular component of membrane raft" (FDR= 3.8E-20), were all significantly enriched in the CPSR genes (Table 2.3).

To identify strain-specific responses, particularly JV3- and JCMS-specific responses, we identified genes that were differentially expressed in response to JV3 and JCMS as compared to all other strains. We found 31 genes differentially expressed in response to JCMS vs the nonpathogenic strains and 327 genes differentially expressed in response to JV3 vs the nonpathogenic strains (Figure 2.4). We found that 14 of the 31 JCMS vs nonpathogenic strains genes were also differentially expressed between JV3 and JCMS. These genes are specifically regulated upon exposure to *S. maltophilia* JCMS and are therefore referred to as the "JCMS-specific response" (JSR) (Table 2.2). Of the 14 JSR genes, 12 are upregulated in response to JCMS as compared to all other strains, whereas two are downregulated (Table 2.2).

We found that 221 of the 327 JV3 vs nonpathogenic strain genes were also differentially expressed between JV3 and JCMS. These genes are specifically regulated upon exposure to *S. maltophilia* JV3 and are referred to as the "JV3-specific response" (VSR) (Table 2.2). Although most CPSR genes are upregulated in response to JV3, a majority (88%) of the VSR genes are downregulated in response to JV3 as compared to the other strains (Table 2.2; Figure 2.3). This suggests that one virulence mechanism employed by JV3 may be to reduce expression of a variety of genes. GO enrichment analyses of these genes reveals enrichment of several

metabolic processes and enzymes, including “biological process of flavonoid glucuronidation” (FDR = 4.01E-8), “biological process of oxidation-reduction process” (FDR = 5.27E-5), and “molecular function of glucuronosyltransferase activity” (FDR = 0.009) (Table 2.4).

Gene networks were used to prioritize genes based on gene and protein interactions

We next wanted to determine whether the CPSR, JSR, and VSR genes are important for the response to both pathogenic *S. maltophilia* strains (CPSR genes) or to specific strains of *S. maltophilia* (JSR and VSR genes). To do this, we utilized WormNet, a probabilistic gene network model, to prioritize genes for functional analysis (Lee et al., 2010). WormNet uses both direct physical and/or genetic interactions as well as inferred interactions to create a gene network that comprises 75.4% (15,139 genes) of the *C. elegans* genome, resulting in 999,367 functional linkages (Lee et al., 2010). Previously, gene networks have been used to identify genes essential for *C. elegans* development and survival under standard conditions, as well as identification of genes associated with particular diseases (Lee et al., 2008; Özgür et al., 2008). Therefore, we hypothesize that the most connected genes within the gene network play a significant role in *S. maltophilia* response and are therefore better candidates for functional analyses.

In addition to gene network connectivity, we used two additional criteria to choose genes for functional analysis, with the goal of testing at least five genes from each category: 1) connected genes within the gene network with previously available alleles or alleles previously generated in our lab were preferentially chosen for functional analysis, 2) expression levels were used to choose additional alleles for mutant generation using CRISPR/Cas9, with higher overall expression preferred.

Of the 145 CPSR genes, 73 were connected within the gene network with an AUC of 0.6972 (p=1.8137e-16) (Figure 2.5; Table 2.5). The AUC is the area under the receiver operating characteristic (ROC) curve and provides a measure for the recovery of true-positive genes as compared to false-positive genes (Lee et al., 2010). A random network would have an AUC of 0.5, whereas a network representing perfect prediction of all connections would have an AUC of one; therefore, an AUC of 0.6972 (p=1.8137e-16) suggests relatively high predictive power of gene connections. Each connected gene is ranked based on the number of connections as well as the strength of the evidence for those connections (Lee et al., 2008), with some of the highest-ranking CPSR genes including *lys-1*, *lys-2*, *dod-22*, *dod-19*, and *clec-67* (Table 2.5). Previous studies have identified these genes as downstream effectors of defense pathways or directly involved in response to bacterial pathogen challenge (Alper et al., 2007; reviewed in Dierking, Yang, & Schulenburg, 2016; Mallo et al., 2002; White et al., 2016). These genes, along with alleles of several other genes highly connected within the network, including *F55G11.8*, *ZK6.11*, *T24B8.5*, and *scl-2*, had available mutant alleles. In addition, we had previously generated mutant alleles for *K08D8.4*, *B0024.4*, and *F08G2.5* using CRISPR/Cas9 (Figure 2.1). Of the 221 VSR genes, 103 are connected within the network (AUC=0.6698, p=5.8695e-20) (Figure 2.6; Table 2.6). Available alleles of several of the highest-ranking genes in this network, including *sodh-1*, *dhs-3*, *F13D12.6*, *pho-1*, *acdh-1*, *C55A6.7*, *dhs-2*, and *F08A8.4* were used for functional analysis.

Because of the small number of JSR genes, WormNet was not used to prioritize these genes for functional analysis. Many of these genes had very low overall expression. Therefore, genes were chosen for functional analysis if the total fragments per kilobase per million mapped reads (FPKM) for all four treatments was greater than 30. We used an available *nhr-110* allele

for functional analysis, and although we attempted to generate mutations in several other JSR genes using CRISPR/Cas9, we were successful for only one gene, *W02A2.8* (Figure 2.1).

Functional analysis of common *S. maltophilia* and strain-specific genes

Survivorship, quantified by Cox proportional hazards test, was used for functional analysis to determine whether candidate genes were important for response to control and treatment bacteria. We hypothesized that CPSR genes are important for response to both JCMS and JV3, JSR genes are important for response to JCMS, and VSR genes are important for response to JV3; therefore, mutants of these genes will result in increased or decreased susceptibility to JCMS and JV3, just JCMS, or just JV3, respectively, as compared to *wild-type*.

Mutations in four of the 12 CPSR candidate genes (*T24B8.5*, *dod-19*, *K08D8.4*, and *lys-1*) caused increased susceptibility to JCMS and mutations in three caused significant differences in survival of *C. elegans* fed JV3, with *scl-2* mutants less susceptible and *clec-67* and *T24B8.5* mutants more susceptible (Figure 2.7; Table 2.7). All of these genes, apart from *scl-2*, were previously reported to play a role in innate immune response based on GO terms. Interestingly, five of the six mutations that caused significant differences in survival upon exposure to JCMS or JV3 also showed significant differences in survival upon exposure to K279a, while one mutant *scl-2* mutants resulted in significant survival differences upon exposure to OP50 (Figure 2.7; Table 2.7). However, because worms with mutations in *scl-2* were less susceptible to both OP50 and K279a and more susceptible to JV3, these differences may be due to different responses on each bacterial treatment. One explanation for the correlation between mutants that result in susceptibility to pathogenic strains and K279a is that K279a may employ virulence mechanisms against which *C. elegans* can defend; therefore, mutations in innate immune genes cause increased susceptibility to K279a as well as pathogenic strains.

Mutations in three of the eight VSR candidate genes (*acox-1.4*, *dhs-3*, *dhs-2*) caused significant susceptibility to JV3 (Figure 2.8; Table 2.7). However, worms with mutations in all three of these alleles are also more susceptible to at least one other bacterial strain tested, suggesting that although these genes are specifically differentially expressed in response to JV3, they are also important for survival under other conditions. Although none of these genes have been reported to play a role in defense or immune response, *dhs-3* and *acox-1.4* are localized to the intestine (Nykamp, Lee, & Kimble, 2008; Steinbaugh et al., 2015) (Figure 2.10), the site of *S. maltophilia* accumulation and proposed pathogenesis (White et al., 2016). In addition, mutations in six of the eight VSR genes tested displayed differences in survival on at least one *S. maltophilia* strain. Therefore, although these genes do not seem to be important for strain-specific survival, they do seem to be important for response to *S. maltophilia* overall.

Only two JSR genes were functionally analyzed, *nhr-110* and *W02A2.8*. While *nhr-110* mutants did not display differences in survival to any *S. maltophilia* strains, mutations in *W02A2.8* resulted in worms that were more susceptible to both JCMS and JV3 (Figure 2.9; Table 2.7). However, *W02A2.8* mutants were 4.8 more times likely to die when fed JCMS and only 1.9 times more likely to die when fed JV3 as compared to *wild-type* worms, suggesting a more significant role for this gene upon JCMS infection (Table 2.7).

In addition to survival analyses, we were able to determine the expression patterns for several CPSR and VSR genes. Several transgenic strains were available from stock centers, including transcriptional reporters for *T24B8.5*, *acdh-1*, *sodh-1*, and a translational reporter for *dhs-3*. In addition, we generated translational reporters for *K08D8.4* and *F19B2.5*. All of these genes were expressed in the intestine upon exposure to *E. coli* OP50 (Figure 2.10). Localization of expression was also seen in the hypodermis (*sodh-1* and *acdh-1*), muscle (*sodh-1*), nervous

system (*sodh-1*), and head (*F19B2.5, sodh-1,* and *acdh-1*) (Figure 2.10). The intestine and hypodermis are common sites of pathogen infection, whereas the nervous system has also been shown to play a role in pathogen recognition and immune response in *C. elegans* (Kawli & Tan, 2008; Shtonda & Avery, 2006; Styer et al., 2008; Y. Zhang, 2008). Therefore, expression of differentially expressed genes in response to *S. maltophilia* correlates with common tissues involved in innate immune response.

Many genes differentially expressed in response to *S. maltophilia* are common to other pathogens

To determine whether the CPSR is common to other pathogen responses in *C. elegans* or whether these genes are unique to pathogenic *S. maltophilia*, we utilized WormExp (Yang, Dierking, Schulenburg, 2015). This database contains a variety of *C. elegans'* transcriptomic and proteomic datasets that can be queried with a list of genes to identify overlap with other datasets. WormExp uses the adjusted Fisher exact test from the program EASE to determine whether gene sets contain significant overlap (Yang, Dierking, Schulenburg, 2015). This database was used to query the CPSR genes for overlap with other microbe exposure experiments, many of which are pathogens. The CPSR genes significantly (FDR<0.05) overlapped with 105 of 212 microbe exposure experiments (Table 2.8), suggesting that these genes in general do correlate with genes regulated by other pathogens. However, to determine which individual CPSR genes in *C. elegans* are commonly regulated by pathogens and which genes are more uniquely differentially expressed in response to *S. maltophilia*, we used the WormExp database to manually determine how many pathogens regulate each individual CPSR gene, resulting in genes regulated by anywhere from 1 to 16 other pathogens (Figure 2.11; Table 2.9). Not surprisingly, many of the

CPSR genes regulated by many pathogens are also higher ranked within the CPSR gene network, validating the use of gene networks to prioritize gene importance.

Functional analysis of common pathogen genes and *S. maltophilia* specific genes

To identify genes for functional analysis, CPSR genes were ranked based on how many pathogens result in differential expression of that gene (Figure 2.11; Table 2.9). Several of the highest-ranking CPSR genes (differentially expressed in response to more than 11 other pathogens) and lowest-ranking CPSR genes (differentially expressed in response to less than 6 other pathogens) were then chosen for functional analysis (Figure 2.11; Table 2.9). High-ranking genes represent genes commonly differentially expressed in response to pathogens, and low-ranking genes represent genes more uniquely differentially expressed in response to pathogenic *S. maltophilia*. From the highest-ranked genes, mutant alleles were available for *dod-22*, *T24B8.5*, *scl-2*, *clec-67*, *lys-1*, *ZK6.11*, and *cpr-4*. From the lowest-ranking genes, alleles were generated using CRISPR/Cas9 based on expression levels, with a chosen cutoff of total FPKM greater than 30 across all four treatments. This criterion was used because genes with very low overall expression may not have as much biological significance. Successful mutant alleles were generated for *K08D8.12*, *Y82E9BR.5*, *fbxa-77*, *F08G2.5*, and *C25F9.11* (Figure 2.1).

We again used Cox proportional hazards tests of survivorship to compare strains containing mutant alleles of each candidate gene to *wild-type* worms for functional analysis. However, in addition to determining differences in survival upon exposure to *S. maltophilia* strains, we also tested several other common bacterial pathogens: *P. aeruginosa* PA14, *S. marcescens* DB10, *S. aureus* NCTC8325, and *X. nematophila* X-1462. These pathogens were chosen because they represent strains that result in differential expression of the low-ranking candidate genes (Table 2.9). We hypothesize that mutations in genes commonly differentially

expressed in response to pathogens, or the higher-ranked genes, will be important for survival upon exposure to all pathogens, whereas mutations in genes more uniquely regulated by *S. maltophilia*, or the lower-ranked genes, will only result in differences in survival when exposed to pathogenic *S. maltophilia* JCMS and JV3.

Of the common pathogen genes, mutations in *clec-67, dod-22*, and *T24B8.5* were significantly less susceptible to all non-*S. maltophilia* pathogens as compared to *wild-type* (Figure 2.12; Table 2.10). Interestingly, *T24B8.5* mutants were significantly more susceptible to all *S. maltophilia* strains (Figure 2.7; Table 2.7). In addition, *scl-2, ZK6.11*, and *cpr-4* mutants were significantly less susceptible to *S. aureus* NCTC8325 and *ZK6.11* was more susceptible to *P. aeruginosa* PA14 and *X. nematophila* X-1462 (Figure 2.12; Table 2.10). Overall, four of the six mutants resulted in differences in survival in response to at least two non-*S. maltophilia* pathogens.

Of the lower-ranking, or more *S. maltophilia*-specific genes, only two of the five mutants caused significant differences on *S. maltophilia* strains, with *fbxa-77* and *Y82E9BR.5* mutants being more susceptible to JCMS (Figure 2.13; Table 2.10). However, as expected, fewer strains showed differences on other pathogens, with *fbxa-77* and *K08D8.12* mutants causing decreased susceptibility to *S. aureus* NCTC8325 and *F08G2.5* mutants causing increased susceptibility to *P. aeruginosa* PA14 and *X. nematophila* X-1462 (Figure 2.13; Table 2.10). Therefore, although these genes result in differential expression in response to pathogenic *S. maltophilia*, mutations in only a few of these genes affect survival in response to *S. maltophilia* or other pathogens.

Discussion

Common and strain-specific genetic changes could provide insight into host responses and virulence mechanisms

This study utilized a transcriptomic approach to identify genetic responses to *S. maltophilia* strains of differing pathogenicity. We found 145 genes that were differentially expressed between all pathogenic and non-pathogenic comparisons (Figure 2.4) and represent a common response to pathogenic *S. maltophilia*. GO enrichment analysis identified processes involved in defense response, particularly response to Gram-negative bacteria, as well as a molecular function in carbohydrate binding and the cellular component of membrane raft (Table 2.3). Genes with the GO term "molecular function of carbohydrate binding" all belong to *clec* or *lec* family, which are named for their structural similarity to carbohydrate binding proteins. Although not all nematode *clec* and *lec* genes encode molecules that bind carbohydrates, many are predicted to be secreted proteins that may act as immune effectors (Dierking, Yang, & Schulenburg, 2016; Pees et al., 2016). In addition, many *clec* genes are expressed in the *C. elegans* intestine and are differentially expressed in response to pathogens; a recent review determined that 237 of 283 *clec* genes are differentially expressed during pathogen infection (Pees et al., 2016). Membrane rafts, or lipid rafts, are membrane domains that contain high concentrations of cholesterol and glycosphingolipids (Pike, 2003). Membrane rafts also serve as sites of colocalization between membrane proteins and signaling pathway components, such as components of MAPK and insulin-like signaling pathways, both of which are known to play roles in innate immunity and defense in *C. elegans* (Aballay et al., 2003; Kim et al., 2002; Murphy, Lee, & Kenyon, 2007; Pike, 2003; Sifri et al., 2003). To further support this, many of the CPSR membrane raft genes correlate with genes with biological processes in innate immune

response. Overall, this supports our hypothesis that the CPSR genes are involved in pathogen recognition and response.

A majority of the CPSR genes (89%) were up-regulated in response to JCMS and JV3. This is consistent with previous transcriptomic patterns of genes differentially expressed in *C. elegans* upon pathogen exposure (Boehnisch et al., 2011; Engelmann et al., 2011; Troemel et al., 2006). However, one study that compared transcriptomic responses to a variety of bacterial pathogens found that a majority of genes were down-regulated in response to Gram-negative bacteria and up-regulated in response to Gram-positive bacteria (Sinha et al., 2012). Therefore, this could be time- or strain-specific. In fact, when looking at genes specifically differentially expressed in response to JV3 (VSR), many genes are downregulated (88%) (Table 2.2). This supports the idea that directionality of gene expression in *C. elegans* may be strain specific, and different virulence mechanisms or host responses may play a role in these patterns. For example, GO enrichment analysis of VSR genes identified many processes and functions involved in metabolism (Table 2.4). Genes with GO terms "flavonoid glucuronidation, "flavonoid metabolic process", and "UPD-glycosyltransferase" consist almost entirely of glycosyltransferase family proteins. This is a large family of proteins in *C. elegans*, comprising 265 genes (Yonekura-Sakakibara & Hanada, 2011). However, a direct linkage between glycosyltransferases and innate immunity has not yet been observed. Genes with GO terms "oxidation-reduction process" and "oxidoreductase activity" encoded many dehydrogenase and oxidase enzymes. Interestingly, mitochondria, the location of many dehydrogenases, has been shown to be involved in pathogen recognition (reviewed in Sancho, Enamorado, & Garaude, 2017). Specifically, $FADH_2$-dependent dehydrogenase activity in mice macrophages increases upon exposure to *E. coli* (Garaude et al., 2016). Mutations in all three JV3-specific genes that caused significant

susceptibility to JV3 (*dhs-2*, *dhs-3*, and *acox-1.4*) were enzymes involved in oxidoreductase activity (Figure 2.8; Table 2.5). Furthermore, many genes involved in oxidoreductase activity and flavanoid metabolic processes are also differentially expressed upon exposure to other pathogens, such as *B. thuringiensis*, *S marcescens*, and *X. nematophila* (Engelmann et al., 2011; Sinha et al., 2012; Zhang et al., 2015). As many of these enzymes are involved in general metabolic processes, one explanation might be that JV3 utilizes a mechanism to decrease overall metabolism in *C. elegans*, resulting in its own increased virulence. On the other hand, this downregulation may be a defense mechanism used by *C. elegans* to decrease metabolites needed for JV3 survival or pathogenesis. Therefore, further analyses of these genes in *C. elegans* are needed to fully understand their role in response to *S. maltophilia* JV3 and other pathogens. These analyses, along with JV3 genome sequence analysis, could provide insight into unique virulence mechanisms employed by JV3.

In comparison to VSR genes (221), there are very few genes (14) specifically differentially expressed in response to JCMS (JSR) (Table 2.2). In addition, most of these genes have very low overall expression. Overall, this suggests that JCMS does not have unique virulence mechanisms as compared to JV3 and therefore does not lead to unique host responses. In order to further understand virulence mechanisms of different strains of *S. maltophilia*, comparative analysis of genomes is necessary. However, it is clear that there are strain-specific differences in *S. maltophilia* that drive differences in both phenotypic and genotypic responses in *C. elegans*.

Overall, data from the survival analyses do not support our hypothesis that CPSR genes are necessary for survival on JCMS and JV3, while JSR and VSR genes are necessary for survival on only JCMS or JV3, respectively, as mutations in a majority of genes did not affect

survival of *C. elegans* in a strain-specific manner (Table 2.7). However, it appears that overall, these genes do play a role in response to *S. maltophilia*, as 13 of 22 candidate CPSR, JSR, and VSR gene mutants display significant differences in survival upon exposure to at least one *S. maltophilia* strain. One explanation for lack of observed survival differences among the mutations in these genes as compared to *wild-type* is functional redundancy among families of similar proteins. In fact, many of these genes, including C-type lectins and lysozymes, all belong to large gene families with structural similarity (Pees et al., 2016; Schulenburg & Boehnisch, 2008). Therefore, even though these genes may play a role in innate immune response, mutations in one of these genes alone may not result in an effect on survival upon pathogen-challenge.

Mutations in many CPSR genes that result in significant susceptibility to JCMS and JV3 also show significant differences in survival upon exposure to K279a (Table 2.7). As K279a has previously been shown to be virulent to *C. elegans* (Fouhy et al., 2007) and genome sequencing has identified a variety of virulence factors in *S. maltophilia* K279a (Crossman et al., 2008), it may in fact be utilizing virulence mechanisms. Whereas *wild-type C. elegans* may be able to defend against K279a infection, mutations in innate immune and defense genes may cause *C. elegans* to become more susceptible to K279a.

Many genes involved in response to *S. maltophilia* are common responses to other pathogens

In addition to identifying strain specific responses, we were able to identify genes that were commonly differentially expressed in response to other pathogens as well as genes uniquely differentially expressed upon *S. maltophilia* infection. WormExp, a recently developed database, provides a resource for querying and comparing a variety of transcriptomic datasets. This database allowed us to identify *S. maltophilia*-regulated genes that were also regulated by a

variety of other pathogens, as well as responses that were specific to pathogenic *S. maltophilia* strains. WormExp contains data for transcriptomic responses of *C. elegans* to 28 bacterial strains, of which our CPSR genes were regulated by a maximum of 16 other pathogens (Figure 2.11; Table 2.9). 89% of the 145 CPSR genes are differentially expressed in response to at least 5 other pathogens, suggesting that a majority of these genes are commonly regulated by other pathogens in *C. elegans* (Figure 2.11; Table 2.9). In addition, many genes that are commonly regulated by other pathogens are highly connected in the gene network of CPSR genes and are associated with GO terms involved in defense response, including *dod-22, clec-67, cpr-3,* ZK6.11, and *lys-1*. Therefore, this study further validates the use of gene networks for identification of important genes, in this case genes involved in response to pathogen infection.

Functional analysis of genes differentially expressed by many pathogens revealed that a majority of these genes are involved in response to both pathogenic *S. maltophilia* and other pathogens (Figure 2.7; Figure 2.12; Table 2.7; Table 2.10). However, genes that are more uniquely differentially expressed in response to *S. maltophilia* do not seem important for survival on pathogenic *S. maltophilia* or other pathogens (Figure 2.13; Table 2.10). Interestingly, we observed the most differences in survival in response to *S. aureus*. In addition, several genes that seem to be important for survival on *S. aureus* are up-regulated in response to *S. maltophilia,* but down-regulated in response to *S. aureus*. *S. aureus* is the only Gram-positive bacteria used in this experiment, suggesting differences in the responses to Gram-positive and Gram-negative bacteria.

Using transcriptomic data to understand and analyze responses to pathogens can provide insight into overall response patterns and pathogen virulence mechanisms. This study illustrates the previously supported idea that there are common signatures of pathogen infection in *C.*

elegans, but also unique species and even strain specific responses. Therefore, to fully understand virulence of bacteria and pathogenesis in *C. elegans* and begin to expand these finding to other animals, a variety of bacterial species and strains need to be investigated.

References

Aballay, A., Drenkard, E., Hilbun, L. R., & Ausubel, F. M. (2003). *Caenorhabditis elegans* innate immune response triggered by *Salmonella enterica* requires intact LPS and is mediated by a MAPK signaling pathway. *Current Biology, 13*(1), 47–52. https://doi.org/10.1016/S0960-9822(02)01396-9

Adamek, M., Linke, B., & Schwartz, T. (2014). Virulence genes in clinical and environmental *Stenotrophomonas maltophilia* isolates: A genome sequencing and gene expression approach. *Microbial Pathogenesis, 67–68*, 20–30. https://doi.org/10.1016/j.micpath.2014.02.001

Adamek, M., Overhage, J., Bathe, S., Winter, J., Fischer, R., & Adamek, M. (2011). Genotyping of environmental and clinical *Stenotrophomonas maltophilia* isolates and their pathogenic potential. *PLoS ONE, 6*(11), e27615. https://doi.org/10.1371/journal.pone.0027615

Alper, S., Mcbride, S. J., Lackford, B., Freedman, J. H., & Schwartz, D. A. (2007). Specificity and complexity of the *Caenorhabditis elegans* innate immune response. *Molecular and Cellular Biology, 27*(15), 5544–5553. https://doi.org/10.1128/MCB.02070-06

Arribere, J. A., Bell, R. T., Fu, B. X. H., Artiles, K. L., Hartman, P. S., & Fire, A. Z. (2014). Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in *Caenorhabditis elegans*. *Genetics, 198*(3), 837–846. https://doi.org/10.1534/genetics.114.169730

Berg, G., Roskot, N., & Smalla, K. (1999). Genotypic and phenotypic relationships between clinical and environmental isolates of *Stenotrophomonas maltophilia. Journal of Clinical Microbiology, 37*(11), 3594–600.

Berg, M., Stenuit, B., Ho, J., Wang, A., Parke, C., Knight, M., Alvarez-Cohen, L., & Shapira, M. (2016). Assembly of the *Caenorhabditis elegans* gut microbiota from diverse soil microbial environments. *The ISME Journal, 10*, 1998–2009. https://doi.org/10.1038/ismej.2015.253

Boehnisch, C., Wong, D., Habig, M., Isermann, K., Michiels, N. K., Roeder, T., May, R. C., & Schulenburg, H. (2011). Protist-type lysozymes of the nematode *Caenorhabditis elegans* contribute to resistance against pathogenic *Bacillus thuringiensis. PLoS ONE, 6*(9), e24619. https://doi.org/10.1371/journal.pone.0024619

Brooke, J. S. (2012). *Stenotrophomonas maltophilia*: An emerging global opportunistic pathogen. *Clinical Microbiology Reviews, 25*(1), 2–41. https://doi.org/10.1128/CMR.00019-11

Cabreiro, F., & Gems, D. (2013). Worms need microbes too: Microbiota, health and aging in *Caenorhabditis elegans. EMBO Molecular Medicine, 5*, 1300–1310. https://doi.org/10.1002/emmm.201100972

Chang, Y.-T., Lin, C.-Y., Chen, Y.-H., & Hsueh, P.-R. (2015). Update on infections caused by *Stenotrophomonas maltophilia* with particular attention to resistance mechanisms and therapeutic options. *Frontiers in Microbiology, 6*(893), 1–20. https://doi.org/10.3389/fmicb.2015.00893

Crossman, L. C., Gould, V. C., Dow, J. M., Vernikos, G. S., Okazaki, A., Sebaihia, M., Saunders, D., Arrowsmith, C., Carver, T., Peters, N., Adlem, E., Kerhornou, A., Lord, A., Murphy, L., Seeger, K., Squares, R., Rutter, S., Quail, M., Rajandream, M., Harris, D., Churcher, C., Bentley, S. D., Parkhill, J., Thomson, N. R., & Avison, M. B. (2008). The complete genome, comparative and functional analysis of *Stenotrophomonas maltophilia* reveals an organism heavily shielded. *Genome Biology*, *9*(4), R74.1-R74.13. https://doi.org/10.1186/gb-2008-9-4-r74

Dierking, K., Yang, W., & Schulenburg, H. (2016). Antimicrobial effectors in the nematode *Caenorhabditis elegans*: an outgroup to the Arthropoda. *Philosophical Transactions of the Royal Society B, 371*(1695), 20150299. https://doi.org/10.1098/rstb.2015.0299

Dirksen, P., Marsh, S. A., Braker, I., Heitland, N., Wagner, S., Nakad, R., Mader, S., Petersen, C., Kowallik, V., Rosenstiel, P., Felix, M-A., & Schulenburg, H. (2016). The native microbiome of the nematode *Caenorhabditis elegans*: gateway to a new host-microbiome model. *BMC Biology, 14*(38). https://doi.org/10.1186/s12915-016-0258-1

Elting, L. S., & Bodey, G. P. (1990). Septicemia due to *Xanthomonas* species and non-aeruginosa *Pseudomonas* species: Increasing incidence of catheter-related infections. *Medicine (Baltimore), 69*(5), 296–306.

Engelmann, I., Griffon, A., Tichit, L., Montanana-Sanchis, F., Wang, G., Reinke, V., Waterston, R. H., Hillier, L. W., & Ewbank, J. J. (2011). A comprehensive analysis of gene expression changes provoked by bacterial and fungal infection in *C. elegans. PLoS ONE, 6*(5), e19055. https://doi.org/10.1371/journal.pone.0019055

Evans, T. C., ed. Transformation and microinjection (April 6, 2006), WormBook, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/ wormbook.1.108.1, http://www.wormbook.org.

Ferrer-Navarro, M., Planell, R., Yero, D., Mongiardini, E., Torrent, G., Huedo, P., Martinez., P., Roher, N., Mackenzie, S., Gibert, I., & Daura, X. (2013). Abundance of the quorum-sensing factor Ax21 in four strains of *Stenotrophomonas maltophilia* correlates with mortality rate in a new zebrafish model of infection. *PLoS ONE, 8*(6), e67207. https://doi.org/10.1371/journal.pone.0067207

Fouhy, Y., Scanlon, K., Schouest, K., Spillane, C., Crossman, L., Avison, M. B., Ryan, R. P., & Dow, J. M. (2007). Diffusible signal factor-dependent cell-cell signaling and virulence in the nosocomial pathogen *Stenotrophomonas maltophilia. Journal of Bacteriology, 189*(13), 4964–4968. https://doi.org/10.1128/JB.00310-07

Friedland, A. E., Tzur, Y. B., Esvelt, K. M., Colaiacovo, M. P., Church, G. M., & Calarco, J. A. (2013). Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nature Methods, 10*(8), 741–743. https://doi.org/10.1038/nmeth.2532.Heritable

Garaude, J., Acín-Pérez, R., Martínez-Cano, S., Enamorado, M., Ugolini, M., Nistal-Villán, E., Hervas-Stubbs, S., Pelegrin, P., Enriquez, J. A., & Sancho, D. (2016). Mitochondrial respiratory-chain adaptations in macrophages contribute to antibacterial host defense. *Nature Immunology, 17*(9), 1037–1045. https://doi.org/10.1038/ni.3509

Garigan, D., Hsu, A., Fraser, A. G., Kamath, R. S., Ahringer, J., & Kenyon, C. (2002). Genetic analysis of tissue aging in *Caenorhabditis elegans*: A role for heat-shock factor and bacterial proliferation. *Genetics, 161*, 1101–1112.

Huang, D. W., Lempicki, R. A., & Sherman, B. T. (2009a). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols, 4*(1), 44–57. https://doi.org/10.1038/nprot.2008.211

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009b). Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research, 37*(1), 1–13. https://doi.org/10.1093/nar/gkn923

Irazoqui, J. E., Troemel, E. R., Feinbaum, R. L., Luhachack, L. G., Cezairliyan, B. O., & Ausubel, F. M. (2010). Distinct pathogenesis and host responses during infection of *C. elegans* by *P. aeruginosa* and *S. aureus*. *PLoS Pathogens, 6*(7), e1000982. https://doi.org/10.1371/journal.ppat.1000982

Jägevall, S., Rabe, L., & Pedersen, K. (2011). Abundance and diversity of biofilms in natural and artificial aquifers of the Äspö Hard Rock Laboratory, Sweden. *Microbial Ecology, 61*(2), 410–422. https://doi.org/10.1007/s00248-010-9761-z

Kaiser, S., Biehler, K., & Jonas, D. (2009). A *Stenotrophomonas maltophilia* multilocus sequence typing scheme for inferring population structured. *Journal of Bacteriology, 191*(9), 2934–2943. https://doi.org/10.1128/JB.00892-08

Kawli, T., & Tan, M. (2008). Neuroendocrine signals modulate the innate immunity of *Caenorhabditis elegans* through insulin signaling. *Nature Immunology, 9*(12), 1415–1424.

Kim, D. H., Feinbaum, R., Alloing, G., Emerson, F. E., Garsin, D. A., Inoue, H., Tanaka-Hino, M., Hisamoto, N., Matsumoto, K., Tan, M-W., & Ausubel, F. M. (2002). A conserved p38 MAP Kinase pathway in *Caenorhabditis elegans* innate immunity. *Science, 297*, 623–626.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology, 14*(4), R36. https://doi.org/10.1186/gb-2013-14-4-r36

Lee, I., Lehner, B., Crombie, C., Wong, W., Fraser, A. G., & Marcotte, E. M. (2008). A single gene network accurately predicts phenotypic effects of gene perturbation in *Caenorhabditis elegans. Nature Genetics, 40*(2), 181–8. https://doi.org/10.1038/ng.2007.70

Lee, I., Lehner, B., Vavouri, T., Shin, J., Fraser, A. G., & Marcotte, E. M. (2010). Predicting genetic modifier loci using functional gene networks. *Gemome Research, 20*(8), 1143–1153. https://doi.org/10.1101/gr.102749.109.et

Lira, F., Berg, G., & Martínez, J. L. (2017). Double-face meets the bacterial world: The opportunistic pathogen *Stenotrophomonas maltophilia*. *Frontiers in Microbiology, 8*, 2190. https://doi.org/10.3389/fmicb.2017.02190

Mallo, G. V., Kurz, C. L., Couillault, C., Pujol, N., Granjeaud, S., Kohara, Y., & Ewbank, J. J. (2002). Inducible antibacterial defense system in *C. elegans. Current Biology, 12*(14), 1209–1214. https://doi.org/10.1016/S0960-9822(02)00928-4

McGhee, J. D., Sleumer, M. C., Bilenky, M., Wong, K., McKay, S. J., Goszczynski, B., Tian, H., Krich, N. D., Khattra, J., Holt, R. A., Baillie, D. L., Kohara, Y., Marra, M A., Jones, S. J.M., Moerman, D. G., & Robertson, A. G. (2007). The ELT-2 GATA-factor and the global regulation of transcription in the *C. elegans* intestine. *Developmental Biology, 302*(2), 627–645. https://doi.org/10.1016/j.ydbio.2006.10.024

Mouridi, S. El, Lecroisey, C., Tardy, P., Mercier, M., Leclercq-Blondel, A., Zariohi, N., & Boulin, T. (2017). Reliable CRISPR/Cas9 genome engineering in *Caenorhabditis elegans* using a single efficient sgRNA and an easily recognizable phenotype. *G3, 7*(5), 1429–1437. https://doi.org/10.1534/g3.117.040824/-/DC1.1

Murphy, C. T., Lee, S., & Kenyon, C. (2007). Tissue entrainment by feedback regulation of insulin gene expression in the endoderm of *Caenorhabditis elegans. Proceedings of the National Academy of Sciences, 104*(48), 19046–19050.

Murphy, C. T., Mccarroll, S. A., Bargmann, C. I., Fraser, A., Kamath, R. S., Ahringer, J., Li, H., & Kenyon, C. (2003). Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans. Nature, 424*, 277–284.

Nykamp, K., Lee, M-H., & Kimble, J. (2008). *C. elegans* La-related protein, LARP-1, localizes to germline P bodies and attenuates Ras-MAPK signaling during oogenesis. *RNA, 14*(7), 1378–1389. https://doi.org/10.1261/rna.1066008

Özgür, A., Vu, T., Erkan, G., & Radev, D. R. (2008). Identifying gene-disease associations using centrality on a literature mined gene-interaction network. *Bioinformatics, 24*(13), 277–285. https://doi.org/10.1093/bioinformatics/btn182

Pees, B., Yang, W., Zárate-potes, A., Schulenburg, H., & Dierking, K. (2016). High innate immune specificity through diversified C-type lectin-like domain proteins in invertebrates. *Journal of Innate Immunity, 8*, 129–142. https://doi.org/10.1159/000441475

Pike, L. J. (2003). Lipid rafts: bringing order to chaos. *Journal of Lipid Research, 44*(4), 655–667. https://doi.org/10.1194/jlr.R200021-JLR200

Rocco, F., De Gregorio, E., Colonna, B., & Di Nocera, P. P. (2009). *Stenotrophomonas maltophilia* genomes: A start-up comparison. *International Journal of Medical Microbiology, 299*(8), 535–546. https://doi.org/10.1016/j.ijmm.2009.05.004

Samuel, B. S., Rowedder, H., Braendle, C., Félix, M.-A., & Ruvkun, G. (2016). *Caenorhabditis elegans* responses to bacteria from its natural habitats. *Proceedings of the National Academy of Sciences, 113*(27), E3941–E3949. https://doi.org/10.1073/pnas.1607183113

Sancho, D., Enamorado, M., & Garaude, J. (2017). Innate immune function of mitochondrial metabolism. *Frontiers in Immunology, 8*, 1–9. https://doi.org/10.3389/fimmu.2017.00527

Schulenburg, H., & Boehnisch, C. (2008). Diversification and adaptive sequence evolution of *Caenorhabditis* lysozymes (Nematoda: Rhabditidae). *BMC Evolutionary Biology, 8*(114). https://doi.org/10.1186/1471-2148-8-114

Shtonda, B. B., & Avery, L. (2006). Dietary choice behavior in *Caenorhabditis elegans. The Journal of Experimental Biology, 209*, 89–102. https://doi.org/10.1242/jeb.01955

Sifri, C. D., Begun, J., Ausubel, F. M., & Calderwood, S. B. (2003). *Caenorhabditis elegans* as a model host for *Staphylococcus aureus* pathogenesis. *Infection and Immunity, 71*(4), 2208–2217. https://doi.org/10.1128/IAI.71.4.2208

Sinha, A., Rae, R., Iatsenko, I., & Sommer, R. J. (2012). System wide analysis of the evolution of innate immunity in the nematode model species *Caenorhabditis elegans* and *Pristionchus pacificus. PLoS ONE, 7*(9). https://doi.org/10.1371/journal.pone.0044255

Steinbaugh, M.J., Narasimhan, S. D., Robida-Stubbs, S., Moronetti Mazzeo, L. E., Dreyfuss, J. M., Hourihan, J. M., Raghavan, P., Operaña, T. N., Esmaillie, R., & Blackwell, T. K. (2005). Lipid-mediated regulation of SKN-1/Nrf in response to germ cell absence. *Elife, 4*(e07836). https://doi.org/10.7554/eLife.07836

Styer, K. L., Singh, V., Macosko, E., Steele, S. E., Bargmann, C. I., & Aballay, A. (2008). Innate immunity in *Caenorhabditis elegans* is regulated by neurons expressing NPR-1/GPCR. *Science, 322*(5900), 460–465. https://doi.org/10.1126/science.1163673

Trapnell, C., Hendrickson, D. G., Sauvageau, M., Goff, L., Rinn, J. L., & Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature Biotechnology, 31*(1), 46–53. https://doi.org/10.1038/nbt.2450

Troemel, E. R., Chu, S. W., Reinke, V., Lee, S. S., Ausubel, F. M., & Kim, D. H. (2006). p38 MAPK regulates expression of immune response genes and contributes to longevity in *C. elegans. PLoS Genetics, 2*(11), e183. https://doi.org/10.1371/journal.pgen.0020183

Turrientes, M. C., Baquero, M. R., Sánchez, M. B., Valdezate, S., Escudero, E., Berg, G., Canton, R., Baquero, F., Galan, J. C., & Martínez, J. L. (2010). Polymorphic mutation

frequencies of clinical and environmental *Stenotrophomonas maltophilia* populations. *Applied and Environmental Microbiology*, *76*(6), 1746–1758. https://doi.org/10.1128/AEM.02817-09

White, C. V, Darby, B. J., Bredden, R. J., & Herman, M. A. (2016). A *Stenotrophomonas maltophilia* strain evades a major *Caenorhabditis elegans* defense pathway. *Infection and Immunity*, *84*(2), 524–536.

Yang, W., Dierking, K., Esser, D., Tholey, A., Leippe, M., Rosenstiel, P., & Schulenburg, H. (2015). Overlapping and unique signatures in the proteomic and transcriptomic responses of the nematode *Caenorhabditis elegans* toward pathogenic *Bacillus thuringiensis. Developmental and Comparative Immunology*, *51*, 1–9. https://doi.org/10.1016/j.dci.2015.02.010

Yang, W., Dierking, K., & Schulenburg, H. (2015). WormExp: a web-based application for *Caenorhabditis elegans*-specific gene expression enrichment analysis. *Bioinformatics*, *32(*6): 943-945.

Yonekura-Sakakibara, K., & Hanada, K. (2011). An evolutionary view of functional diversity in family 1 glycosyltransferases. *The Plant Journal*, *66*(1), 182–193. https://doi.org/10.1111/j.1365-313X.2011.04493.x

Zhang, F., Berg, M., Dierking, K., Félix, M., Shapira, M., Samuel, B. S., & Schulenburg, H. (2017). *Caenorhabditis elegans* as a model for microbiome research. *Frontiers in Microbiology*, *8*(485), 1–10. https://doi.org/10.3389/fmicb.2017.00485

Zhang, W., Yu, Y., Hertwig, F., Thierry-Mieg, J., Zhang, W., Thierry-Mieg, D., … Fischer, M. (2015). Comparison of RNA-seq and microarray-based models for clinical endpoint prediction. *Genome Biology*, *16*, 133. https://doi.org/10.1186/s13059-015-0694-1

Zhang, Y. (2008). Neuronal mechanisms of *Caenorhabditis elegans* and pathogenic bacteria interactions. *Current Opinion in Microbiology*, *11*(3), 257–261. https://doi.org/10.1016/j.mib.2008.04.003

Figures

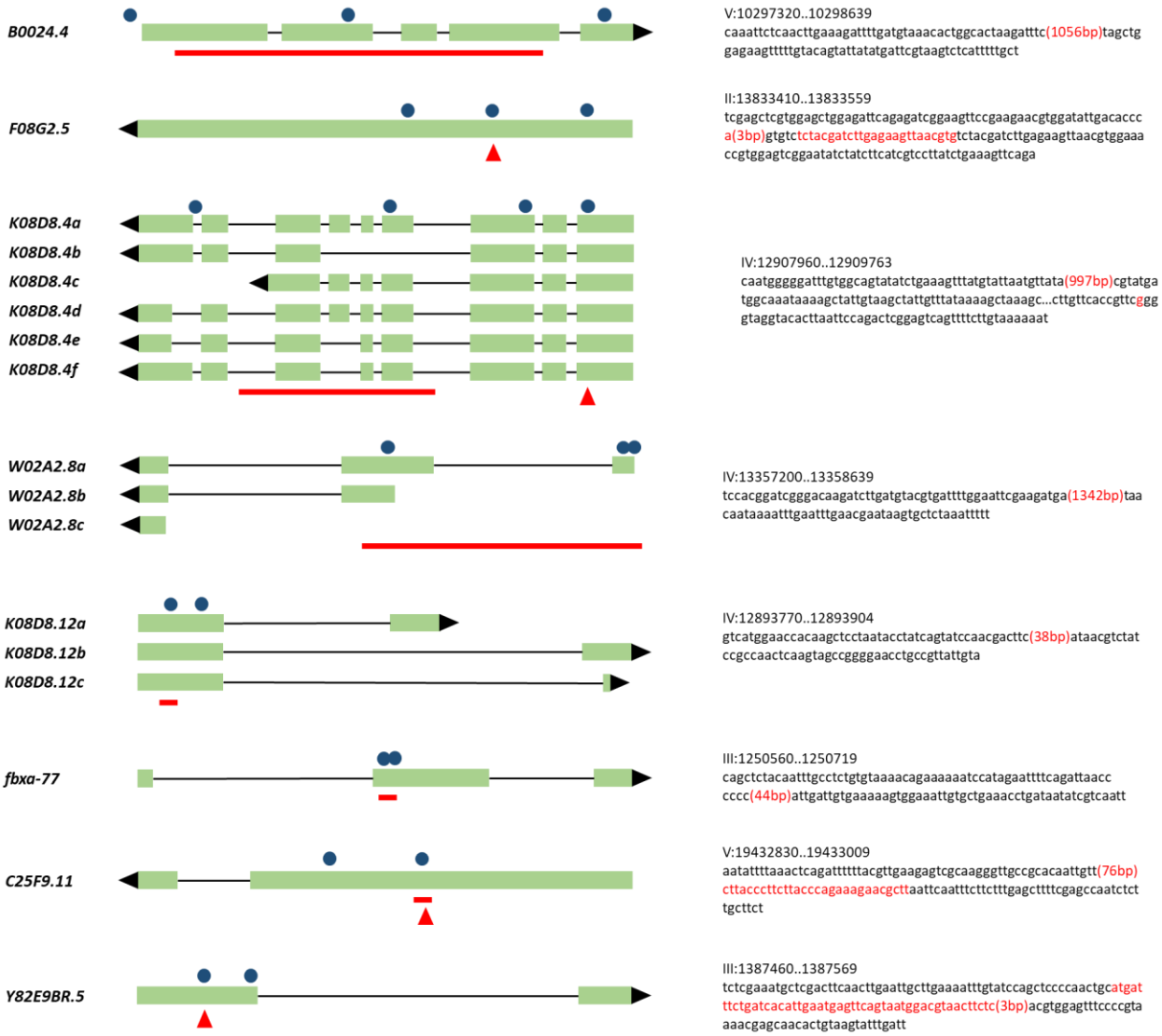

Figure 2.1 CRISPR/Cas9 generated alleles

Genes with generated CRISPR/Cas9 mutation alleles are shown above. Exons are indicated by green boxes and introns are indicated by black lines. Gene sizes are not to scale, but exon/intron size within genes is to scale. Relative location of gRNAs is indicated by circles above the gene, and location of mutation is indicated in red (lines for deletions, triangles for insertions). All isoforms of *W02A2.8*, *K08D8.4*, and *K08D8.12* are shown. *K08D8.4* and *K08D8.12* mutations are predicted to result in loss of function of all isoforms; *W02A2.8c* may be expressed but is not differentially expressed between treatments. Mutation sequence and flanking sequence is shown on the right, with mutation sequence shown in red font (number in parentheses represents size of deletion).

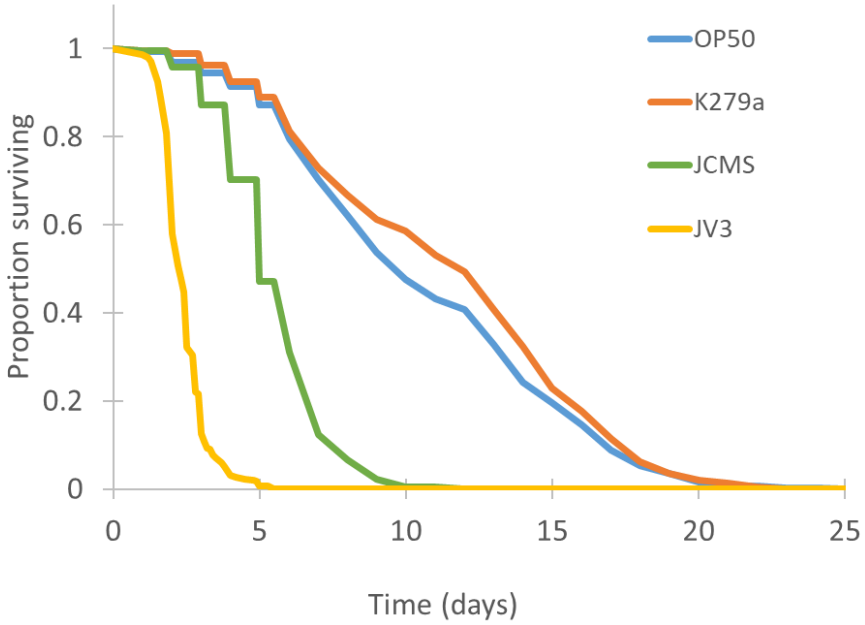

Figure 2.2 *S. maltophilia* strains show varying pathogenicity to *C. elegans*

Survivorship of *wild-type* nematodes on *S. maltophilia* JCMS, K279a, JV3, and *E. coli* OP50. Survival estimates were determined using Kaplan-Meier estimates generated in R. This data contains all *wild-type* data collected from experiments in this study, representing 21 individual experiments and n= 428-546 for each bacterial treatment. Sample sizes, hazard ratios and p-values generated form Cox proportional hazards tests are shown in Table 2.5.

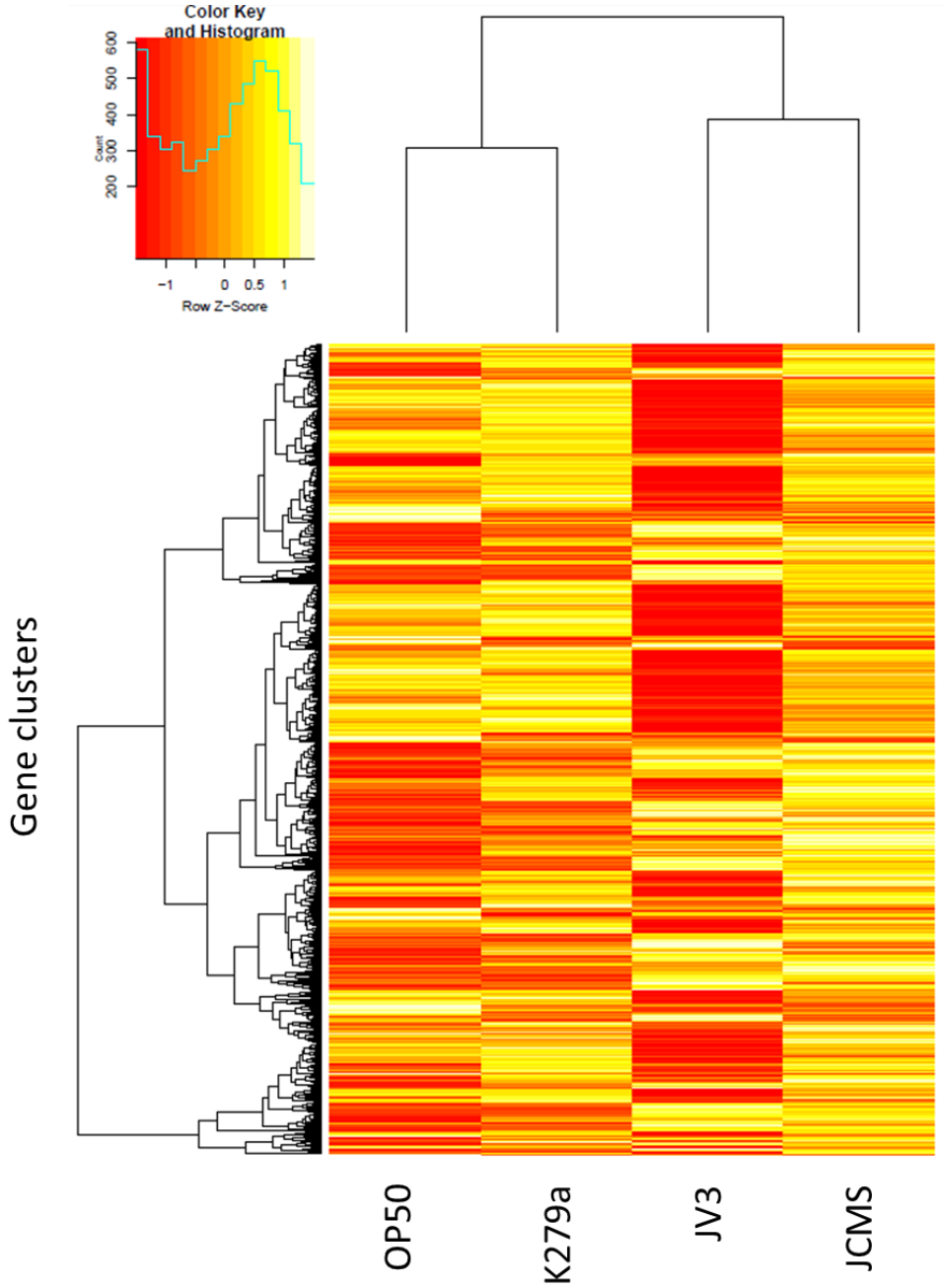

Figure 2.3 Heatmap of *C. elegans* differentially expressed genes in response to *S. maltophilia*

Differentially expressed genes from RNA-sequencing include genes with fold-change >2 and FDR-adjusted p-value <0.05 between any treatment comparisons. Fragments per kilobase per million (FPKM) values for each gene and treatment were log transformed. Vertical distances on dendrogram represent degree of similarity of gene expression profiles between treatments. Gene expression is color coded, with red indicating lower expression and yellow indicating higher expression. Heatmap was generated and visualized using the gplots package in R.

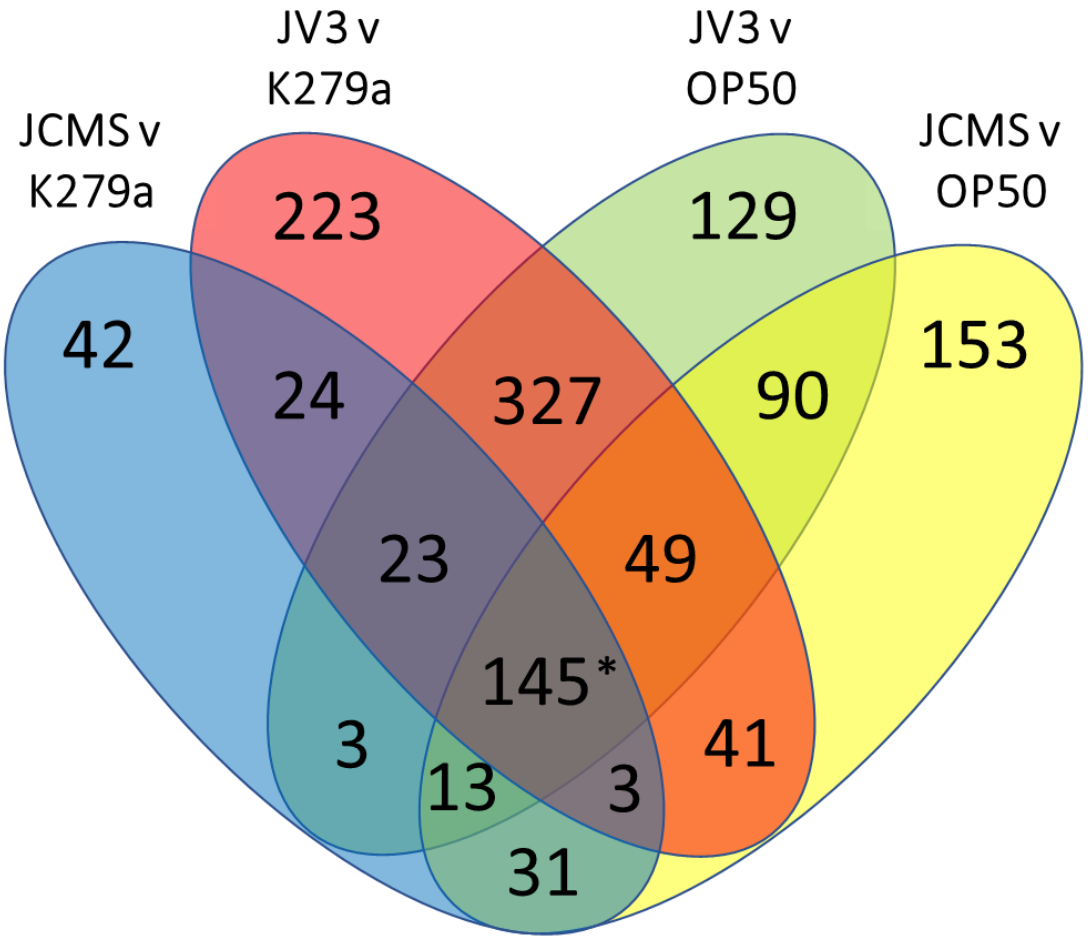

Figure 2.4 *C. elegans* expresses a common set of 145 genes in response to pathogenic *S. maltophila* strains

Differential expression was determined between each pathogenic and nonpathogenic comparison, with the number of significantly differentially expressed genes indicated between each set of comparisons. Genes included are differentially expressed between the specified treatments with fold-change >2 and FDR-adjusted p-value <0.05. 145 genes were commonly differentially expressed between all pathogenic (JCMS and JV3) and nonpathogenic (K279a and OP50) treatments, or the common pathogenic *S. maltophilia* response (CPSR), indicated by the asterisk.

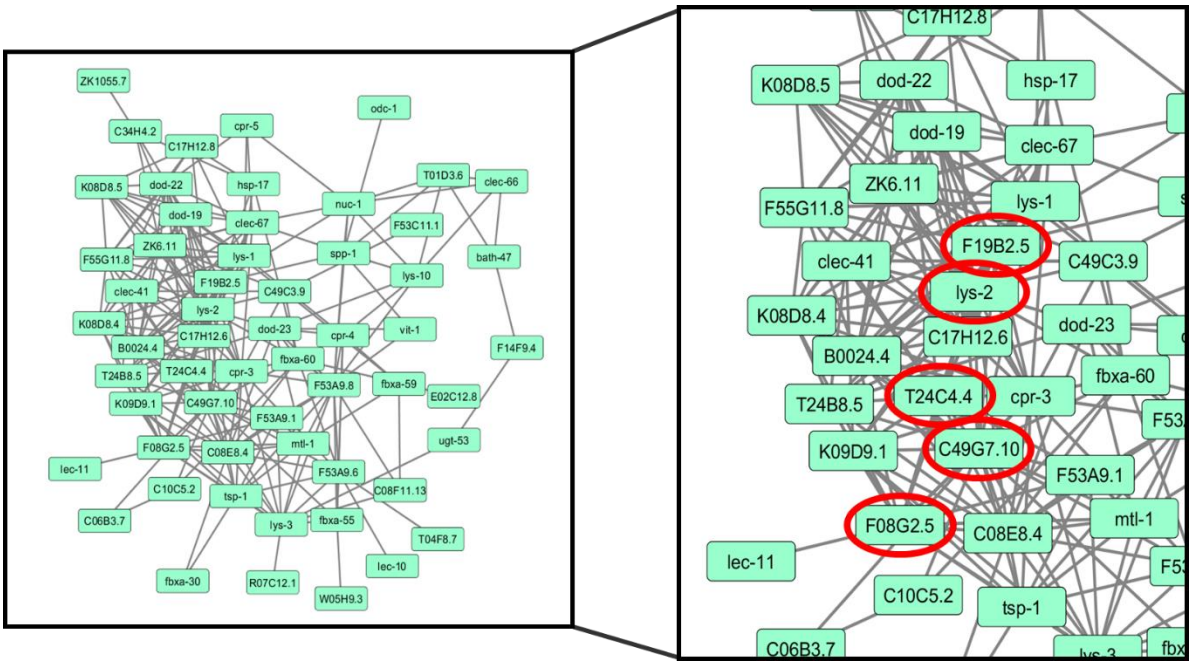

Figure 2.5 Gene network analysis was used to prioritize CPSR genes

WormNet v2, a probabilistic functional gene network model, was queried with the 145 genes that were differentially expressed in response to non-pathogenic vs. pathogenic strains. 73 of the 145 genes are connected to one another (AUC=0.6942, p=1.8137e-16), with the 5 highest-ranking genes circled in the inset. Network was visualized using Cytoscape 3.5.1. Green boxes represent individual genes and grey lines represent functional connections between genes.

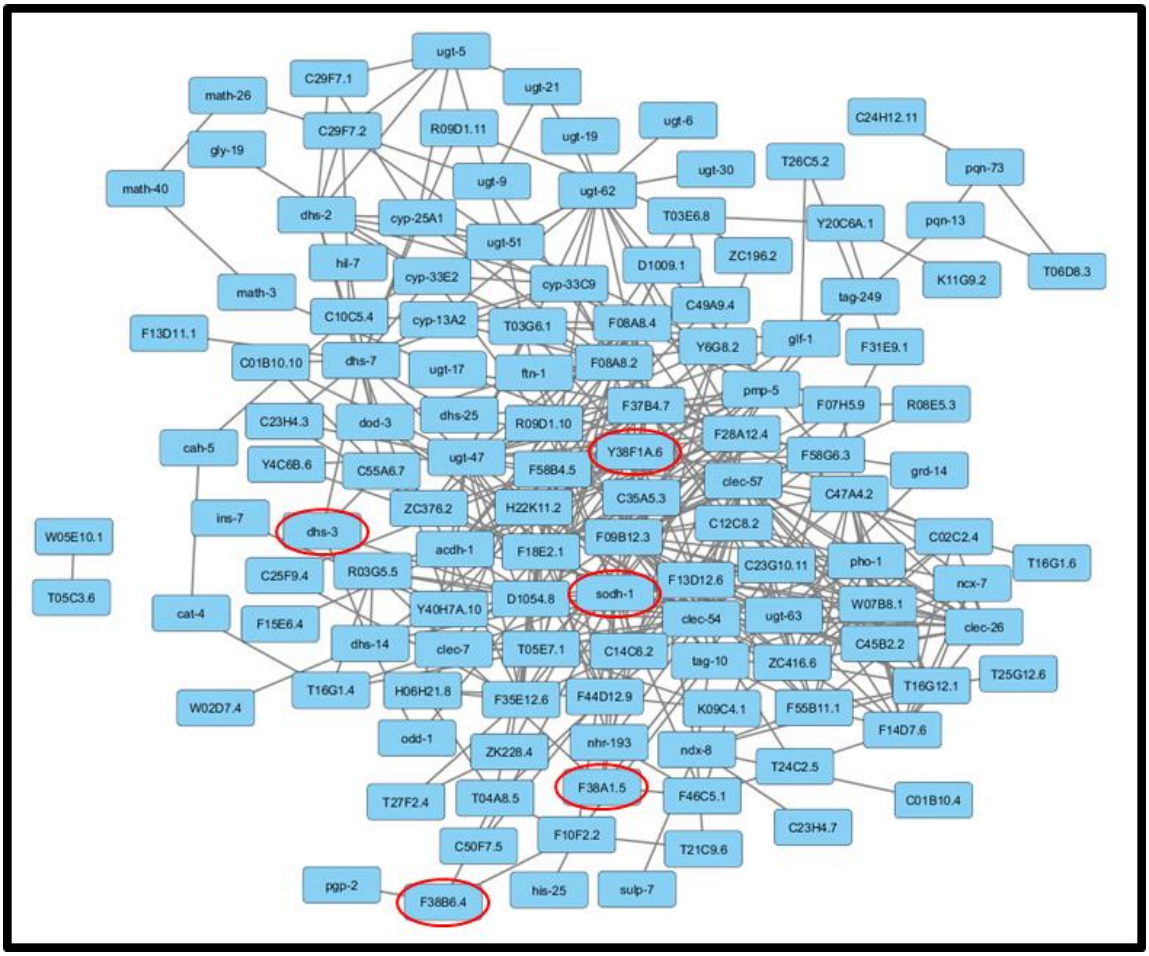

Figure 2.6 Gene network analysis was used to prioritize VSR genes

WormNet v2, a probabilistic functional gene network model was queried with the 221 JV3-specific response (VSR) genes. 103 of the 221 genes are connected to one another (AUC=0.6687, *p*=5.8695e-20), with the 5 highest-ranking genes circled. Network was viewed using Cytoscape 3.5.1. Blue boxes represent individual genes and grey lines represent functional connections between genes.

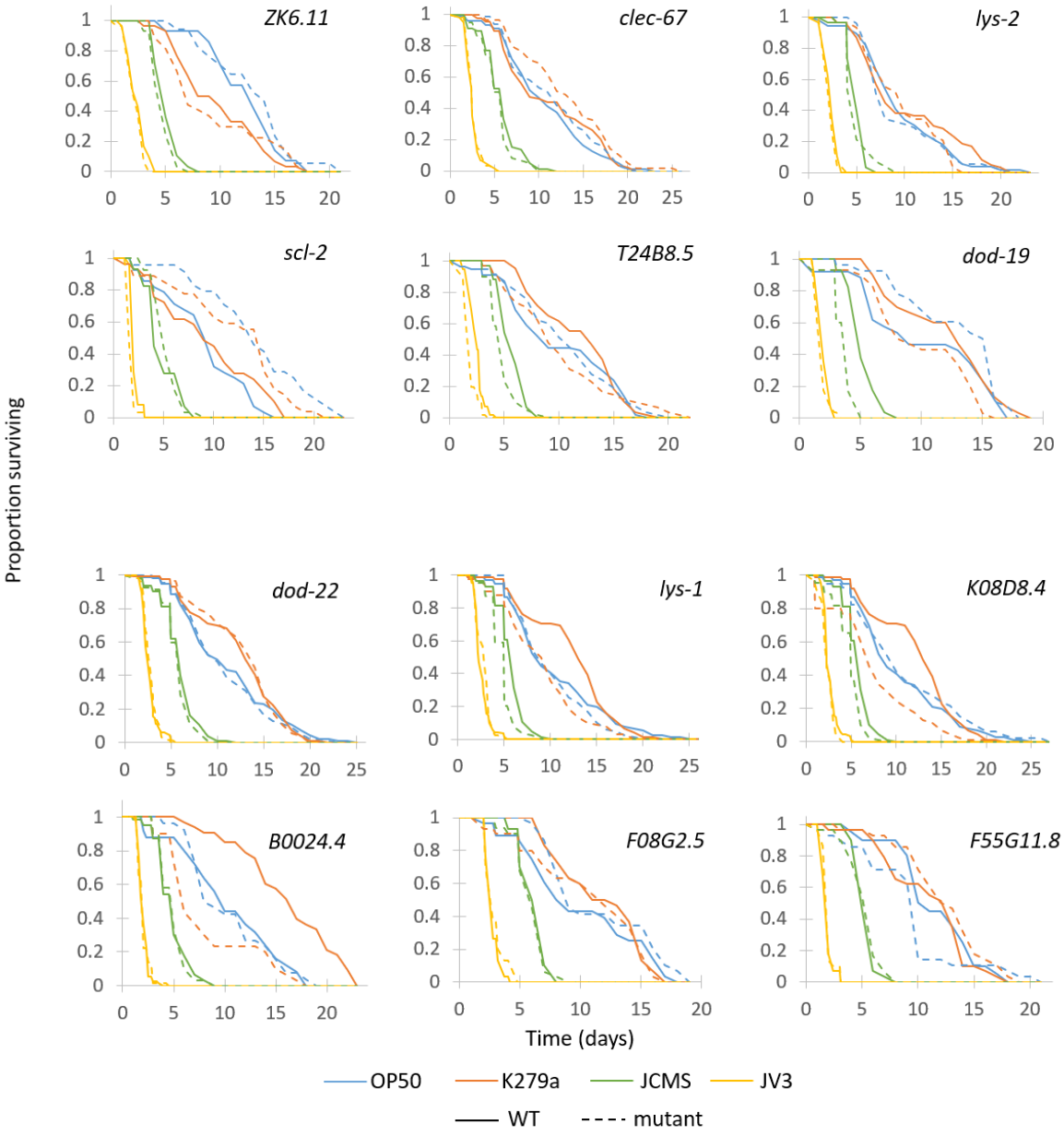

Figure 2.7 Mutations in CPSR genes result in a variety of survival patterns upon *S. maltophilia* exposure

Survivorship of *wild-type* nematodes and CPSR mutants on *S. maltophilia* JCMS, K279a, JV3, and *E. coli* OP50. Survival estimates were determined using Kaplan-Meier estimates generated in R. For these experiments, 10 - 12 worms were synchronized, picked onto each treatment bacterial lawn and the number of living worms was recorded daily. 1 - 3 replicates were completed for all bacterial and *C. elegans* strain combinations. Sample sizes, hazard ratios and p-values generated form Cox proportional hazards tests are shown in Table 2.7.

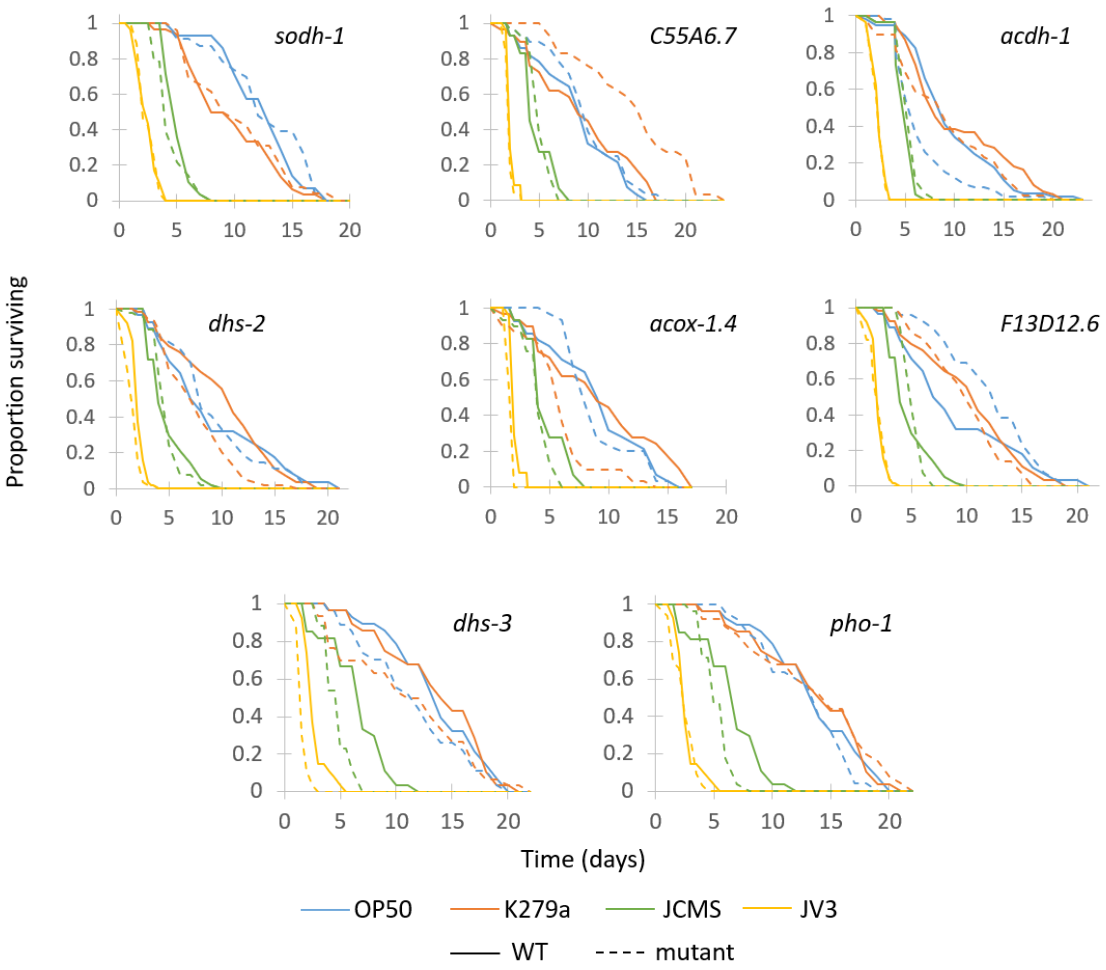

Figure 2.8 Mutations in VSR genes result in a variety of survival patterns upon *S. maltophilia* exposure

Survivorship of *wild-type* nematodes and VSR mutants on *S. maltophilia* JCMS, K279a, JV3, and *E. coli* OP50. Survival estimates were determined using Kaplan-Meier estimates generated in R. For these experiments, 10 - 12 worms were synchronized, picked onto each treatment bacterial lawn and the number of living worms was recorded daily. 1 - 3 replicates were completed for all bacterial and *C. elegans* strain combinations. Sample sizes, hazard ratios and p-values generated form Cox proportional hazards tests are shown in Table 2.7.

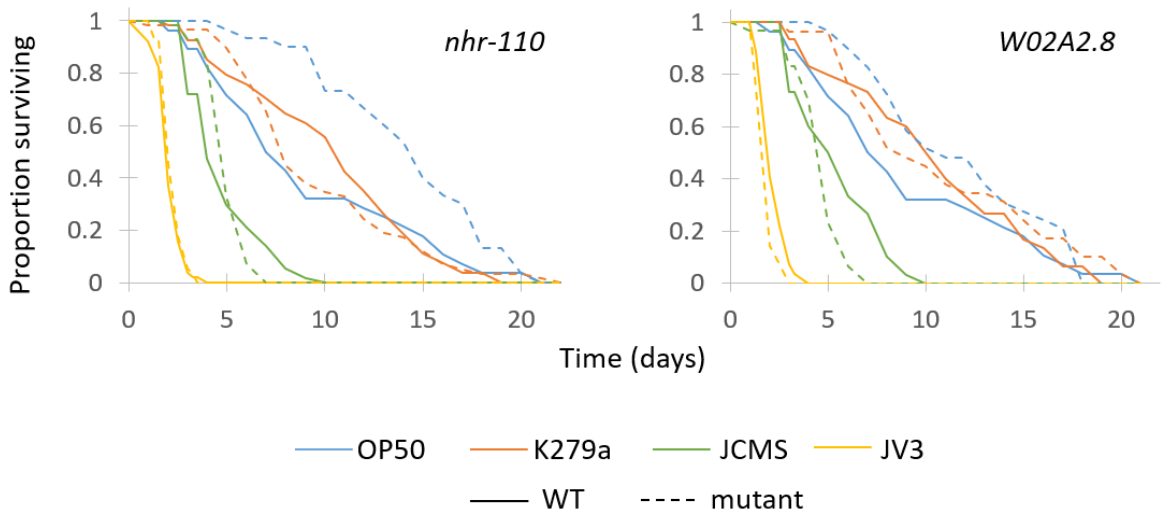

Figure 2.9 Mutations in JSR genes result in a variety of survival patterns upon *S. maltophilia* exposure

Survivorship of *wild-type* nematodes and JSR mutants on *S. maltophilia* JCMS, K279a, JV3, and *E. coli* OP50. Survival estimates were determined using Kaplan-Meier estimates generated in R. For these experiments, 10 - 12 worms were synchronized, picked onto each treatment bacterial lawn and the number of living worms was recorded daily. 1 - 3 replicates were completed for all bacterial and *C. elegans* strain combinations. Sample sizes, hazard ratios and p-values generated form Cox proportional hazards tests are shown in Table 2.7.

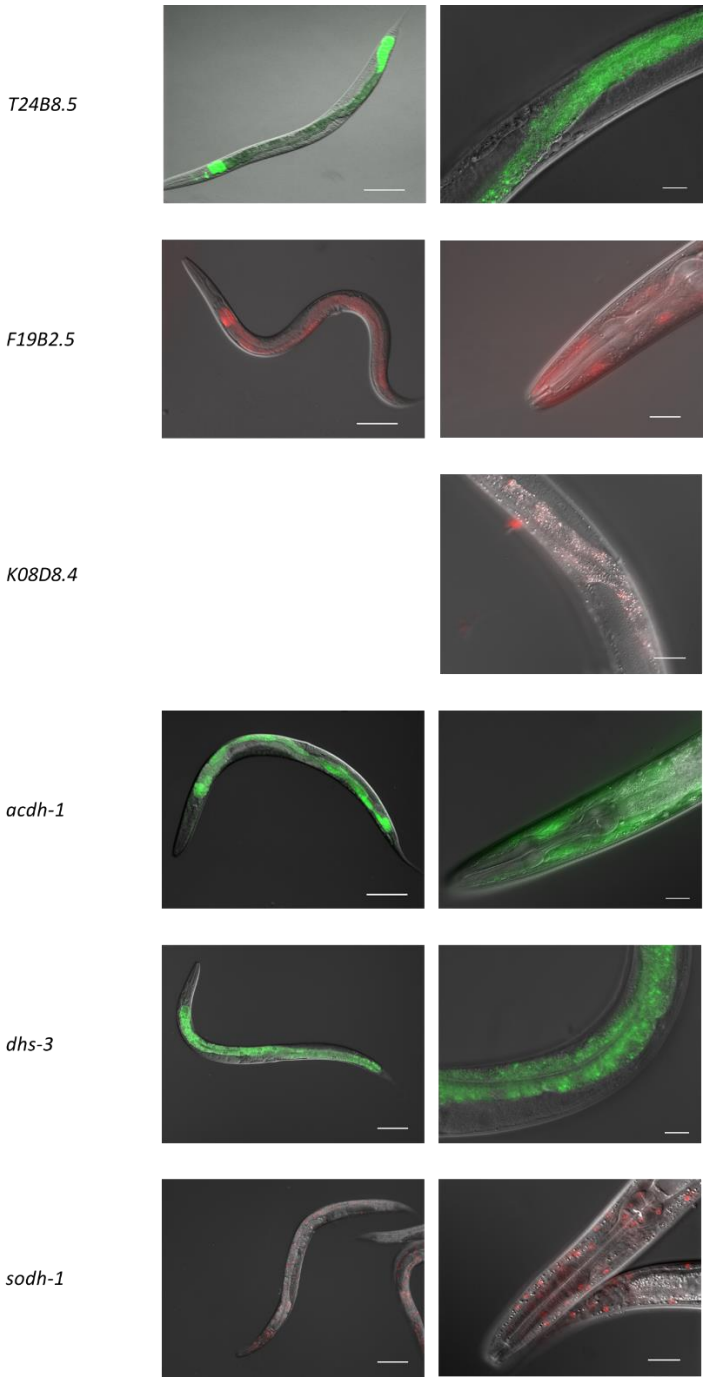

Figure 2.10 A majority of tested CPSR and VSR genes are expressed in innate immune response tissues

Expression of several CPSR (*T24B8.5*, *F19B2.5*, *K08D8.4*) and VSR (*sodh-1*, *acdh-1*, *dhs-3*) genes using transcriptional or translational fluorescent protein fusions upon exposure to *E. coli* OP50. 100x magnification is shown on the left, 400x magnification is shown on the right. Scale bar indicates 100 μm for 100x pictures and 20 μm for 40x pictures. Expression for *K08D8.4* was too dim to be observed at 100x. Note that expression in the AIY interneuron in the *T24B8.5* transgenic strain is due to a *ttx-3*:GFP marker and not *T24B8.5* expression.

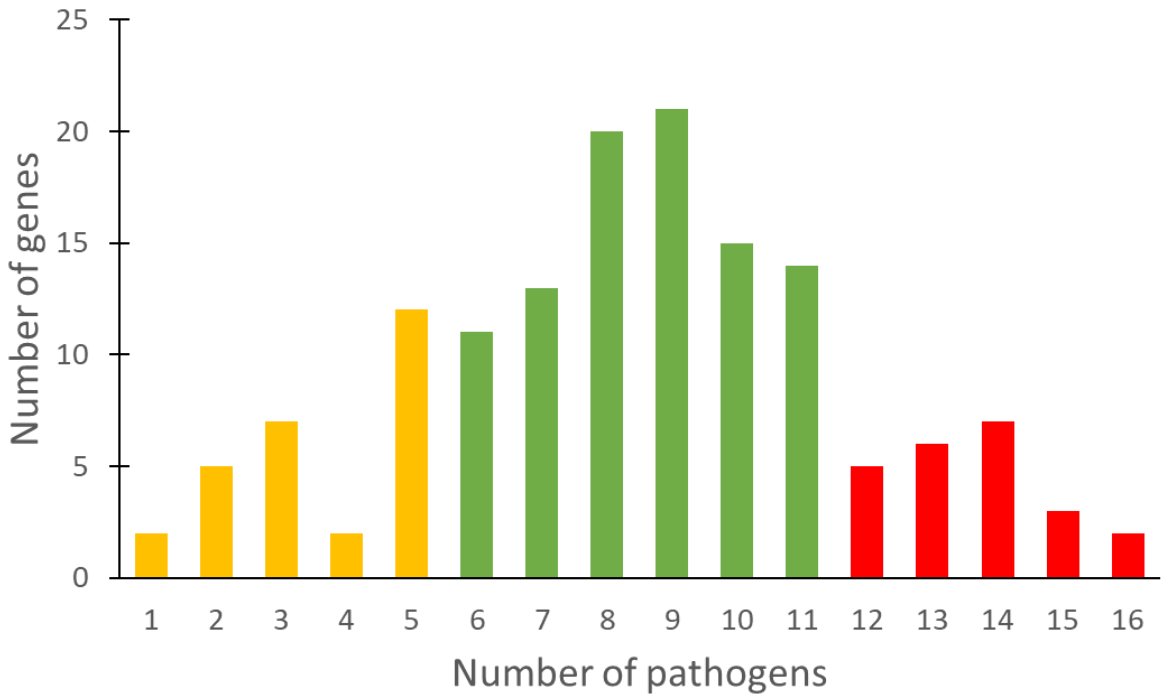

Figure 2.11 Many CPSR genes are commonly differentially expressed in response to other pathogens

The WormExp microbe exposure dataset was used to determine how many pathogens each CPSR gene is differentially expressed in response to, with the number of pathogens ranging from 1-16 pathogens, excluding *S. maltophilia*. This bar graph depicts how many CPSR genes are regulated by each number of pathogens. Note that the pathogens represented by each x-value are not the same for each gene. Genes represented by the yellow bars represent low-ranking, or *S. maltophilia*-specific genes, and genes represented in red bars indicate high-ranking, or common pathogen genes. Candidate genes were chosen from the yellow and red bars for functional analysis.

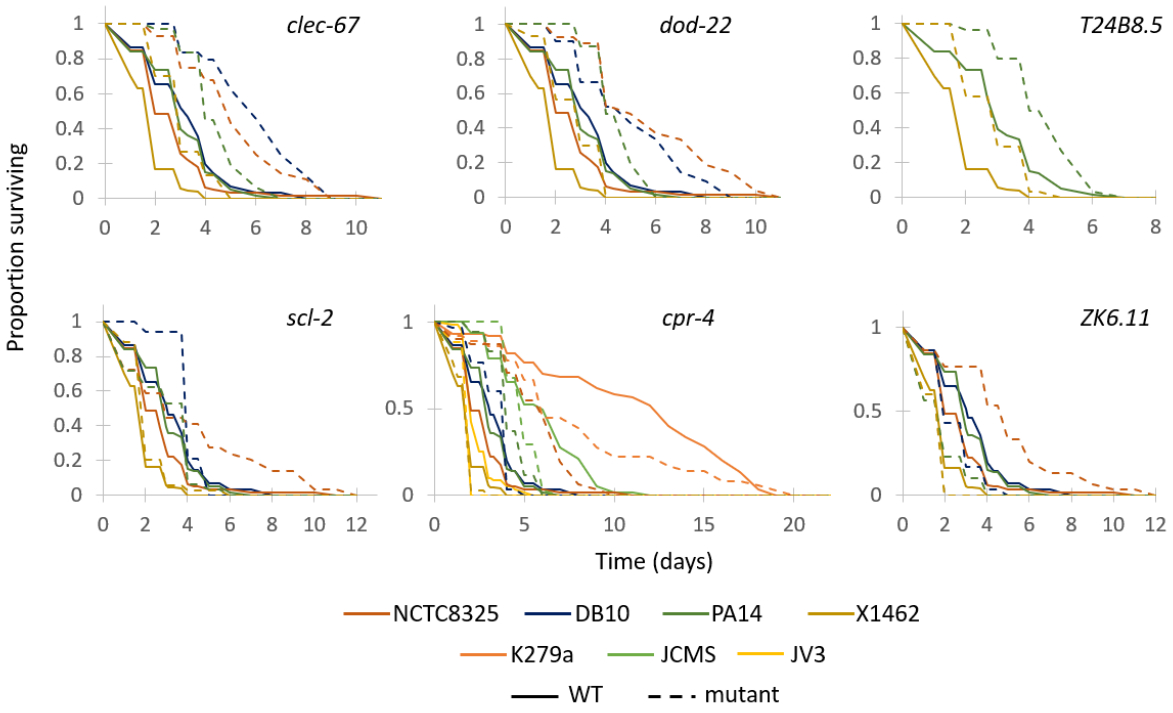

Figure 2.12 Survival analyses of mutations in common pathogen response genes

Survivorship of *wild-type* nematodes and mutants on *S. maltophilia* JCMS, *S. maltophilia* K279a, *S. maltophilia* JV3, *S. aureus* NCTC8325, *S. marcescens* DB10, *P. aeruginosa* PA14, and *X. nematophila* X1462. *S. maltophilia* data not shown can be found in Figure 2.7. Survival estimates were determined using Kaplan-Meier estimates generated in R. For these experiments, 10 - 12 worms were synchronized, picked onto each treatment bacterial lawn and the number of living worms was recorded daily. Sample sizes, hazard ratios and p-values generated form Cox proportional hazards tests are shown in Table 2.10.

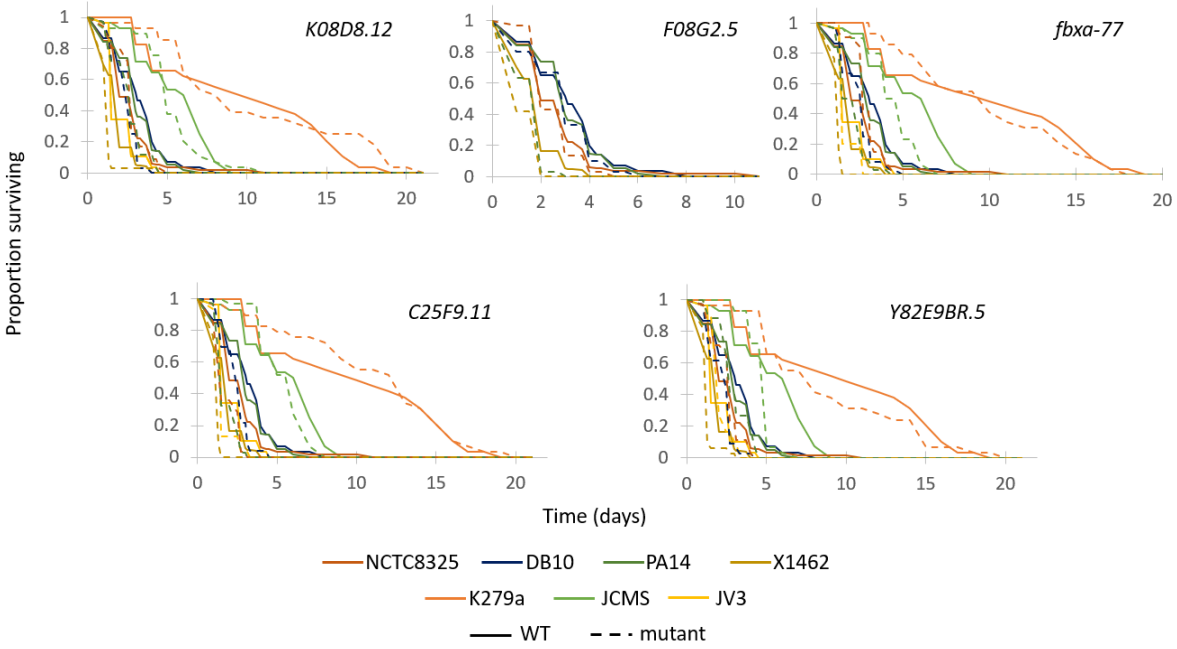

Figure 2.13 Survival analyses of mutations in *S. maltophilia*-specific genes

Survivorship of *wild-type* nematodes and mutants on *S. maltophilia* JCMS, *S. maltophilia* K279a, *S. maltophilia* JV3, *S. aureus* NCTC8325, *S. marcescens* DB10, *P. aeruginosa* PA14, and *X. nematophila* X1462. *S. maltophilia* data not shown can be found in Figure 2.7. Survival estimates were determined using Kaplan-Meier estimates generated in R. For these experiments, 10 - 12 worms were synchronized, picked onto each treatment bacterial lawn and the number of living worms was recorded daily. Sample sizes, hazard ratios and p-values generated form Cox proportional hazards tests are shown in Table 2.10.

Tables

Table 2.1 CRISPR/Cas9 target gene primers

Forward and reverse primers for each gRNA were annealed and ligated into pRB1017. gRNA sequence (20 bp prior to the PAM site) are underlined for each primer, non-underlined bases are included in the primer sequence for proper ligation into *Bsa*I cut pRB1017. Primers flanking gRNA target loci were used to amplify DNA from possible mutant worms to detect insertions or deletions based on amplicon size. If an odd number of flanking primers are listed they were tested in combination (forward primer was tested with each reverse primers). All primers are listed 5' to 3'.

Gene	gRNA Primers	Flanking Primers
B0024.4	1 F: TCTTGCTAGAAATAGAAAGTTCGTT 1 R: AAACAACGAACTTTCTATTTCTAGC 2 F: TCTTGTCCTTCATACAACTTTACAG 2 R: AAACCTGTAAAGTTGTATGAAGGAC 3 F: TCTTGGGACGTCAAACTACATCACG 3 R: AAACCGTGATGTAGTTTGACGTCCC	F: CCATTTACACTCCTCCTC R: TTTACATCAAAATCTTTCAAGTTGAG R: AATTGTAATGATAAATGACGTGAATAG
F08G2.5	1 F: TCTTGTCGAAGAATCCGTCTCCAAG 1 R: AAACCTTGGAGACGGATTCTTCGAC 2 F: TCTTGAAGATCGTAGAGACACCCAA 2 R: AAACTTGGGTGTCTCTACGATCTTC 3 F: TCTTGACAGAGATCGAAGAGAAAGT 3 R: AAACACTTTCTCTTCGATCTCTGTC	F: TAAAACCAGCACCTCTCACC R: ATCATCAGAGTCATCAGAAGAG
K08D8.4	1 F: TCTTGATTAAGTGTACCTACCCGAA 1 R: AAACTTCGGGTAGGTACACTTAATC 2 F: TCTTGATATACATCGACCTTCCGT 2 R: AAACGACGGAAGGTCGATGTATATC 3 F: TCTTGGATTACTTGACTCTTCCGAA 3 R: AAACTTCGGAAGAGTCAAGTAATCC 4 F: TCTTGTTAAAAATAGAACAATACTT 4 R: AAACAAGTATTGTTCTATTTTTAAC	F: CAGATAAATGTTCCTGAAGGC R: GCATCACTTGATTCACAGC R: CAGTGTTGGGAATGTTGTTG
W02A2.8	1 F: TCTTGAGCGGATTCCCGATTCACGA 1 R: AAACTCGTGAATCGGGAATCCGCTC 2 F: TCTTGGGAAGCGTCCTCATTCAACG 2 R: AAACCGTTGAATGAGGACGCTTCCC 3 F: TCTTGATTTTGACCCCCCATGACGG 3 R: AAACCCGTCATGGGGGGTCAAAATC	F: GTATTTCTTGTGATTCTAGAGTCACC R: GGAAGAAAATAGCGGAATAGGTTAC R: AACAACCAAAGACGAACCTC
fbxa-77	1 F: TCTTGGTTTCTAGACACATCCCGCA 1 R: AAACTGCGGGATGTGTCTAGAAACC 2 F: TCTTGGAGCAGCTGTCGATAACGTG 2 R: AAACCACGTTATCGACAGCTGCTCC	F: GGAGTGGAGCTACTGGAAAAC R: ATGGGCTAATCGAGCCTTC
K08D8.12	1 F: TCTTGGTTGAATCAGAATAAGCGGG 1 R: AAACCCCGCTTATTCTGATTCAACC 2 F: TCTTGGGTACAACTACAATAACGGC 2 R: AAACGCCGTTATTGTAGTTGTACCC	F: TGCTTTTTCGTTAGTCTCTACTTCG R: AAGATAGGATGGTGCGCGAG

C25F9.11	1 F: TCTTGTTTACGTTGAAGAGTCGCAA 1 R: AAACTTGCGACTCTTCAACGTAAAC 2 F: TCTTGCGTTCTTTCTGGGTAAGGGT 2 R: AAACACCCTTACCCAGAAAGAACGC	F: ACCTTTATTAGTTTCCCGATGGC R: CGACGGCAGATTGTTCATGG
Y82E9BR.5	1 F: TCTTGACAGTGTTGCTCGTTTTACG 1 R: AAACCGTAAAACGAGCAACACTGTC 2 F: TCTTGAGCTCCCCAACTGCAACACG 2 R: AAACCGTGTTGCAGTTGGGGAGCTC	F: GCATTCTACAATGTCTTTTTGTTCCC R: ATCAGAGCCGACGAGGATG

Table 2.2 All differentially expressed genes between treatment comparisons

*TO VIEW TABLE DOWNLOAD SUPPLEMENTAL FILE "LeahRadeke2018-Table2.2-differentially expressed genes"

List of all differentially expressed genes between all pathogenic and nonpathogenic comparisons (represented in Figure 2.4 *C. elegans* expresses a common set of 145 genes in response to pathogenic *S. maltophila* strains. Fragments per kilobase per million (FPKM) values are shown for each treatment. Additional sheets contain CPSR, VSR, and JSR genes with FPKM values and up/down-regulation in reference to the pathogenic strains (up/down-regulated in response to JV3 and/or JCMS). Asterisk indicates different direction of regulation depending on the comparison considered.

Table 2.3 Innate immune response GO terms are significantly enriched in common pathogenic *S. maltophilia* response (CPSR) genes

Gene ontology (GO) enrichment analysis was performed on the 145 CPSR genes using DAVID Bioinformatics Resources 6.8. GO analysis identifies terms relating to the biological process, molecular function, or cellular component that are significantly enriched among a list of genes. Indented terms indicate child terms, or subcategories, of the term listed above, with the parent term left-aligned. Note that the degree of indention of each term does not reflect absolute GO term level within each category. Only terms with FDR <.05 and the most descriptive term for each unique gene list are shown.

	Term	Count	%	FDR
Biological process	defense response	52	37.68	5.07E-48
	innate immune response	50	36.23	4.04E-53
	defense response to other organism	16	11.59	4.43E-11
	defense response to bacterium	16	11.59	3.35E-12
	defense response to Gram-negative bacterium	14	10.14	2.72E-11
Molecular function	carbohydrate binding	12	8.70	1.28E-04
Cellular component	membrane raft	17	12.32	3.80E-20

Table 2.4 Metabolism and enzyme GO terms are significantly enriched in *S. maltophilia* JV3-specific response (VSR) genes

Gene ontology (GO) enrichment analysis was performed on the 221 VSR genes using DAVID Bioinformatics Resources 6.8. GO analysis identifies terms relating to the biological process, molecular function, or cellular component that are significantly enriched among a list of genes. Indented terms indicate child terms, or subcategories, of the term listed above, with the parent term left-aligned. Note that the degree of indention of each term does not reflect absolute GO term level within each category. Only terms with FDR <.05 and the most descriptive term for each unique gene list are shown.

	Term	Count	%	FDR
Biological process	metabolic process	35	16.51	6.68E-05
	flavonoid biosynthetic process	13	6.13	3.33E-08
	oxidation-reduction process	21	9.91	3.94E-03
	organic acid metabolic process	21	9.91	3.64E-05
	carboxylic acid metabolic process	20	9.43	5.27E-05
	monocarboxylic acid metabolic process	17	8.02	3.30E-06
	flavonoid glucuronidation	13	6.13	3.33E-08
	transmembrane transport	19	8.96	0.002122
	transition metal ion transport	7	3.30	0.035263
Molecular Function	transferase activity, transferring glycosyl groups	16	7.55	0.003022
	transferase activity, transferring hexosyl groups	15	7.08	0.002674
	glucuronosyltransferase activity	11	5.19	3.07E-05
	UDP-glycosyltransferase activity	15	7.08	2.68E-06
	carboxylic ester hydrolase activity	11	5.19	3.06E-04
	oxidoreductase activity	20	9.43	0.009001
Cellular component	extracellular region	20	9.43	2.99E-04
	extracellular space	13	6.13	0.003195
	intracellular membrane-bounded organelle	14	6.60	6.37E-07

Table 2.5 CPSR genes ordered based on gene network rank

WormNet v2 was queried with the 145 genes that were significantly differentially expressed in response to nonpathogenic vs pathogenic strains. 73 of the 145 genes were connected within the network and are ordered based on WormNet score, which is based on the number of connections that gene has and the strength of the evidence for those connections (Score). C= number of connected CPSR genes to the listed gene. Up/down-regulated= direction of differential expression in response to pathogenic compared to nonpathogenic strains. Asterisk indicates different direction of regulation depending on the comparison considered.

Gene name	Score	C	Connected genes	Up/Down-regulated
lys-2	3.67	26	B0024.4 clec-41 C08E8.4 C17H12.6 C17H12.8 C34H4.2 C49C3.9 C49G7.10 F08G2.5 F08G5.6 F19B2.5 F35E12.7 dod-23 hsp-17 dod-22 F55G11.8 clec-67 K08D8.4 K08D8.5 K09D9.1 cpr-3 fbxa-60 T24B8.5 T24C4.4 dod-19 ZK6.11	Up
C49G7.10	3.65	23	B0024.4 tsp-1 C06B3.7 C08E8.4 C10C5.2 C17H12.6 F08G2.5 F08G5.6 F19B2.5 F53A9.1 F53A9.6 K08D8.4 K09D9.1 mtl-1 R10D12.9 cpr-3 fbxa-60 fbxa-55 T24B8.5 T24C4.4 lys-1 lys-2 lys-3	Up
F19B2.5	3.52	20	B0024.4 B0496.7 nuc-1 C17H12.6 C49G7.10 F08G2.5 F08G5.6 dod-22 F55G11.8 clec-67 K08D8.4 K08D8.5 K09D9.1 spp-1 cpr-3 T24C4.4 lys-1 lys-2 dod-19 ZK6.11	Up
T24C4.4	3.36	17	B0024.4 clec-41 C08E8.4 C17H12.6 C49G7.10 F08G2.5 F08G5.6 F19B2.5 F53A9.1 dod-22 F55G11.8 K08D8.4 K09D9.1 cpr-3 T24B8.5 lys-1 lys-2	Up
F08G2.5	3.26	17	B0024.4 tsp-1 C08E8.4 C49G7.10 F08F8.5 F19B2.5 lec-11 F53A9.1 F53A9.2 K08D8.4 K09D9.1 mtl-1 cpr-3 T24B8.5 T24C4.4 lys-2 lys-3	Up
K09D9.1	3.19	11	clec-41 tsp-1 C08E8.4 C49G7.10 F08G2.5 F19B2.5 K08D8.4 cpr-3 T24C4.4 lys-2 lys-3	Up
C17H12.6	3.1	16	B0024.4 clec-41 C49G7.10 F19B2.5 dod-22 F55G11.8 clec-67 K08D8.4 K08D8.5 cpr-3 T24B8.5 T24C4.4 lys-1 lys-2 dod-19 ZK6.11	Up
dod-22	3.09	15	B0024.4 C17H12.6 C17H12.8 F08G5.6 F19B2.5 F35E12.7 F55G11.8 clec-67 K08D8.4 K08D8.5 T24C4.4 cpr-5 lys-2 dod-19 ZK6.11	Up
F55G11.8	3.08	14	B0024.4 C17H12.6 C17H12.8 F08G5.6 F19B2.5 dod-22 clec-67 K08D8.4 K08D8.5 T24C4.4 lys-1 lys-2 dod-19 ZK6.11	Up

ZK6.11	3.06	15	B0024.4 C17H12.6 C34H4.2 C49C3.9 F08G5.6 F19B2.5 F35E12.7 dod-22 F55G11.8 clec-67 K08D8.4 K08D8.5 cpr-3 lys-1 lys-2	Up
C08E8.4	3.03	18	B0024.4 tsp-1 C49G7.10 F08G2.5 F49E11.10 dod-23 F53A9.1 F53A9.6 F53A9.8 K09D9.1 mtl-1 cpr-3 fbxa-60 T24B8.5 T24C4.4 lys-2 lys-3 fbxa-30	Up
T24B8.5	2.96	9	clec-41 C08E8.4 C17H12.6 C49G7.10 F08G2.5 dod-23 T24C4.4 lys-1 lys-2	Up
K08D8.4	2.92	15	B0024.4 C17H12.6 C49G7.10 F08G2.5 F08G5.6 F19B2.5 dod-22 F55G11.8 clec-67 K09D9.1 cpr-3 T24C4.4 lys-2 dod-19 ZK6.11	Up
F08G5.6	2.89	19	B0024.4 cyp-35A2 C49C3.9 C49G7.10 F19B2.5 F35E12.7 F53A9.6 dod-22 F55G11.8 clec-67 cyp-35A5 K08D8.4 K08D8.5 cpr-3 T24C4.4 cpr-5 lys-2 dod-19 ZK6.11	Up
B0024.4	2.85	16	C08E8.4 C17H12.6 C49G7.10 F08G2.5 F08G5.6 F19B2.5 dod-22 F55G11.8 clec-67 K08D8.4 cpr-3 T24C4.4 lys-1 lys-2 dod-19 ZK6.11	Up
cpr-3	2.82	21	B0024.4 tsp-1 C08E8.4 C17H12.6 C49C3.9 C49G7.10 F08G2.5 F08G5.6 F19B2.5 F49E11.10 F53A9.6 F53A9.8 clec-67 K08D8.4 K09D9.1 fbxa-60 T24C4.4 lys-2 lys-3 dod-19 ZK6.11	Up
lys-1	2.81	15	B0024.4 clec-41 C17H12.6 C17H12.8 C49G7.10 F19B2.5 cpr-4 hsp-17 F55G11.8 K08D8.5 T24B8.5 T24C4.4 cpr-5 dod-19 ZK6.11	Up
dod-19	2.75	16	B0024.4 clec-41 C17H12.6 C34H4.2 C49C3.9 F08G5.6 F19B2.5 F35E12.7 dod-22 F55G11.8 clec-67 K08D8.4 K08D8.5 cpr-3 lys-1 lys-2	Up
clec-67	2.73	16	B0024.4 nuc-1 C17H12.6 F08G5.6 F19B2.5 F49E11.10 dod-22 F55G11.8 K08D8.4 K08D8.5 spp-1 cpr-3 cpr-5 lys-2 dod-19 ZK6.11	Up
scl-2	2.55	17	tsp-1 nuc-1 C08E8.4 C56A3.2 cpr-4 dod-23 F53A9.1 F53A9.6 F53A9.8 clec-67 vit-1 odc-1 mtl-1 spp-1 cpr-3 fbxa-60 lys-3	Down
F53A9.8	2.51	17	tsp-1 nuc-1 C08E8.4 C08F11.13 lys-10 cpr-4 F49E11.10 dod-23 F53A9.1 F53A9.6 vit-1 mtl-1 spp-1 cpr-3 fbxa-60 fbxa-55 lys-3	Up
lys-3	2.42	18	B0496.7 tsp-1 C08E8.4 C08F11.13 C49G7.10 F08G2.5 F09B9.1 F49E11.10 F53A9.1 F53A9.2 F53A9.6 F53A9.8 K09D9.1 mtl-1 R07C12.1 ugt-53 cpr-3 fbxa-55	Up
clec-41	2.37	8	C17H12.6 C49C3.9 K09D9.1 T24B8.5 T24C4.4 lys-1 lys-2 dod-19	Up

K08D8.5	2.25	11	C17H12.6 C17H12.8 F08G5.6 F19B2.5 dod-22 F55G11.8 clec-67 lys-1 lys-2 dod-19 ZK6.11	Up
T01D3.6	2.19	5	nuc-1 lys-10 clec-66 spp-1 bath-47	Up
mtl-1	2.08	10	tsp-1 C08E8.4 C49G7.10 F08G2.5 F49E11.10 dod-23 F53A9.1 F53A9.8 fbxa-59 lys-3	Down
tsp-1	2.03	15	B0496.7 C08E8.4 C49G7.10 F08G2.5 F49E11.10 F53A9.1 F53A9.2 F53A9.6 F53A9.8 K09D9.1 mtl-1 R10D12.9 cpr-3 lys-3 fbxa-30	Up
clec-66	2.03	4	nuc-1 lys-10 T01D3.6 bath-47	Up
dod-23	2.02	10	C08E8.4 lys-10 cpr-4 F49E11.10 F53A9.8 vit-1 mtl-1 fbxa-60 T24B8.5 lys-2	Up
nuc-1	1.97	9	lys-10 F19B2.5 clec-66 F49E11.10 F53A9.8 clec-67 T01D3.6 spp-1 cpr-5	Down
spp-1	1.97	9	nuc-1 lys-10 F19B2.5 F49E11.10 F53A9.8 clec-67 vit-1 odc-1 T01D3.6	Down
C17H12.8	1.93	6	hsp-17 dod-22 F55G11.8 K08D8.5 lys-1 lys-2	Up
F53A9.6	1.89	13	tsp-1 C08E8.4 C49G7.10 F08G5.6 cpr-4 F49E11.10 F53A9.8 T04F8.7 cpr-3 fbxa-60 fbxa-55 lec-10 lys-3	Up
F35E12.7	1.88	5	F08G5.6 dod-22 lys-2 dod-19 ZK6.11	Up
B0496.7	1.88	5	tsp-1 F19B2.5 F53A9.1 R07C12.1 lys-3	Up
F53A9.1	1.81	10	B0496.7 tsp-1 C08E8.4 C49G7.10 F08G2.5 F49E11.10 F53A9.8 mtl-1 T24C4.4 lys-3	Up
fbxa-55	1.73	6	C49G7.10 cpr-4 F53A9.6 F53A9.8 W05H9.3 lys-3	Up
C08F11.13	1.62	4	F09B9.1 F53A9.8 fbxa-59 lys-3	Up
fbxa-60	1.6	14	C08E8.4 C10C5.2 C49C3.9 C49G7.10 E02C12.8 cpr-4 F49E11.10 dod-23 hsp-17 F53A9.6 F53A9.8 R10D12.9 cpr-3 lys-2	Up
cpr-4	1.57	10	lys-10 F49E11.10 dod-23 F53A9.6 F53A9.8 fbxa-60 fbxa-55 fbxa-59 W02D7.2 lys-1	Down
vit-1	1.54	4	F49E11.10 dod-23 F53A9.8 spp-1	Down
C49C3.9	1.49	10	clec-41 F08G5.6 F53C11.1 cpr-3 fbxa-60 fbxa-59 W02D7.2 lys-2 dod-19 ZK6.11	Up
lys-10	1.47	7	nuc-1 clec-66 cpr-4 dod-23 F53A9.8 T01D3.6 spp-1	*
C34H4.2	1.46	5	hsp-17 lys-2 ZK1055.7 dod-19 ZK6.11	Up
cpr-5	1.43	6	nuc-1 F08G5.6 dod-22 clec-67 W02D7.2 lys-1	Down
C10C5.2	1.39	4	C06B3.7 C49G7.10 R10D12.9 fbxa-60	Up
hsp-17	1.34	6	C17H12.8 C34H4.2 R10D12.9 fbxa-60 lys-1 lys-2	Up
fbxa-59	1.31	4	C08F11.13 C49C3.9 cpr-4 mtl-1	Up
bath-47	1.31	3	F14F9.4 clec-66 T01D3.6	Up

clec-218	1.3	4	C49C3.9 cpr-4 K11H12.4 cpr-5	Down
R10D12.9	1.3	5	tsp-1 C10C5.2 C49G7.10 hsp-17 fbxa-60	Up
odc-1	1.27	2	F49E11.10 spp-1	Up
F53A9.2	1.25	3	tsp-1 F08G2.5 lys-3	Up
F14F9.4	1.18	2	ugt-53 bath-47	Up
C06B3.7	1.12	2	C10C5.2 C49G7.10	Up
cyp-35A5	1.11	2	C45B11.2 F08G5.6	Down
ugt-53	1.09	2	F14F9.4 lys-3	Down
R07C12.1	1.08	2	B0496.7 lys-3	Up
F09B9.1	1.03	2	C08F11.13 lys-3	Up
fbxa-30	1.01	2	tsp-1 C08E8.4	Up
C45B11.2	0.99	3	cyp-35A2 cyp-35A5 cyp-14A2	Up
cyp-35A2	0.97	2	C45B11.2 F08G5.6	Down
F53C11.1	0.9	1	C49C3.9	Up
lec-11	0.87	1	F08G2.5	Up
E02C12.8	0.85	1	fbxa-60	Up
F08F8.5	0.82	1	F08G2.5	Up
ttr-44	0.79	1	F49E11.10	Down
ZK1055.7	0.79	1	C34H4.2	Up
K11H12.4	0.78	1	W02D7.2	Up
lec-10	0.78	1	F53A9.6	Up
T04F8.7	0.78	1	F53A9.6	Up
W05H9.3	0.78	1	fbxa-55	Up
cyp-14A2	0.72	1	C45B11.2	Down

Table 2.6 VSR genes ordered based on gene network rank

WormNet v2 was queried with the 221 genes that were specifically differentially expressed in response to JV3 vs. all other strains. 103 of the 221 genes were connected within the network and are ordered based on WormNet score, which is based on the number of connections that gene has and the strength of the evidence for those connections (Score). C= number of connected CPSR genes to the listed gene. Up/down-regulated= direction of differential expression in response to pathogenic compared to nonpathogenic strains.

Gene name	Score	C	Connected genes	Up/Down-regulated
F10F2.2	3.9	5	tag-10 F38B6.4 sodh-1 T21C9.6 his-25	Down
sodh-1	3.85	16	C14C6.2 C23G10.11 dod-3 C50F7.5 ftn-1 C55A6.7 D1054.8 clec-54 F10F2.2 clec-7 F13D12.6 F46C5.1 nhr-193 dhs-14 Y38F1A.6 Y6G8.2	Up
dhs-3	3.67	4	C55A6.7 D1054.8 dhs-7 dhs-14	Down
F38B6.4	3.62	2	pgp-2 F10F2.2	Down
Y38F1A.6	3.53	13	acdh-1 D1054.8 F08A8.4 F09B12.3 dhs-25 F13D12.6 F18E2.1 glf-1 sodh-1 ugt-47 dhs-14 T03G6.1 pmp-5	Down
clec-57	3.35	19	ugt-17 C14C6.2 tag-10 C45B2.2 pho-1 F08A8.4 clec-54 F09B12.3 F13D12.6 F14D7.6 F18E2.1 F55B11.1 F58B4.5 F58G6.3 ugt-62 T16G12.1 clec-26 W07B8.1 ZC416.6	Down
glf-1	3.35	4	ugt-47 T03G6.1 T26C5.2 Y38F1A.6	Up
F13D12.6	3.34	15	ugt-63 tag-10 C50F7.5 pho-1 F08A8.4 clec-57 clec-54 F09B12.3 F55B11.1 sodh-1 T16G12.1 clec-26 W07B8.1 Y38F1A.6 ZC416.6	Down
dhs-7	3.28	8	C01B10.10 C23H4.3 C55A6.7 dhs-25 dhs-2 ugt-47 dhs-3 Y40H7A.10	Down
F09B12.3	3.27	15	C14C6.2 tag-10 D1054.8 pho-1 clec-57 clec-54 F13D12.6 F55B11.1 F58G6.3 R08E5.3 T16G12.1 clec-26 W07B8.1 Y38F1A.6 ZC416.6	Down
ZC416.6	3.24	13	ugt-63 ncx-7 tag-10 pho-1 clec-57 clec-54 F09B12.3 F13D12.6 F14D7.6 F55B11.1 T16G12.1 clec-26 W07B8.1	Down
T03G6.1	3.21	3	glf-1 ugt-47 Y38F1A.6	Up
pho-1	3.18	14	ugt-63 ncx-7 tag-10 C45B2.2 clec-57 clec-54 F09B12.3 F13D12.6 F14D7.6 F55B11.1 T16G12.1 clec-26 W07B8.1 ZC416.6	Down
acdh-1	3.11	7	ugt-63 C14C6.2 F08A8.2 F08A8.4 F15E6.4 pmp-5 Y38F1A.6	Down

C55A6.7	3.09	4	C01B10.10 dhs-7 sodh-1 dhs-3	Down
D1054.8	3	7	odd-1 ugt-63 F09B12.3 F58B4.5 sodh-1 dhs-3 Y38F1A.6	Down
F46C5.1	2.83	4	C23G10.11 F14D7.6 nhr-193 sodh-1	Up
dhs-2	2.82	6	ugt-51 dod-3 C29F7.1 C29F7.2 dhs-7 ugt-5	Down
F08A8.4	2.82	10	acdh-1 clec-57 F13D12.6 F18E2.1 F58B4.5 ugt-47 T05E7.1 pmp-5 Y38F1A.6 ZC376.2	Down
clec-26	2.81	11	ugt-63 ncx-7 C45B2.2 pho-1 clec-57 clec-54 F09B12.3 F13D12.6 F14D7.6 W07B8.1 ZC416.6	Down
dhs-14	2.8	4	odd-1 sodh-1 dhs-3 Y38F1A.6	Down
C23G10.11	2.6	3	F46C5.1 sodh-1 Y40H7A.10	Up
tag-10	2.58	10	ftn-1 pho-1 clec-57 F09B12.3 F10F2.2 F13D12.6 F55B11.1 W07B8.1 ndx-8 ZC416.6	Down
clec-54	2.54	20	C14C6.2 dod-3 C45B2.2 pho-1 clec-57 F09B12.3 clec-7 F13D12.6 F18E2.1 F35E12.6 nhr-193 H22K11.2 sodh-1 ugt-47 T05E7.1 T16G1.4 clec-26 T24C2.5 Y40H7A.10 ZC416.6	Down
F08A8.2	2.5	3	acdh-1 T05E7.1 ndx-8	Down
dhs-25	2.41	2	dhs-7 Y38F1A.6	Down
W07B8.1	2.39	7	tag-10 pho-1 clec-57 F09B12.3 F13D12.6 clec-26 ZC416.6	Down
T16G12.1	2.26	11	ugt-63 ncx-7 pho-1 clec-57 F09B12.3 F13D12.6 F14D7.6 F55B11.1 T25G12.6 ndx-8 ZC416.6	Down
F55B11.1	2.24	10	ncx-7 tag-10 pho-1 clec-57 F09B12.3 F13D12.6 F14D7.6 T16G12.1 ndx-8 ZC416.6	Down
C29F7.2	2.1	6	ugt-51 C10C5.4 math-26 dhs-2 ugt-9 ugt-5	Down
ugt-5	2.09	5	ugt-51 C29F7.1 C29F7.2 ugt-21 dhs-2	Down
ndx-8	2.09	9	C01B10.4 C23H4.7 tag-10 C35A5.3 F08A8.2 F55B11.1 T16G12.1 T21C9.6 sulp-7	Down
C35A5.3	2.07	7	F18E2.1 F35E12.6 F58B4.5 ugt-47 pmp-5 ndx-8 ZK228.4	Down
F35E12.6	1.97	9	C25F9.4 C35A5.3 clec-54 F18E2.1 F58B4.5 H06H21.8 H22K11.2 ugt-47 ZC376.2	Up
ugt-47	1.92	18	C01B10.10 C23H4.3 C35A5.3 dhs-7 F08A8.4 clec-54 F18E2.1 F35E12.6 F58B4.5 glf-1 H06H21.8 ugt-62 T03G6.1 T05E7.1 pmp-5 Y38F1A.6 Y40H7A.10 ZC376.2	Down
ugt-21	1.91	3	ugt-62 ugt-9 ugt-5	Down
ugt-63	1.88	11	ncx-7 acdh-1 D1054.8 pho-1 F13D12.6 F14D7.6 F58G6.3 grd-14 T16G12.1 clec-26 ZC416.6	Down
ugt-9	1.83	2	C29F7.2 ugt-21	Down

F14D7.6	1.81	10	ugt-63 ncx-7 C45B2.2 pho-1 clec-57 F46C5.1 F55B11.1 T16G12.1 clec-26 ZC416.6	Down
C14C6.2	1.79	7	acdh-1 clec-57 clec-54 F09B12.3 clec-7 H22K11.2 sodh-1	Down
F58G6.3	1.78	6	ugt-63 clec-57 F09B12.3 F58B4.5 ugt-62 grd-14	Down
ugt-62	1.77	9	C10C5.4 ugt-21 clec-57 F58B4.5 F58G6.3 ugt-19 ugt-47 ugt-30 ugt-6	Down
F58B4.5	1.68	13	C25F9.4 C35A5.3 C49A9.4 D1054.8 F08A8.4 clec-57 F18E2.1 F35E12.6 F58G6.3 ugt-62 ugt-47 ZC376.2 ZK228.4	Down
ncx-7	1.66	7	ugt-63 pho-1 F14D7.6 F55B11.1 T16G12.1 clec-26 ZC416.6	Down
F18E2.1	1.61	12	C35A5.3 F08A8.4 clec-57 clec-54 F35E12.6 F58B4.5 H06H21.8 ugt-47 T05E7.1 pmp-5 Y38F1A.6 ZC376.2	Down
C25F9.4	1.57	2	F35E12.6 F58B4.5	Down
Y40H7A.10	1.55	7	C23G10.11 dhs-7 clec-54 H22K11.2 ugt-47 W02D7.4 ins-7	Down
C29F7.1	1.54	2	dhs-2 ugt-5	Down
ugt-51	1.52	5	C29F7.2 ftn-1 dhs-2 Y6G8.2 ugt-5	Down
dod-3	1.51	5	hil-7 ftn-1 clec-54 dhs-2 sodh-1	Down
Y6G8.2	1.5	7	ugt-51 ftn-1 nhr-193 H22K11.2 sodh-1 T03E6.8 ZC196.2	Down
T05E7.1	1.47	7	F08A8.2 F08A8.4 clec-54 clec-7 F18E2.1 K09C4.1 ugt-47	Down
pmp-5	1.47	6	C35A5.3 acdh-1 F08A8.4 F18E2.1 ugt-47 Y38F1A.6	Down
ZC376.2	1.46	6	C23H4.3 F08A8.4 F18E2.1 F35E12.6 F58B4.5 ugt-47	Down
C10C5.4	1.45	3	C29F7.2 ugt-62 cah-5	Down
ftn-1	1.44	5	ugt-51 dod-3 tag-10 sodh-1 Y6G8.2	Down
grd-14	1.44	2	ugt-63 F58G6.3	Down
C45B2.2	1.43	5	pho-1 clec-57 clec-54 F14D7.6 clec-26	Down
C01B10.10	1.43	3	C55A6.7 dhs-7 ugt-47	Down
clec-7	1.37	5	C14C6.2 clec-54 sodh-1 T05E7.1 T16G1.4	Down
ugt-17	1.33	2	math-3 clec-57	Down
nhr-193	1.28	5	clec-54 F46C5.1 H22K11.2 sodh-1 Y6G8.2	Down
C50F7.5	1.26	2	F13D12.6 sodh-1	Up
H22K11.2	1.25	6	C14C6.2 clec-54 F35E12.6 nhr-193 Y40H7A.10 Y6G8.2	Down
R08E5.3	1.22	1	F09B12.3	Down
C23H4.3	1.2	3	dhs-7 ugt-47 ZC376.2	Down
math-3	1.2	2	ugt-17 math-40	Down
T16G1.4	1.16	3	clec-54 clec-7 cat-4	Down

T03E6.8	1.16	3	Y20C6A.1 Y6G8.2 ZC196.2	Down
T06D8.3	1.16	2	pqn-13 pqn-73	Up
H06H21.8	1.16	3	F18E2.1 F35E12.6 ugt-47	Down
pqn-13	1.15	2	T06D8.3 pqn-73	Up
pqn-73	1.15	3	pqn-13 C24H12.11 T06D8.3	Up
ZK228.4	1.13	2	C35A5.3 F58B4.5	Down
cat-4	1.08	2	cah-5 T16G1.4	Down
math-26	1.06	2	C29F7.2 math-40	Down
Y20C6A.1	1.03	2	K11G9.2 T03E6.8	Down
T21C9.6	1.01	2	F10F2.2 ndx-8	Down
W02D7.4	1.01	1	Y40H7A.10	Down
math-40	1.01	2	math-3 math-26	Down
ZC196.2	0.99	2	T03E6.8 Y6G8.2	Down
cah-5	0.97	2	C10C5.4 cat-4	Down
ugt-19	0.88	1	ugt-62	Down
odd-1	0.85	2	D1054.8 dhs-14	Down
T24C2.5	0.82	1	clec-54	Down
K09C4.1	0.82	1	T05E7.1	Down
T25G12.6	0.81	1	T16G12.1	Down
hil-7	0.81	1	dod-3	Up
ins-7	0.8	1	Y40H7A.10	Up
C49A9.4	0.8	1	F58B4.5	Down
T26C5.2	0.79	1	glf-1	Up
F15E6.4	0.78	1	acdh-1	Down
C24H12.11	0.78	1	pqn-73	Down
K11G9.2	0.77	1	Y20C6A.1	Down
pgp-2	0.69	1	F38B6.4	Down
ugt-6	0.68	1	ugt-62	Down
C23H4.7	0.68	1	ndx-8	Down
sulp-7	0.65	1	ndx-8	Down
C01B10.4	0.63	1	ndx-8	Down
W05E10.1	0.63	1	T05C3.6	Down
T05C3.6	0.63	1	W05E10.1	Down
ugt-30	0.63	1	ugt-62	Down
his-25	0.58	1	F10F2.2	Up

Table 2.7 Cox proportional hazard ratios for common *S. maltophilia* and strain-specific genes

Mean survival (M), standard error of the mean (SE), and sample size (N), are given for each allele and bacterial treatment combination. *Wild-type* statistics were determined from combining all *wild-type* data from all experiments. Hazard ratios indicate the treatment hazard divided by the hazard of *wild-type* (first column) or the hazard of that allele on OP50 (last column) within that experiment. The hazard is defined as the probability of a nematode dying at a given time. Hazard ratios and associated p values for each comparison were determined using Cox proportional hazards tests in R. Asterisk and yellow shading indicate significant p-values ($p<0.05$).

						Relative to *wild-type*		Relative to OP50	
	Nematode	Bacteria	N	M	SE	Hazard Ratio	p value	Hazard Ratio	p value
	wild-type	OP50	428	10.62	0.22	NA		NA	
		K279a	545	11.38	0.08	NA		0.878	0.09
		JCMS	546	5.53	0.03	NA		5.81	<2E-16*
		JV3	524	2.43	0.20	NA		63.45	<2E-16*
CPSR genes	*B0024.4*	OP50	26	8.08	0.24	1.05	0.868	NA	
	(mh82)	K279a	58	7.09	0.24	2.32	7.60E-02	1.37	6.15E-6*
		JCMS	90	4.74	0.16	1.03	0.858	7.02	1.59E-7*
		JV3	58	2.06	0.07	0.95	0.78	128.40	<2E-16*
	F08G2.5	OP50	29	10.52	0.54	0.73	0.25	NA	
	(mh86)	K279a	30	10.30	0.75	1.10	0.72	1.51	0.135
		JCMS	31	6.06	0.24	1.08	0.76	9.77	4.04E-11*
		JV3	30	2.99	0.14	0.73	0.25	181.93	<2E-16*
	ZK6.11	OP50	17	11.47	0.55	0.80	0.54	NA	
	(ok3738)	K279a	27	8.48	0.63	1.00	0.97	1.69	0.1
		JCMS	29	4.79	0.18	1.38	0.23	15.22	1.32E-10*
		JV3	29	2.31	0.14	1.41	0.20	893.41	1.55E-15*

T24B8.5	OP50	58	9.90	0.47	0.91	0.431	NA	
(ok3236)	K279a	61	9.36	0.42	1.21	3.48E-13*	1.04	0.89
	JCMS	60	4.88	0.16	1.73	<2E-16*	8.46	<2E-16*
	JV3	57	1.94	0.07	2.27	<2E-16*	252.53	<2E-16*
dod-19	OP50	28	10.54	0.44	0.64	0.109	NA	
(ok2679)	K279a	30	8.93	0.58	1.72	0.0415*	3.08	.000234*
	JCMS	30	3.63	0.18	5.89	4.53E-9*	78.55	3.34E-14*
	JV3	29	1.97	0.09	1.31	0.3197	754.76	<2E-16*
dod-22	OP50	101	10.50	0.38	1.10	0.486	NA	
(ok1918)	K279a	113	12.47	0.35	1.03	0.7	0.77	0.0967
	JCMS	116	5.71	0.16	1.25	0.32	8.10	<2E-16*
	JV3	121	2.76	0.06	0.91	0.292	124.80	<2E-16*
K08D8.4	OP50	79	9.86	0.46	0.86	0.65	NA	
(mh101)	K279a	90	7.46	0.47	2.38	6.22E-9*	1.88	0.0316
	JCMS	87	4.82	0.17	1.76	<2E-16*	6.81	<2E-16*
	JV3	85	2.36	0.08	1.46	0.1342	45.13	<2E-16*
lys-1	OP50	82	9.84	0.37	1.20	0.514	NA	
(ok2445)	K279a	92	8.93	0.40	1.97	2.46E-8*	1.16	0.343
	JCMS	88	4.76	0.14	2.07	6.65E-7*	8.68	6.83E-12*
	JV3	86	2.77	0.08	0.85	0.6677	82.02	<2E-16*
clec-67	OP50	79	11.15	0.47	0.90	0.184	NA	
(ok2770)	K279a	55	12.38	0.54	0.72	0.063	0.72	.0235*
	JCMS	57	5.46	0.24	1.24	0.112	7.75	<2E-16*
	JV3	53	2.67	0.10	0.96	.000419*	89.57	<2E-16*
lys-2	OP50	55	8.68	0.35	1.11	0.683	NA	
(tm2398)	K279a	29	9.24	0.54	1.35	0.2836	0.98	0.883
	JCMS	28	4.96	0.26	0.99	0.962	6.71	<2E-16*
	JV3	27	2.40	0.09	1.20	0.503	508.63	<2E-16*
F55G11.8	OP50	28	8.71	0.60	1.44	0.23	NA	
(gk3130)	K279a	28	11.29	0.46	0.66	0.132	0.55	.0335*
	JCMS	30	5.37	0.27	0.83	0.471	9.07	1.65E-8*
	JV3	29	2.04	0.08	0.89	0.662	148.11	<2E-16*

VSR genes

scl-2	OP50	24	10.48	0.45	0.28	6.89E-5*	NA	
(tm2428)	K279a	27	9.48	0.63	0.56	.035*	1.52	0.152
	JCMS	28	5.39	0.24	0.86	0.58	17.55	2.95E-11*
	JV3	29	1.76	0.07	2.24	.0036*	359.85	<2E-16*
acdh-1	OP50	58	6.55	0.34	2.2985	<2E-16*	NA	
(ok1489)	K279a	39	8.22	0.59	1.3757	0.2077	0.55	0.0276
	JCMS	29	5.48	0.2	0.7439	0.258	1.73	1.67E-15*
	JV3	28	2.46	0.1	1.06	0.799	58.33	<2E-16*
sodh-1	OP50	23	11.26	0.52	0.78	0.48	NA	
(ok2799)	K279a	26	9.62	0.59	0.87	0.61	1.61	0.108
	JCMS	28	4.57	0.26	1.36	0.25	16.86	2.9E-12*
	JV3	26	2.62	0.15	1.02	0.93	163.65	<2E-16*
pho-1	OP50	25	12.52	0.63	1.37	0.269	NA	
(tm5302)	K279a	25	12.3	0.79	0.87	0.615	0.57	0.0697
	JCMS	28	5.36	0.23	2.94	5.95E-4*	16.90	2.48E-10*
	JV3	32	2.55	0.17	1.17	0.554	234.60	<2E-16*
C55A6.7	OP50	28	9.14	0.5	0.82	0.46	NA	
(tm6807)	K279a	29	10.5	0.35	0.31	6.93E-5*	0.30	9.2E-5*
	JCMS	29	5.24	0.21	1.00	0.98	10.93	2.69E-9*
	JV3	29	1.99	0.07	1.34	0.28	360.03	<2E-16*
acox-1.4	OP50	30	8.63	0.34	1.19	0.507	NA	
(tm6415)	K279a	30	5.95	0.49	2.78	.00029*	2.83	.00011*
	JCMS	29	4.14	0.24	2.10	.0085*	15.57	5.22E-14*
	JV3	29	1.75	0.05	2.67	.00065*	205.48	<2E-16*
dhs-3	OP50	27	11.09	0.73	1.42	0.198	NA	
(tm6151)	K279a	30	10.47	0.87	1.33	0.286	0.95	0.86
	JCMS	26	4.77	0.23	4.30	8.7E-6*	8.00	1.37E-8*
	JV3	27	1.67	0.09	4.65	4.7E-7*	399.60	<2E-16*

JSR genes

F13D12.6	OP50	26	11.08	0.51	0.68	0.17	NA	
(tm7051)	K279a	29	9.76	0.53	1.25	0.00154*		
	JCMS	27	5.41	0.15	0.93	0.863		
	JV3	28	2.05	0.14	1.00	0.972		
dhs-2	OP50	27	8.7	0.61	1.18	1.26E-05*	NA	
(tm7516)	K279a	53	7.62	0.40	2.11	<2E-16*	1.62	<2E-16*
	JCMS	53	4.91	0.20	1.15	0.625	5.80	<2E-16*
	JV3	53	1.64	0.09	1.93	.0115*	139.04	<2E-16*
nhr-110	OP50	30	11.77	0.39	0.46	<2E-16*	NA	
(gk987)	K279a	58	8.97	0.40	1.11	0.566	2.14	<2E-16*
	JCMS	58	5.16	0.13	1.07	0.858	24.60	<2E-16*
	JV3	57	2.29	0.07	0.88	0.73	1006.00	<2E-16*
W02A2.8	OP50	29	10.86	0.57	0.66	0.125	NA	
(mh87)	K279a	29	9.83	0.65	0.81	0.434	0.99	0.971
	JCMS	30	4.77	0.23	2.13	.00985*	15.84	2.98E-13*
	JV3	28	1.93	0.09	1.91	.0205*	464.06	<2E-16*

Table 2.8 CPSR genes overlap with other *C. elegans* transcriptome datasets upon microbe exposure

WormExp was used to determine significant overlap of the 145 CPSR genes with differentially expressed genes upon exposure to other microbes. Exp. count indicates the number of differentially expressed genes in each experiment, and Input count indicates the number of CPSR genes that overlap with the Exp. Count genes. Experiments that resulted in enrichment of gene lists with FDR <0.05 are shown.

Experiment	Exp. Count	Input Count	FDR
UP by PA14, 24h	534	107	9.46E-152
UP by B. thuringiensis at 6h (BT247, 1:2) (Yang)	644	91	5.67E-110
UP by PA14, 12h	285	75	1.97E-109
UP by B. thuringiensis (BT247), 24h	1288	106	1.55E-107
UP by B. thuringiensis (BT247), 12h	1240	104	8.06E-106
UP by B. thuringiensis at 12h (BT247, 1:2) (Yang)	1058	98	2.52E-102
UP by PA14 (Miller)	236	65	7.00E-95
UP by B. thuringiensis at 6h (BT247, 1:10) (Yang)	751	85	2.25E-93
UP on S. marcescens	1238	95	1.08E-90
UP exposed to Vibrio cholerae with intact hemolysin A gene	267	56	2.96E-73
UP exposed to Vibrio cholerae E7946	276	56	2.01E-72
UP by B. thuringiensis at 12h (BT247, 1:10) (Yang)	1415	86	1.50E-71
UP exposed to gacA Mutant vs. Wild-Type P. aeruginosa Strain PA14	196	51	2.19E-71
UP on X. nematophila	745	62	1.58E-56
UP by P. aeruginosa PA14 (Head)	123	38	4.68E-55
PA14 Infection induced	195	42	3.99E-54
UP to PA14	258	45	5.34E-54
UP in Slow Killing, P. aeruginosa PA14	157	39	1.40E-52
UP by Bt toxin,Cry5B	369	45	6.98E-47
UP S. marcescens 24h, RNASeq	2263	77	3.27E-44
UP P.luminescens 24h, RNASeq	3017	85	8.56E-44
UP S. marcescens 24h, TillingArray	2649	73	2.57E-35
UP P.luminescens 24h, TillingArray	3782	84	2.57E-35
UP Harposporium 24h, RNASeq	2062	66	5.62E-35
UP by B. licheniformis 141	1484	55	2.80E-31
UP infected by PA14 octr-1 mutant vs. N2	1028	48	4.08E-31
UP by hemolysin mutant of Vibrio cholerae E7496	128	25	5.89E-30
Responding to P. aeruginosa	354	33	9.19E-30
UP by B. thuringiensis (BT247), C. elegans strain MY15	232	28	5.84E-28
UP treated by RPW-24	267	29	9.29E-28
UP exposed to Vibrio cholerae without hemolysin A gene	114	23	9.29E-28
UP by P. aeruginosa (Bond)	258	28	9.91E-27

UP by B. thuringiensis (BT247), C. elegans strain MY18	155	23	1.12E-24
UP Y. pestis vs. E. coli	98	19	2.71E-22
UP by B. thuringiensis (BT247), proteomics	171	21	1.00E-20
UP by P. aeruginosa PA14 at 12h	104	18	3.50E-20
UP by P. aeruginosa PA14 (McEwan)	66	16	3.99E-20
Changed Y. pestis vs. E. coli	255	21	2.40E-17
down by N. parisii at 40h	185	19	2.79E-17
down by N. parisii at 64h	348	23	4.21E-17
UP by B. thuringiensis (BT247), C. elegans strain N2	160	18	5.55E-17
UP PMK-1 and confirmed for resistance to PA14	38	12	3.75E-16
Changed infected by Microbacterium nematophilum	100	15	9.62E-16
UP by P. luminescens (Wong)	640	27	1.33E-15
UP by N. parisii at 64h	79	13	5.33E-14
down by C. albicans (Pukkila-Worley)	179	16	1.59E-13
UP by heat-killed C. albicans (Pukkila-Worley)	121	14	3.73E-13
UP in Liquid Killing, P. aeruginosa PA14	167	15	1.19E-12
UP by S. aureus (Visvikis)	821	26	3.02E-12
UP by N. parisii at 40h	64	11	6.90E-12
UP by S. aureus, dependent on hlh-30 (Visvikis)	633	22	5.17E-11
UP by N. parisii, array	80	11	6.43E-11
UP infected by S. aureus 8H	187	14	8.48E-11
UP infected by Microbacterium nematophilum	88	11	1.67E-10
Confirmed infected by PA14 in wildtype	26	8	3.02E-10
Responding to M. nematophilum AND P. aeruginosa NOT S. Aureus	14	7	3.04E-10
UP by N. parisii at 30h	103	11	7.68E-10
Changed infected to S. aureus and exposure to B. thuringiensis PFT	19	7	2.51E-09
Down D. coniospora 12h, RNASeq	2280	36	7.44E-09
UP by biofilms of Y. pseudotuberculosis at 1h	187	12	1.67E-08
down by MC-BA (M. aeruginosa)	1328	26	5.00E-08
Down infected by X. Nematophila	620	18	7.83E-08
UP Bacillus strain 67	412	15	1.15E-07
UP by C. albicans (Pukkila-Worley)	119	9	8.65E-07
Responding to S. aureus AND P. aeruginosa NOT M. Nematophilum	34	6	3.91E-06
down by PA14, 12h	109	8	6.77E-06
down by S. aureus (Bond)	291	11	1.07E-05
response to PA14	20	5	1.40E-05
UP D. coniospora 12h, RNASeq	1516	23	2.75E-05
UP by virus Orsay	53	6	3.37E-05
Down infected by S. Aureus 8H	198	9	3.43E-05
UP on S. aureus	266	10	3.76E-05
down by P. aeruginosa (Bond)	206	9	4.50E-05
down by heat-killed C. albicans (Pukkila-Worley)	149	8	4.85E-05
down by virus Orsay	61	6	6.48E-05

UP fed by E. Coli HT115 vs. OP50	31	5	7.93E-05
Down fed with L. rhamnosus CNCM I-3690 vs. OP50,3 days	756	15	1.13E-04
down by C. albicans vs. heat-killed C. albicans (Pukkila-Worley)	36	5	1.40E-04
UP on B. thuringiensis DB27	268	9	2.65E-04
down by B. thuringiensis (BT247), 24h	524	12	2.76E-04
UP by E. faecalis (Wong)	625	13	2.96E-04
down by N. parisii at 30h	25	4	1.10E-03
Confirmed infected by PA14 in spe-26 mutant	26	4	1.23E-03
UP fed with L. rhamnosus CNCM I-4137 vs. OP50,10 days	118	6	1.32E-03
UP by N. parisii at 8h	123	6	1.57E-03
Down infected by S. Aureus	192	7	1.66E-03
Responding to S. aureus AND P. aeruginosa AND M. Nematophilum	10	3	4.20E-03
down by P. aeruginosa PA14 at 12h	42	4	4.60E-03
UP by E. coli strain LF82 at 72h, proteomics	94	5	4.80E-03
Down infected by S. Marcescens	165	6	5.36E-03
UP after 8h infection with S. aureus	46	4	5.89E-03
UP E. faecalis 24h, TillingArray	4051	34	8.22E-03
UP infected with hly(+) V. cholerae strains	54	4	9.09E-03
down by B. thuringiensis at 12h (BT247, 1:10) (Yang)	1208	15	9.81E-03
Down PA14 vs. OP50	121	5	1.14E-02
down by PA14, 24h	608	10	1.21E-02
down D. coniospora 12 h (cDNA)	209	6	1.40E-02
Down D. coniospora 12h, cDNAArray	209	6	1.40E-02
UP by N. parisii at 16h	68	4	1.67E-02
down by B. thuringiensis at 6h (BT247, 1:10) (Yang)	890	12	1.67E-02
down by B. thuringiensis at 12h (BT247, 1:2) (Yang)	1044	13	1.97E-02
UP Micrococcus luteus vs. OP50	26	3	2.53E-02
UP to soil bacteria vs OP50	27	3	2.70E-02
down by B. thuringiensis at 6h (BT247, 1:2) (Yang)	841	11	3.04E-02
Down S. marcescens 24h, cDNAArray	167	5	3.21E-02

Table 2.9 CPSR genes ranked based on how many pathogens result in differential expression

Genes are listed in order based on the number of pathogens that result in differential expression of that gene, from lowest-ranking to highest-ranking. N= number of pathogens, and pathogens that result in differential expression of that gene are listed.

Gene	# of pathogens	Pathogens
K08D8.12	1	*B. thuringiensis*
nhr-220	1	*Harposporium*
F27E5.9	2	*B. thuringiensis, P. aeruginosa*
bath-47	2	*B. licheniformis, B. thuringiensis*
linc-40	2	*B. thuringiensis, P. aeruginosa*
Y37H2B.1	2	*E. faecalis, S. marcescens*
Y54G2A.28	2	*B. thuringiensis, M. aeruginosa*
C45B11.2	3	*B. thuringiensis, P. aeruginosa, X. nematophila*
arl-7	3	*B. thuringiensis, P. aeruginosa, X. nematophila*
W05H9.3	3	*B. licheniformis, S. aureus, S. marcescens*
Y37F4.8	3	*B. thuringiensis, D. coniospora, E. faecalis*
Y82E9BL.18	3	*B. thuringiensis, D. coniospora, P. aeruginosa*
Y82E9BR.5	3	*B. thuringiensis, P. aeruginosa, S. marcescens*
fbxa-77	3	*B. licheniformis, M. aeruginosa, S. marcescens*
ZK1055.7	4	*B. thuringiensis, P. aeruginosa, P. luminescens, S. marcescens*
fbxa-80	4	*B. thuringiensis, Harposporium, P. aeruginosa, S. marcescens*
nhr-112	5	*B. thuringiensis, P. aeruginosa, P. luminescens, S. marcescens, X. nematophila*
dod-23	5	*B. thuringiensis, E. faecalis, L. rhamnosus, N. parisii, S. marcescens*
nhr-63	5	*B. thuringiensis, P. aeruginosa, P. luminescens, S. marcescens, X. nematophila*
F08G2.5	5	*B. thuringiensis, D. coniospora, N. parisii, P. aeruginosa, P. luminescens*
Y69A2AR.25	5	*B. thuringiensis, L. rhamnosus, N. parisii, S. marcescens, V. cholerae*
Y39A3A.4	5	*B. thuringiensis, D. coniospora, P. aeruginosa, P. luminescens, X. nematophila*
F43C1.7	5	*B. thuringiensis, P. aeruginosa, P. luminescens, S. marcescens, X. nematophila*
Y113G7B.14	5	*B. thuringiensis, C. albicans, P. aeruginosa, S. marcescens, X. nematophila*
twk-28	5	*B. thuringiensis, Harposporium, P. aeruginosa, P. luminescens, S. marcescens*
C25F9.11	5	*B. licheniformis, B. thuringiensis, S. aureus, S. marcescens, X. nematophila*
srd-64	5	*B. thuringiensis, D. coniospora, P. aeruginosa, P. luminescens, S. marcescens*
fbxa-55	5	*E. carotovora, Harposporium, S. aureus, S. marcescens, X. nematophila*
fbxa-74	6	*B. thuringiensis, Harposporium, P. aeruginosa, P. luminescens, S. marcescens, X. nematophila*
oac-20	6	*B. licheniformis, B. thuringiensis, Bacillus* 67*, Harposporium, P. aeruginosa, S. marcescens*

F33H12.7	6	*B. licheniformis, B. thuringiensis, N. parisii, P. aeruginosa, S. marcescens, X. nematophila*
C49G7.7	6	*B. thuringiensis, D. coniospora, E. faecalis, P. aeruginosa, S. marcescens, X. nematophila*
Y94H6A.10	6	*B. thuringiensis, D. coniospora, Harposporium, P. aeruginosa, P. luminescens, S. marcescens*
clec-232	6	*B. licheniformis, B. thuringiensis, Bacillus* 67, *D. coniospora, S. aureus, S. marcescens*
T23F11.6	6	*B. licheniformis, B. thuringiensis, E. faecalis, P. aeruginosa, S. marcescens, X. nematophila*
odc-1	6	*B. licheniformis, B. thuringiensis, D. coniospora, P. aeruginosa, S. aureus, X. nematophila*
cyp-14A2	6	*B. licheniformis, B. thuringiensis, E. faecalis, N. parisii, P. aeruginosa, S. marcescens*
fbxa-30	6	*B. thuringiensis, D. coniospora, E. faecalis, Harposporium, P. aeruginosa, P. luminescens*
numr-1	6	*B. licheniformis, N. parisii, P. aeruginosa, P. luminescens, S. marcescens, X. nematophila*
C49C3.9	7	*B. thuringiensis, Harposporium, L. rhamnosus, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae*
clec-41	7	*B. licheniformis, B. thuringiensis, N. parisii, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae*
C49G7.10	7	*B. thuringiensis, Harposporium, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, X. nematophila*
Y22D7AL.15	7	*B. licheniformis, B. thuringiensis, P. aeruginosa, S. aureus, S. marcescens, Y. pestis, Y. pseudotuberculosis*
Y58A7A.3	7	*B. thuringiensis, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae, X. nematophila*
C49G7.12	7	*B. thuringiensis, E. faecalis, Harposporium, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae*
ugt-31	7	*B. thuringiensis, Harposporium, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, X. nematophila*
M01G12.9	7	*B. thuringiensis, D. coniospora, E. faecalis, N. parisii, P. aeruginosa, P. luminescens, X. nematophila*
T05F1.9	7	*B. licheniformis, B. thuringiensis, E. faecalis, Harposporium, P. aeruginosa, P. luminescens, S. marcescens*
nhr-50	7	*B. thuringiensis, D. coniospora, M. aeruginosa, P. aeruginosa, P. luminescens, S. marcescens, X. nematophila*
C18H7.11	7	*B. thuringiensis, E. faecalis, Harposporium, P. aeruginosa, P. luminescens, S. marcescens, X. nematophila*
K11G9.3	7	*B. thuringiensis, D. coniospora, E. faecalis, P. aeruginosa, P. luminescens, S. aureus, S. marcescens*
Y9C9A.8	7	*B. thuringiensis, D. coniospora, Harposporium, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae*
clec-66	8	*B. licheniformis, B. thuringiensis, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, X. nematophila, Y. pestis*

C06B3.7	8	*B. thuringiensis, D. coniospora, Harposporium, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, X. nematophila*
dct-17	8	*B. licheniformis, B. thuringiensis, D. coniospora, Harposporium, M. nematophilum, P. aeruginosa, P. luminescens, S. marcescens*
Y46D2A.2	8	*B. thuringiensis, C. albicans, D. coniospora, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, Y. pestis*
F19B2.5	8	*B. thuringiensis, D. coniospora, E. faecalis, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, X. nematophila*
Y47H10A.5	8	*B. thuringiensis, N. parisii,* Orsay virus, *P. aeruginosa, P. luminescens, S. aureus, S. marcescens, X. nematophila*
F14F9.4	8	*B. thuringiensis, E. faecalis, Harposporium, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, X. nematophila*
F25A2.1	8	*B. thuringiensis, D. coniospora, E. faecalis, Harposporium, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae*
K09D9.1	8	*B. thuringiensis, E. faecalis, Harposporium, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, X. nematophila*
cpr-5	8	*B. licheniformis, B. thuringiensis, E. faecalis, L. rhamnosus, P. aeruginosa, P. luminescens, S. aureus, X. nematophila*
Y58A7A.4	8	*B. thuringiensis, Harposporium, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae, X. nematophila*
F55G11.8	8	*B. licheniformis, B. thuringiensis, L. rhamnosus, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, Y. pestis*
Y75B8A.28	8	*B. thuringiensis, D. coniospora, E. faecalis, Harposporium, P. aeruginosa, P. luminescens, S. aureus, S. marcescens*
C08E8.4	8	*B. licheniformis, B. thuringiensis, Harposporium, L. rhamnosus, M. aeruginosa, P. aeruginosa, P. luminescens, V. cholerae*
clec-83	8	*B. thuringiensis, L. rhamnosus, M. nematophilum, N. parisii, P. aeruginosa, S. aureus, S. marcescens, Y. pestis*
C34H4.2	8	*B. thuringiensis, Harposporium, L. rhamnosus, M. aeruginosa, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae*
cutl-18	8	*B. thuringiensis, D. coniospora, Harposporium, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, X. nematophila*
fbxa-163	8	*D. coniospora, Harposporium, L. rhamnosus, N. parisii, P. aeruginosa, P. luminescens, S. aureus, S. marcescens*
hsp-17	8	*B. licheniformis, B. thuringiensis, Harposporium, N. parisii, P. aeruginosa, P. luminescens, S. marcescens, X. nematophila*
fbxa-37	8	*B. thuringiensis, D. coniospora, E. faecalis, Harposporium, P. aeruginosa, P. luminescens, S. marcescens, X. nematophila*
B0024.4	9	*B. thuringiensis, M. aeruginosa, M. nematophilum, N. parisii,* Orsay virus, *P. aeruginosa, P. luminescens, S. marcescens, V. cholerae*
clec-265	9	*B. thuringiensis, C. albicans, D. coniospora, N. parisii, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae*
ttr-44	9	*B. thuringiensis, E. faecalis, N. parisii, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, X. nematophila, Y. pseudotuberculosis*
dod-19	9	*B. thuringiensis, D. coniospora, E. coli* LF82, *Harposporium, L. rhamnosus, M. aeruginosa, P. aeruginosa, P. luminescens, S. marcescens*

F20G2.5	9	*B. licheniformis, B. thuringiensis, D. coniospora, Harposporium, M. aeruginosa, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae*
clec-4	9	*B. thuringiensis, E. faecalis, Harposporium, M. aeruginosa, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae*
cyp-35A2	9	*B. licheniformis, B. thuringiensis, C. albicans, E. faecalis, N. parisii, P. aeruginosa, S. aureus, S. marcescens, X. nematophila*
C10C5.2	9	*B. thuringiensis, E. faecalis, Harposporium, M. aeruginosa, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, X. nematophila*
clec-85	9	*B. thuringiensis, D. coniospora, E. coli* LF82, *L. rhamnosus, N. parisii, P. aeruginosa, Pseudomonas, S. marcescens, V. cholerae*
T04F8.7	9	*B. licheniformis, B. thuringiensis, Bacillus* 67, *D. coniospora, P. aeruginosa, S. aureus, S. marcescens, X. nematophila, Y. pseudotuberculosis*
swt-6	9	*B. thuringiensis, E. carotovora, E. faecalis, N. parisii, P. aeruginosa, S. aureus, S. marcescens, V. cholerae, X. nematophila*
R07C12.1	9	*B. thuringiensis, L. rhamnosus, M. nematophilum, N. parisii, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, Y. pseudotuberculosis*
nuc-1	9	*B. licheniformis, B. thuringiensis, D. coniospora, N. parisii, P. luminescens, S. aureus, S. marcescens, X. nematophila, Y. pestis*
Y51H4A.25	9	*B. thuringiensis, Bacillus* 67, *D. coniospora, E. faecalis, Harposporium, P. aeruginosa, P. luminescens, S. aureus, S. marcescens*
lys-10	9	*B. licheniformis, B. thuringiensis, Bacillus* 67, *D. coniospora, E. faecalis, M. luteus, P. luminescens, Pseudomonas, S. aureus*
tba-7	9	*B. licheniformis, B. thuringiensis, E. faecalis, Harposporium, M. aeruginosa, P. aeruginosa, P. luminescens, S. marcescens, X. nematophila*
W02A2.9	9	*B. thuringiensis, C. albicans, D. coniospora, E. faecalis, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, X. nematophila*
fbxa-60	9	*B. licheniformis, B. thuringiensis, Harposporium, N. parisii, P. aeruginosa, S. aureus, S. marcescens, V. cholerae, X. nematophila*
F15B9.6	9	*B. licheniformis, B. thuringiensis, D. coniospora, E. faecalis, L. rhamnosus, N. parisii, P. aeruginosa, P. luminescens, S. marcescens*
Y94H6A.2	9	*B. thuringiensis, D. coniospora, E. faecalis, Harposporium, N. parisii,* Orsay virus, *P. aeruginosa, P. luminescens, S. marcescens*
C08F11.13	9	*B. licheniformis, B. thuringiensis, E. faecalis, N. parisii, P. aeruginosa, S. marcescens, V. cholerae, X. nematophila, Y. pseudotuberculosis*
T24C4.4	10	*B. thuringiensis, D. coniospora, E. carotovora, E. faecalis, Harposporium, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, X. nematophila*
oac-31	10	*B. thuringiensis, C. albicans, E. faecalis, Harposporium, M. nematophilum, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, X. nematophila*
K11H12.4	10	*B. licheniformis, B. thuringiensis, D. coniospora, Harposporium, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae, X. nematophila*
lec-11	10	*B. licheniformis, B. thuringiensis, E. faecalis, Harposporium, L. rhamnosus, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, Y. pestis*
tsp-1	10	*B. licheniformis, B. thuringiensis, D. coniospora, Harposporium, N. parisii, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, X. nematophila*

spp-1	10	*B. licheniformis, B. thuringiensis, M. aeruginosa, N. parisii, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae, X. nematophila*
F35E12.9	10	*B. licheniformis, B. thuringiensis, D. coniospora, Harposporium,* Orsay virus*, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, X. nematophila*
lys-3	10	*B. thuringiensis, D. coniospora, E. carotovora, E. faecalis, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae, X. nematophila*
valv-1	10	*B. licheniformis, B. thuringiensis, C. albicans, E. faecalis, Harposporium, N. parisii, P. aeruginosa, P. luminescens, S. aureus, S. marcescens*
Y58A7A.5	10	*B. thuringiensis, D. coniospora, E. faecalis, Harposporium, N. parisii,* Orsay virus, *P. aeruginosa, P. luminescens, S. marcescens, X. nematophila*
cyp-35A5	10	*B. thuringiensis, Bacillus* 67, *C. albicans, E. faecalis, N. parisii, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, X. nematophila*
clec-218	10	*B. licheniformis, B. thuringiensis, Bacillus* 67*, C. albicans, N. parisii, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, X. nematophila*
E02C12.8	10	*B. thuringiensis, E. faecalis, Harposporium, M. aeruginosa, N. parisii, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, X. nematophila*
fbxa-59	10	*B. licheniformis, B. thuringiensis, Bacillus* 67, *C. albicans, E. faecalis, Harposporium, P. aeruginosa, S. aureus, S. marcescens, X. nematophila*
K11H12.3	10	*B. thuringiensis, D. coniospora, E. carotovora, E. faecalis, Harposporium, M. aeruginosa, P. aeruginosa, P. luminescens, S. marcescens, X. nematophila*
T01D3.6	11	*B. licheniformis, B. thuringiensis, C. albicans, N. parisii, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae, X. nematophila, Y. pestis*
lys-2	11	*B. thuringiensis, Bacillus* 67*, C. albicans, E. faecalis, Harposporium, M. aeruginosa, N. parisii, P. aeruginosa, P. luminescens, S. aureus, S. marcescens*
oac-6	11	*B. licheniformis, B. thuringiensis, Bacillus* 67*, Harposporium, M. nematophilum, N. parisii, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, X. nematophila*
K08D8.5	11	*B. thuringiensis, C. albicans, E. coli* HT115*, Harposporium, L. rhamnosus, N. parisii, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, Y. pestis*
oac-14	11	*B. thuringiensis, C. albicans, D. coniospora, M. aeruginosa, M. nematophilum, N. parisii, P. aeruginosa, S. aureus, S. marcescens, V. cholerae, X. nematophila*
tsp-2	11	*B. licheniformis, B. thuringiensis, D. coniospora, E. faecalis, Harposporium, N. parisii,* Orsay virus*, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae*
F53B2.8	11	*B. licheniformis, B. thuringiensis, Harposporium, M. aeruginosa, N. parisii, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae, X. nematophila*
K08D8.4	11	*B. thuringiensis, D. coniospora, E. faecalis, Harposporium,* Orsay virus*, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae, X. nematophila*
F53C11.1	11	*B. thuringiensis, C. albicans, D. coniospora, E. coli* LF82*, Harposporium, L. rhamnosus, M. aeruginosa, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae*

C17H12.6	11	*B. thuringiensis, D. coniospora, E. coli* HT115*, E. faecalis, Harposporium, L. rhamnosus, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae*
lec-10	11	*B. thuringiensis, C. albicans, D. coniospora, E. faecalis, Harposporium, L. rhamnosus, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, X. nematophila*
F35E12.4	11	*B. licheniformis, B. thuringiensis, D. coniospora, E. faecalis, Harposporium, M. nematophilum, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, X. nematophila*
ugt-53	11	*B. licheniformis, B. thuringiensis, C. albicans, D. coniospora, E. faecalis, Harposporium, N. parisii, P. aeruginosa, S. aureus, S. marcescens, X. nematophila*
clec-70	11	*B. licheniformis, B. thuringiensis, C. albicans, E. faecalis, Harposporium, M. nematophilum, P. aeruginosa, S. aureus, S. marcescens, V. cholerae, X. nematophila*
T19D12.4	12	*B. thuringiensis, C. albicans, D. coniospora, Harposporium, L. rhamnosus, M. aeruginosa, N. parisii, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, X. nematophila*
T24E12.5	12	*B. thuringiensis, Bacillus* 67, *E. faecalis, M. aeruginosa, M. nematophilum, N. parisii,* Orsay virus*, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, X. nematophila*
lys-1	12	*B. thuringiensis, Bacillus* 67, *D. coniospora, E. carotovora, E. faecalis, L. rhamnosus, M. luteus, N. parisii, P. aeruginosa, P. luminescens, S. marcescens, X. nematophila*
ZK6.11	12	*B. licheniformis, B. thuringiensis, Bacillus* 67*, D. coniospora, E. coli* LF82, *Harposporium, L. rhamnosus, M. aeruginosa, P. aeruginosa, P. luminescens, S. aureus, S. marcescens*
cpr-4	12	*B. animalis, B. thuringiensis, D. coniospora, E. faecalis, M. aeruginosa, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, X. nematophila, Y. pestis, Y. pseudotuberculosis*
cpr-3	13	*B. licheniformis, B. thuringiensis, Bacillus* 67, *C. albicans, D. coniospora, L. rhamnosus, M. aeruginosa,* Orsay virus, *P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, Y. pestis*
clec-67	13	*B. licheniformis, B. thuringiensis, C. albicans, E. coli* HT115*, E. faecalis, M. nematophilum, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae, X. nematophila, Y. pestis*
mtl-1	13	*B. licheniformis, B. megaterium, B. thuringiensis, C. albicans, D. coniospora, E. faecalis, N. parisii, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae, X. nematophila*
hpo-6	13	*B. thuringiensis, C. albicans, D. coniospora, E. coli* LF82*, Harposporium, M. nematophilum, N. parisii, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, X. nematophila, Y. pestis*
ugt-18	13	*B. licheniformis, B. thuringiensis, Bacillus* 67*, C. albicans, D. coniospora, Harposporium, M. nematophilum, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae, X. nematophila*

irg-4	13	*B. animalis, B. licheniformis, B. thuringiensis, L. rhamnosus, M. aeruginosa, N. parisii,* Orsay virus*, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae, Y. pestis*
F53A9.8	14	*B. licheniformis, B. thuringiensis, Bacillus* 67*, C. albicans, D. coniospora, E. faecalis, M. nematophilum, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, X. nematophila, Y. pestis, Y. pseudotuberculosis*
kreg-1	14	*B. licheniformis, B. thuringiensis, Bacillus* 67*, C. albicans, D. coniospora, E. faecalis, Harposporium, M. aeruginosa, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae, X. nematophila*
T24B8.5	14	*B. licheniformis, B. megaterium, B. thuringiensis, E. carotovora, E. coli* HT115*, E. faecalis, Harposporium, M. aeruginosa, N. parisii, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae*
ech-9	14	*B. licheniformis, B. thuringiensis, Bacillus* 67*, C. albicans, D. coniospora, Harposporium, M. aeruginosa, M. nematophilum, N. parisii, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae*
vit-1	14	*B. licheniformis, B. thuringiensis, Bacillus* 67*, C. albicans, D. coniospora, E. faecalis, M. luteus, N. parisii, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae, X. nematophila*
scl-2	14	*B. licheniformis, B. thuringiensis, C. albicans, D. coniospora, Harposporium, L. rhamnosus, M. nematophilum, N. parisii, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, X. nematophila, Y. pseudotuberculosis*
M02H5.8	14	*B. licheniformis, B. thuringiensis, Bacillus* 67*, C. albicans, D. coniospora, E. faecalis, Harposporium, M. aeruginosa, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, Y. pestis, Y. pseudotuberculosis*
dod-22	15	*B. licheniformis, B. thuringiensis, C. albicans, D. coniospora, E. carotovora, E. coli* HT115*,* Orsay virus*, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae, X. nematophila, Y. pestis, Y. pseudotuberculosis*
C17H12.8	15	*B. animalis, B. thuringiensis, C. albicans, E. carotovora, Harposporium, L. rhamnosus, M. aeruginosa, M. nematophilum,* Orsay virus*, P. aeruginosa, P. luminescens, S. marcescens, V. cholerae, X. nematophila, Y. pestis*
mul-1	15	*B. licheniformis, B. thuringiensis, C. albicans, D. coniospora, E. faecalis, L. rhamnosus, M. nematophilum, N. parisii, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae, X. nematophila, Y. pestis*
F53A9.1	16	*B. licheniformis, B. thuringiensis, Bacillus* 67*, D. coniospora, E. faecalis, Harposporium, M. nematophilum, N. parisii, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae, X. nematophila, Y. pestis, Y. pseudotuberculosis*
F53A9.6	16	*B. licheniformis, B. thuringiensis, C. albicans, D. coniospora, E. faecalis, Harposporium, M. aeruginosa, M. nematophilum, N. parisii, P. aeruginosa, P. luminescens, S. aureus, S. marcescens, V. cholerae, Y. pestis, Y. pseudotuberculosis*

Table 2.10 Cox proportional hazard ratios for common pathogen and *S. maltophilia*-specific genes

Mean survival (M), standard error of the mean (SE), and sample size (N), are given for each allele and bacterial treatment combination. Bacterial treatments include *S. maltophilia* JCMS, *S. maltophilia* K279a, *S. maltophilia* JV3, *S. aureus* NCTC8325, *S. marcescens* DB10, *P. aeruginosa* PA14, and *X. nematophila* X1462. *Wild-type* statistics were determined by combining all *wild-type* data from all experiments. Hazard ratios indicate the treatment hazard divided by the hazard of *wild-type*. The hazard is defined as the probability of a nematode dying at a given time. Hazard ratios and associated p-values for each comparison were determined using Cox proportional hazards tests in R. Asterisk indicates significant p-values ($p<0.05$). *S. maltophilia* data not shown can be found in Table 2.7.

						Relative to *wild-type*	
	Nematode	Bacteria	N	M	SE	Hazard Ratio	p value
	wild-type	K279a	60	7.43	0.33	NA	
		JCMS	61	5.89	0.30	NA	
		JV3	59	2.45	0.10	NA	
		NCTC8325	113	2.68	0.13	NA	
		PA14	114	3.10	0.13	NA	
		DB10	112	3.24	0.15	NA	
		X-1462	127	1.86	0.07	NA	
Common pathogen genes	*dod-22 (ok2679)*	NCTC8325	27	5.70	0.44	0.27	.00045*
		PA14	31	4.55	0.17	0.46	.024*
		DB10	21	5.10	0.48	0.37	.00179*
		X-1462	30	2.80	0.17	0.41	.0056*
	T24B8.5 (ok3236)	PA14	30	4.57	0.22	0.42	.0135*
		X-1462	31	2.90	0.16	0.37	.002*
	scl-2 (tm2428)	NCTC8325	29	3.69	0.44	0.44	.00757*
		PA14	32	2.97	0.26	1.31	0.392
		DB10	33	4.09	0.11	0.69	0.326
		X-1462	34	2.21	0.15	0.88	0.615
	clec-67 (ok2770)	NCTC8325	28	5.29	0.38	0.32	.0005*
		PA14	31	4.52	0.21	0.44	.017*
		DB10	24	6.08	0.39	0.28	5.24E-5*
		X-1462	31	3.10	0.18	0.29	.00011*
	ZK6.11 (ok3738)	NCTC8325	30	4.27	0.32	0.35	.00052*
		PA14	30	1.90	0.18	2.81	.00602*
		DB10	30	3.37	0.22	1.00	9.87E-01
		X-1462	30	1.60	0.09	1.56	0.1998

	cpr-4	K279a	36	5.56	0.31	1.46	0.0815
	(ok3413)	JCMS	34	5.00	0.14	2.36	.026*
		JV3	35	1.89	0.05	3.10	.0293*
		NCTC8325	31	5.35	0.35	0.32	.00011*
		PA14	35	4.26	0.17	0.67	0.114
		DB10	30	3.37	0.17	1.12	0.77
		X-1462	35	1.74	0.08	1.59	0.15
***S. maltophilia* specific genes**	*fbxa-77*	K279a	29	9.48	0.74	1.08	0.768
	(mh94)	JCMS	30	4.47	0.26	2.59	.00072*
		JV3	30	2.00	0.07	1.56	0.1
		NCTC8325	3	4.05	0.18	0.43	.00371*
		PA14	32	2.97	0.16	1.33	0.4563
		DB10	36	3.53	0.18	0.95	0.883
		X-1462	34	1.83	0.09	1.09	0.79
	K08D8.12	K279a	28	9.21	0.80	0.72	0.225
	(mh101)	JCMS	29	5.34	0.37	1.30	0.328
		JV3	28	2.00	0.08	1.54	0.116
		NCTC8325	35	3.68	0.20	0.49	.01706*
		PA14	35	3.56	0.22	0.74	0.3271
		DB10	32	3.32	0.16	1.11	0.772
		X-1462	35	1.87	0.15	1.10	0.78
	C25F9.11	K279a	29	10.40	0.81	0.85	0.541
	(mh97)	JCMS	29	5.45	0.27	1.44	0.1798
		JV3	30	1.90	0.08	1.75	0.0381
		NCTC8325	28	2.57	0.18	1.26	0.44
		PA14	34	2.66	0.12	1.94	0.1184
		DB10	23	3.39	0.22	1.08	0.83
		X-1462	34	1.83	0.09	1.18	0.628
	Y82E9BR.5	K279a	29	8.48	0.73	1.27	0.357
	(mh93)	JCMS	29	4.76	0.14	2.57	.00195*
		JV3	30	2.18	0.10	1.19	0.5
		NCTC8325	35	3.51	0.17	0.59	0.0647
		PA14	34	4.00	0.21	0.56	1.18
		DB10	34	3.14	0.16	1.36	0.425
		X-1462	34	1.96	0.11	0.99	0.965
	F08G2.5	K279a	30	10.30	0.75	1.10	0.72
	(mh86)	JCMS	31	6.06	0.24	1.08	0.76
		JV3	30	2.99	0.14	0.73	0.25
		NCTC8325	30	2.57	0.15	1.18	0.51
		PA14	30	1.67	0.10	4.96	.00038*
		DB10	30	2.93	0.24	1.27	0.489
		X-1462	31	1.42	0.09	2.00	.0488*

Chapter 3 - Comparative genomics of pathogenic and nonpathogenic *S. maltophilia* strains, including a novel environmental isolate *S. maltophilia* JCMS

Introduction

Stenotrophomonas maltophilia is a Gram-negative bacterium that is ubiquitous in the environment, most commonly found in aqueous sources and associated with plant roots (De Boer et al., 2001; Denton & Kerr, 1998; Jägevall, Rabe, & Pedersen, 2011). More recently, it has been isolated from water sources and medical devices within hospitals (Brooke, 2012). Although not highly virulent, it is capable of causing infection in immunocompromised patients, and is commonly associated with respiratory diseases such as cystic fibrosis and lung cancer (Bittar et al., 2008; Brooke, 2012; Chang, 2015). *S. maltophilia* is multi-drug resistant and capable of forming biofilms (Chang et al., 2015; Elting & Bodey, 1990; Jägevall et al., 2011), thus developing treatment methods for this pathogen has become of increasing concern. Although *S. maltophilia* can be detrimental to humans, it has beneficial implications in ecology and biotechnology. For example, *S. maltophilia* plays an important role in the rhizosphere where it enhances plant growth by fixing nitrogen (Park et al., 2005). In addition, its anti-microbial and metal resistance properties are promising for phytoremediation and other biotechnology applications (Berg & Martinez, 2015).

Phylogenetic and antibiotic profile analyses of environmental and clinical isolates revealed heterogeneity between strain origin (Berg, Roskot, & Smalla, 1999; Lira, Berg, & Martínez, 2017), suggesting that environmental isolates of *S. maltophilia* strains have pathogenic potential without acquiring new virulence mechanisms. Furthermore, plant-associated strains and

clinical strains often share similar genetic features, such as adhesion proteins, efflux pumps, and secreted enzymes (Adamek et al., 2011; Alavi et al., 2014; Berg & Martinez, 2015; Berg et al., 1999; Lira, Berg, & Martínez, 2017; Ryan et al., 2009). However, clinical isolates have higher mutation rates than environmental isolates, suggesting clinical isolates might be better able to adapt to stressors such as the immune response and antibiotics (Turrientes et al., 2010).

Although clinical and environmental isolates of *S. maltophilia* share common features, several lines of evidence suggest there are differences between strains. In fact, *S. maltophilia* strains are extremely genetically diverse, and it has even been proposed that *S. maltophilia* be divided into multiple species (Kaiser et al., 2009; Nicoletti et al., 2011; Valdezate et al., 2004). Genome sequence and proteomic analyses of multiple strains of *S. maltophilia* have also identified differences in virulence factors between strains, many contained on potentially mobile genomic islands (Adamek, Linke, & Schwartz, 2014; Ferrer-Navarro et al., 2013; Lira, Berg, & Martínez, 2017; Rocco et al., 2009). Because of these differences, it is not surprising that strains of *S. maltophilia* show varying pathogenicity to *Caenorhabditis elegans*, amoeba (*Dictyostelium discoideum* and *Acanthamoeba castellanii*), zebrafish (*Danio rerio*), and wax moths (*Galleria mellonella*) (Adamek et al., 2011; Ferrer-Navarro et al., 2013; Wh(Pompilio et al., 2016)ite et al., 2016).

While many studies have predicted the pathogenic potential of *S. maltophilia* strains based on genomic virulence factors, to date, no studies have correlated phenotypic pathogenicity to differences in virulence factors. To address this gap, we have previously developed *Caenorhabditis elegans* as a model for studying *S. maltophilia*-host interactions (White et al., 2016). *C. elegans* are bacterivores that are found in rotting organic matter and are constantly in contact with a variety of microorganisms, including bacteria. In fact, recent studies have found

that *Stenotrophomonas* is one of the most abundant genera of bacteria found in the native microbiome of *C. elegans* (Berg et al., 2016; Dirksen et al., 2016; Zhang et al., 2017). Thus using *C. elegans* as a model provides a realistic platform for understanding *S. maltophilia* pathogenesis. We have previously identified transcriptomic responses in *C. elegans* upon exposure to *S. maltophilia* strains with varying pathogenicity (Chapter 2 - ; White and Herman, unpublished). Overall, differences in pathogenicity caused differences in host responses, and increased virulence positively correlated with the expression of defense response genes (Chapter 2 -). However, to better understand the genetic basis of differences in virulence in *S. maltophilia* that lead to distinct host gene expression, we compared draft genome sequences of several *S. maltophilia* strains that vary in pathogenicity to *C. elegans*. One strain of particular interest in this study, *S. maltophilia* JCMS, was locally isolated in association with soil nematodes and is detrimental to the health of *C. elegans*. Therefore, one goal of this study was assembly of the JCMS genome to determine the genetic basis of JCMS pathogenesis. In addition, we sought to compare the draft genomes of several of these strains to their respective reference genome sequences to determine possible differences between the strains we sequenced and the strains used for the reference sequencing.

Methods

Nematode and Bacteria Strains

The *C. elegans N2* strain was obtained from the *Caenorhabditis* Genetics Center (CGC) and used as *wild-type*. Strains were maintained on nematode growth media (NGM) plates seeded with *E. coli* OP50. Bacteria strains include *E. coli* OP50 from the CGC, *S. maltophilia* K279a from R. Ryan (University College Cork), *S. maltophilia* R551-3 from D. van der Lelie (Brookhaven National Laboratory), *S. maltophilia* JV3 from J. Tiedje (Michigan State

University), and *S. maltophilia* JCMS isolated in association with nematodes from Konza Prairie near Manhattan, KS (White et al., 2016).

All bacteria strains were frozen at -80° C upon arrival to the lab and thawed frequently for experimentation. *S. maltophilia* is naturally Ampicillin resistant and was grown on Luria Broth (LB) agar containing 100 µg/mL Ampicillin to selectively isolate and maintain *S. maltophilia* strains and avoid contamination. *E. coli* OP50 was grown on normal LB agar. All plates were incubated at 37° C overnight. *S. maltophilia* strains were grown in liquid LB containing 100 µg/mL Ampicillin, and other strains were grown in liquid LB and shaken overnight at 37° C overnight. For *C. elegans* survival assays, liquid cultures were then seeded onto NGM plates and grown at room temperature overnight.

Bacterial DNA extraction and sequencing

Genomic DNA was isolated from *S. maltophilia* JCMS, JV3, K279a and R551-3 using phenol-chloroform extraction and RNA with digested with RNase A. Illumina sequencing libraries were prepared by the Kansas State University Integrated Genomics Facility (KSU IGF). Sequencing was performed using the Illumina MiSeq Benchtop Sequencer in the KSU IGF to obtain 250 bp paired-end sequences. This resulted in an average of 1.45 million sequences per strain totaling 3,174 Mb. Additional 100 single end sequences were obtained from the Illumina HiSeq 2500 Sequencer at the Genetic Sequencing Facility in the NIH COBRE Center for Molecular Analysis of Disease Pathways (CMADP) at KU. This resulted in an average of 4.49 million sequences per strain totaling 5,080 Mb.

Bacterial genome assembly

Adaptors were trimmed from sequences using bbduk from the BBMAP v37.75 package. Assembly of MiSeq and HiSeq sequences was then performed with SPAdes (version 3.8.1) with

the following parameters: *-k* 21,33,55,77,99,127 *--cov-cutoff auto --careful*. This resulted in 57 contigs for JV3, 75 contigs for R551-3, 56 contigs for K279a, and 61 contigs for JCMS. Contigs were then analyzed using VecScreen in NCBI to identify vector contamination (www.ncbi.nlm.nih.gov/tools/vecscreen). Contigs with >20% overall coverage aligning to vectors were removed. In addition, BLASTn was used with default parameters to align each contig to other similar sequences. Contigs with no significant sequence similarity were removed along with remaining contigs less than 500 bp.

Genome annotation and comparison

Rapid Annotation Subsystem Technology (RAST) was used to annotate draft assemblies of each strain (www.rast.nmpdr.org) (Aziz et al., 2008). This database uses FIGfam protein families based on known functional roles and similarities between proteins in closely related species to assign protein-coding genes to functional groups (Aziz et al., 2008). Subsystems consist of groups of functional roles that make up biological processes, complexes, and protein families, and can be grouped into broader "Categories" and "Subcategories" (Aziz et al., 2008). FASTA files of the reference genomes were also submitted by us for RAST annotation instead of using available annotations in order to maintain consistency, as the RAST database updates its annotations frequently. Lists of subsystem features identified for each strain were then compared using a presence/absence matrix to determine strain-specific functions. BLASTp was also used to determine presence/absence of proteins in the draft genomes based on significant alignment of the CDS of the draft genome to the protein sequence of interest (www.blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins).

Draft and reference sequences were aligned in Mauve using default parameters to look for synteny between sequences (Darling et al., 2004). In addition, progressiveMauve alignment

was used to align all draft and reference genomes, using the complete K279a genome as a reference. Mauve uses local colinear blocks (LCB), or small, highly similar regions between sequences, as anchors to perform alignments of long sequences. Alignments between the LCBs is performed using CLUSTAL W. Mauve is advantageous for genome alignments that may contain a variety of rearrangements, as it allows for sequence fragments to be moved within the alignment. The output tree file from Mauve, based on the Neighbor Joining method, was input in FigTree v1.4.3 for visualization. BRIG, which utilizes BLAST+ alignments to visualize whole genome similarity, was used to align and compare all draft and reference genomes (Alikhan et al., 2011).

For alignment of the quorum sensing protein, RfpF, sequences, protein sequences were found for each draft genome within genbank files obtained from RAST annotations. These protein sequences were aligned using Clustal Omega to the 106 N-terminal residues of each RfpF residue: RfpF-1 (GenBank: KJ149475) and RfpF-2 (GenBank: KJ149552) (Sievers et al, 2011). This region is known to contain variability used for isoform assignment.

Determining the pangenome

The core genome contains genes present in all genomes, the shell genome represents moderately conserved, often dispensable genes, and the cloud genome contains strain specific genes, specific to 1-2 strains (Contreras-Moreira & Vinuesa, 2013). The pangenome consists of the sum of the core, shell, and cloud genomes.

To determine the pangenome of the *S. maltophilia* strains in our laboratory, K279a, R551-3, JV3 and JCMS draft genome sequences were combined with the CSM2 genome sequence (NZ_CP025298) and genome sequences of 24 additional strains previously used for pangenome analysis (Lira, Berg, & Martínez, 2016). These strains include: clinical strains K279a

(NC_010943.1), D457 (NC_017671.1), E729 (NERH00000000), E759 (NERG00000000), E999(NERF00000000), G51 (NERE00000000), E301 (NERD00000000), D388 (NERC00000000), E861 (NERB00000000), C357 (NERA00000000), E539 (NEQZ00000000), E824 (NEQY00000000); and environmental strains: R551-3 (NC_011071.1), JV3 (NC_015947.1), NS26 (NEQO00000000), EP13 (NEQX00000000), EA22 (NEQW00000000), EA1 (NEQV00000000), PS5 (NEQU00000000), EA23 (NEQT00000000), EP20 (NEQS00000000), EP5 (NEQR00000000), EA21 (NEQQ00000000), EA63 (NEQP00000000) (Alonso & Martínez, 1997; De Boer et al., 2001; Lira et al., 2012; Minkwitz & Berg, 2001; Ribbeck-Busch et al., 2005; Suckstorff & Berg, 2003). Pan- and core- genomes were determined by first identifying clusters of homologous genes using the GET_HOMOLOGUES package. These clusters were identified using two popular clustering methods: COGS and OMCL (Kristensen et al., 2010; Li, Stoeckert, & Roos, 2003), with clusters only used for further analysis if they were identified with both algorithms (the intersection). The parse_pangenome_matrix.pl script within GET_HOMOLOGUES was used with the -x option to use the intersection of the two above algorithms to create the pangenome matrix. The pangenome matrix includes the presence/absence of each gene cluster within each strain (Contreras-Moreira & Vinuesa, 2013), and is used to identify the core, shell, and cloud genomes.

C. elegans survival assays

Treatment or control, *E. coli* OP50, bacteria was cultured in liquid LB (with Ampicillin for *S. maltophilia* strains) overnight and 100 µl of bacteria was plated onto NGM agar plates the day prior to use. Worms were bleached to synchronize and reared at 20° C on lawns of *E. coli* OP50. For survival assays, ten fourth larval stage (L4) worms were transferred to each treatment plate, with three replicates of each treatment, and maintained at 25° C. Worms were transferred

to new plates every day until they stopped laying eggs to separate them from their progeny. Surviving worms were recorded each day and dead worms were removed from plates, with live worms determined by lack of movement following prodding their heads and tails with a platinum wire pick. Plates that became contaminated or worms that crawled off the agar and died were removed from data analysis.

Survival probability estimates over time were determined using the Kaplan-meier formula and used to produce survival curves. The Cox-proportional hazards model was then used to compare the effects of independent variables using hazard ratios. Independent variables included bacterial treatments, with the dependent variable being the probability of nematode death on a given day. Differences between times of experimentation were accounted for in the model. These analyses were performed in R (Vienne, Austria: R Foundation for Statistical Computing).

Results and Discussion

S. maltophilia strains display varying pathogenicity

We have previously developed *Caenorhabditis elegans* as a model organism for studying *S. maltophilia* pathogenesis. Because *C. elegans* are bacterivores that encounter *S. maltophilia* in their natural environment (Dirksen et al., 2016; Samuel et al., 2016), this is a suitable platform for studying host-bacterial interactions. The JCMS strain was isolated in association with nematodes at Konza Prairie near Manhattan, KS, USA (White et al., 2016). Other *S. maltophilia* strains utilized in this study include environmental isolates R551-3 and JV3, and clinical isolate K279a. The degree of pathogenicity to *C. elegans* is determined using Cox proportional hazards based on survival over time. Hazards can be defined as the probability of death at a given time and can be compared between conditions across the entire lifespan, resulting in hazard ratios.

Hazard ratios greater than one indicate detrimental, or hazardous, conditions, whereas hazard ratios less than one indicate beneficial conditions. Hazard ratios were determined for each of the above treatment *S. maltophilia* strains by comparing hazards of nematodes exposed to *S. maltophilia* to those of nematodes exposed to standard lab food *E. coli* OP50. This resulted in a hazard ratio of 0.86 for K279a, and hazard ratios of 1.7, 5.7, and 59.3 for R551-3, JV3, and JCMS, respectively. Thus, K279a appears to be nonpathogenic, perhaps even slightly beneficial, whereas R551-3, JCMS, and JV3 display varying degrees of pathogenicity in *C. elegans* (Figure 3.1). The virulence observed in these environmental strains supports previous studies that have identified virulence factors within both clinical and environmental isolates (Adamek et al., 2011; Alavi et al., 2014; Berg & Martinez, 2015; Berg et al., 1999; Lira, Berg, & Martínez, 2017; Ryan et al., 2009), and further suggests that not only are virulence mechanisms present in environmental strains, they actually lead to pathogenicity within the host. In addition, because of the variation in virulence of *S. maltophilia* strains, we were interested in elucidating genomic features within these strains that might be the source of these differences.

Reference and draft genome sequences are highly similar

Genomic DNA from *S. maltophilia* K279a, R551-3, JCMS, and JV3 was sequenced using both MiSeq and HiSeq platforms and assembled used SPAdes, resulting in an average of 62 contigs per strain. Contigs were then filtered using VecScreen and BLASTn (NCBI) to remove contamination, resulting in an average of 46 contigs per strain (Table 3.1).

In addition to the assembled drafts we generated, completely assembled reference sequences are available for K279a (Crossman et al., 2008), R551-3 (Lucas et al., 2008), and JV3 (Lucas et al., 2011). Due to contradicting information between observed virulence of these strains in our hands and previous data from the same strains, we wanted to compare genome

sequences of our lab strains to those of the published reference strains. Specifically, a previous study found that neither K279a or R551-3 were pathogenic to the amoeba *Dictyostelium discoideum* or *Acanthamoeba castellanii* (Adamek et al., 2011), another study revealed that K279a killed *C. elegans* within 24 hours (Fouhy et al., 2007), and our observations reveal that K279a was not pathogenic and R551-3 was mildly pathogenic to *C. elegans* (Figure 3.1).

In order to differentiate the reference genome sequences from our draft assemblies, the strains used to generate our draft assemblies will subsequently be referred to as K279a-KS, R551-3-KS, and JV3-KS. Overall, the -KS assemblies were similar to the reference genome sequences. K279a, the largest reference genome, is 4,851,126 bp with 4,386 predicted protein-coding sequences, whereas K279a-KS resulted in a predicted genome size of 4,803,885 bp and 4,392 protein-coding sequences. R551-3-KS, which resulted in a predicted genome size of 4,547,979 (4,109 protein-coding genes), and JV3-KS, which resulted in a predicted genome size of 4,526,743 (4,126 protein-coding genes), were also similar to their respective references (Table 3.1). GC content of all reference and drafts were also very similar at approximately 66.5% (Table 3.1).

Genome sequences of -KS strains were aligned to their respective references to determine the degree of synteny, resulting in ordered contigs (Figure 3.2). Overall, reference and draft genomes display a high degree of synteny, indicating no large rearrangements have occurred (Figure 3.2). Overall, this degree of similarity suggests published data about these reference genomes should also be relevant to the -KS strains. Genomes were annotated using Rapid Annotation Subsystem Technology (RAST), which classifies protein coding sequences based on subsystems (Aziz et al., 2008). Subsystems are groups of proteins based on their known functional roles. Protein-coding sequences that were classified into subsystems based on

functional roles were compared between reference and -KS genome sequences. Surprisingly, very few functional roles differed between reference and -KS genome sequences, again suggesting similarity between reference and -KS strains. In fact, there were not any differences between the R551-3 reference and R551-3-KS, and only one, a short ribosomal protein, present in the JV3-KS sequence but absent in the reference.

Two functional roles differed between the K279a reference and K279a-KS sequences. Interestingly, one of these differences, present in K279a-KS but absent in the reference, encodes the flagellar hook-associated protein 3, FlgL. Previous research revealed that flagella play a role in biofilm formation in *S. maltophilia*, as mutations in a flagella-specific ATPase, FliI, reduce biofilm formation efficiency and abolish motility (Pompilio et al., 2010). However, because mobility and biofilm formation have been observed in K279a (Fouhy et al., 2007; García, Alcaraz, Franco, & De Rossi, 2015), this likely does not explain phenotypic virulence differences between K279a-KS and K279a used by Fouhy and colleagues (Fouhy et al., 2007).

The other difference, a gene found in K279a reference but not K279a-KS, encodes HipA, a toxin that is part of a Type II toxin-antitoxin (TA) system, in which HipB is the antitoxin. HipA is also found in R551-3, but HipB is not found in any of the reference or -KS genome sequences. Expression of the HipA toxin at appropriate levels leads to an increase in 'persister' cells that are tolerant to antibiotics and other stressors (Lewis, 2007; Rotem et al., 2010). Interestingly, persister cells have been hypothesized to play a role in antibiotic resistance and biofilm formation in *S. maltophilia* (Abda et al., 2015; Di Bonaventura et al., 2010), but functional analysis of this system in *S. maltophilia* has not yet been performed. In addition, toxins do not seem to play a major role in *S. maltophilia* pathogenesis (White et al., 2016), though it is

possible that this toxin indirectly impacts virulence through antibiotic resistance or biofilm formation.

One explanation for differences in observed virulence between K279a and K279a-KS that is not dependent on genomic features could be linked to environmental controls on gene regulation. In fact, one study looked at gene expression of an *S. maltophilia* strain grown at different temperatures and found differential expression of genes involved in stress responses and motility (Alavi et al., 2014). However, a different study determining pathogenic potential of *S. maltophilia* strains based on genomic features found that temperature did not account for changes in virulence in several strains (Adamek et al., 2011). Thus, functional analyses of genetic differences as well as transcriptomic approaches could be utilized to unravel mechanisms behind these phenotypic differences.

S. maltophilia JCMS is closely related to JV3

Assembly of JCMS, a novel, pathogenic, environmental isolate, resulted in a total predicted genome size of 4,919,076 with 4,371 predicted protein-coding sequences and a GC content of 66.4% (Table 3.1). RAST was used to annotate the resulting incomplete assembly. Figure 3.3 summarizes subsystem annotations by more broadly grouping subsystems into categories. Interestingly, JCMS contains more genes classified in the "Virulence, Disease, and Defense" category than the other strains analyzed, with 194 genes as compared to 149, 125, and 131 genes in K279a-KS, R551-3-KS, and JV3-KS, respectively (Figure 3.3). This could indicate novel virulence mechanisms in JCMS, or a combination of virulence mechanisms found in other strains.

From 16S rRNA gene sequencing, we determined that *S. maltophilia* JCMS was most similar to *S. maltophilia* JV3 (White et al., 2016). Therefore, we aligned the JCMS contigs using

JV3 as a reference in Mauve, resulting in a high degree of synteny between JCMS and JV3 (Figure 3.4). Interestingly, when contigs were filtered using BLASTn, we found that JCMS has more contigs that align to a different completely sequenced strain of *S. maltophilia*, CSM2, than to any other *S. maltophilia* strain. CSM2 (NZ_CP025298.1) is an environmental isolate that was found in Morelos, Cuernavaca, Mexico (Castro-Jaimes et al., 2017), but further information on this strain has yet to be reported. To further explore the similarity between these two strains, CSM2 was included in subsequent analyses.

BLAST+ was used to align all sequences to the K279a reference, as it is the largest reference genome. This resulted in an overall alignment between all reference and draft assemblies that shows similarity between strains, with the highest degree of similarity between reference and draft sequences of the same strain (Figure 3.5). This is further supported by a multiple sequence alignment of all references and ordered contigs from Figure 3.2 and Figure 3.4 (Figure 3.6A). Phylogenetic analysis using the Neighbor Joining distance matrix from these whole-genome Mauve alignments revealed that CSM2, JCMS, and JV3 are closely related, while R551-3 is more closely related to K279a than the other pathogenic strains (Figure 3.6B). This supports previous findings that clinical and environmental isolates are heterogenic in evolutionary origin, as clinical and environmental isolates cluster together (Berg et al., 1999; Lira, Berg, & Martínez, 2017).

Many virulence factors are common between strains

We again utilized RAST subsystem annotations to determine whether genomic features differ between strains. Protein-coding sequences for each genome that were classified into functional roles were organized into a presence/absence matrix. Only 11% (179) of the 1,577 functional roles differed between at least two draft genomes (Table 3.2). Of the 179 differences

in functional roles between strains, 65 were different when K279a and R551-3 were compared to JV3 and JCMS (Table 3.2). This is in support of whole genome-sequence phylogenetic data suggesting highest similarity between JV3 and JCMS (Figure 3.6B). Genes assigned to subsystems are more broadly grouped into categories, one of which is "Virulence, Disease, and Defense" (Table 3.2). Interestingly, very few functional roles assigned to the "Virulence, Disease, and Defense" category differed between genomes. Specifically, of 179 functional roles that differed in at least one genome, only 12 were assigned to the category "Virulence, Disease, and Defense" (Table 3.2). Interestingly, 11 of these were involved in metal resistance or homeostasis, with seven present in K279a and JCMS and absent in R551-3 and JV3 (Table 3.2). These genes could be beneficial for environmental strains as heavy metal pollutants are found in many environments common to *S. maltophilia*, such as soils and rhizospheres. These genes also have the potential to play a role in host pathogenesis as metal ions are important for host immune response and metal ion concentration changes in host tissue upon infection with *S. maltophilia* (Pompilio et al., 2014). Therefore, ability to transport metal ions may be a mechanism for increasing virulence as well as increasing the ability to compete for these elements. However, whether these genes affect virulence remains unclear, and their presence in one clinical isolate (K279a) and one environmental isolate (JCMS) but absence in others requires further analysis.

The other "Virulence, Disease, and Defense" categorized gene that differed between genomes was the resistant nodulation division (RND) efflux system membrane-fusion protein (MFP) CmeA, which was absent in K279a but present in R551-3, JV3, and JCMS. This gene, which is homologous to AcrA in *E. coli*, encodes one of three proteins that forms an RND efflux pump. Interestingly, all strains contain genes that encode CmeB and CmeC, which together make the tripartite efflux pump CmeABC. Whereas the outer membrane protein (CmeC) of a tripartite

efflux pump is more indiscriminate in its interactions, the MFP (CmeA) is specific to each RND efflux protein (Crossman et al., 2008). Therefore, without CmeA, it is likely that this efflux pump is inactive in K279a. The CmeABC efflux pump is known to be involved in antibiotic resistance in *Campylobacter* species (Vieira et al., 2017; Yao et al., 2016), suggesting it could also play a role in pathogenesis in *S. maltophilia*.

Another striking difference found between strains was the presence of 11 functional roles involved in the Type IV secretion system in the genome sequences of K279a and R551-3 that were absent in JCMS and JV3. Secretion systems are common virulence mechanisms that transport enzymes and other proteins into their surrounding environment. The functional roles absent in JCMS and JV3, including VirB11, VirB4, VirB8, and VirB10, consist of major components and enzymes required for Type IV secretion system function (Fronzes, Christie, & Waksmas, 2009). Therefore, it is unlikely that this system is functional in JCMS and JV3. Interestingly, these genes were included in the "Membrane Transport" category and thus may not be involved in virulence. Type IV secretion systems are known for transporting a variety of substances (Fronzes, Christie, & Waksmas, 2009), and the roles of many of the Type IV secretion system genes found in K279a and R551-3 are involved in transfer DNA (T-DNA) secretion. However, these systems are also known for transferring toxins (Fronzes, Christie, & Waksmas, 2009). In fact, effectors of the Type IV secretion system in a close relative of *S. maltophilia*, *Xanthomonas citri*, are capable of killing other bacterial cells (Souza et al., 2015). Therefore, it is possible that this secretion system plays a role in transport of virulence factors, but functional analysis of these genes is necessary to understand their possible role in virulence and to provide an explanation for their presence in the less virulent strains, K279a and R551-3.

Although few differences exist between strains that could explain differences in virulence, a variety of important virulence mechanisms are conserved between all strains. In the initial assembly of K279a, nine RND-efflux pumps of the drug-resistance type were identified (Crossman et al., 2008), all of which are also present in all strains. In addition, a quorum sensing system has been identified in *S. maltophilia* that likely facilitates cell-to-cell communication and synchronization of cell behaviors. Quorum-sensing systems are dependent on diffusible signaling factors (DSF), which in *S. maltophilia* is the fatty acid *cis*-11-methyl-2-dodecenoic acid (Fouhy et al., 2007). RfpF, a protein encoded by the regulation of pathogenicity factors (*rpf*) cluster, is necessary for synthesis of the DSF in *S. maltophilia*, which increases motility and virulence (Fouhy et al., 2007; Huedo et al., 2014). While *rfpF* is common in *S. maltophilia* genomes, two isoforms of the *rfpF* gene exist, and only strains encoding RfpF-1 produce DSF under standard conditions (Huedo et al., 2014). Multiple sequence alignment of the RfpF sequences of these strains revealed that K279a-KS and R551-3-KS contain the RfpF-1 variant, while JCMS, JV3-KS, and CSM2-KS contain identical RfpF-2 variants (Figure 3.7). This suggests that the more virulent strains, JCMS and JV3, lack functional quorum-sensing systems.

One limitation of the RAST annotation database is that many protein-coding genes lack functional annotation. For example, of the *S. maltophilia* strains annotated, approximately 54% of genes are not categorized into subsystems. This could provide an explanation for the lack of differences identified using the subsystem approach, as genes without subsystem annotation may play an important role in virulence.

Core and cloud genome analysis reveals strain specificity

For a genome-wide analysis of unique and shared genes within each genome, we utilized the GET_HOMOLOGUS software to identify the core, shell, and cloud genomes of each *S.*

maltophilia strain (Contreras-Moreira & Vinuesa, 2013). Briefly, the core genome is shared between all strains, the cloud genome is unique to each strain, and the shell is moderately conserved and contains dispensable genes. Because a majority of genomes used for this analysis were not completely sequenced and could contain genetic features that fall between contigs, we also characterized the soft core genome, which is shared between almost all (95%) of the genomes. The pangenome contains all genes that make up the core, soft core, cloud, and shell genome. To obtain a more accurate representation of the core genome, we also utilized *S. maltophilia* strains from a previous study (Lira, Berg, & Martínez, 2017) that analyzed the pangenome of 24 environmental and clinical strains. We used these 24 genome sequences as well as those of our -KS strains and CSM2 for this analysis. These sequences were aligned in Mauve to determine relatedness of strains and the resulting phylogenetic tree recapitulates phylogenetic clustering determined by MLST and whole CDS alignment performed by Lira and colleagues (Lira, Berg, & Martínez, 2017) (Figure 3.8). JV3, JCMS, and CSM2 again form a closely related cluster, while K279a and R551-3 fall into different clusters with the addition of these strains (Figure 3.8).

Pangenome analysis resulted in 9,161 total gene clusters determined by homology of CDSs. The core genome contains 2,803 orthologous clusters that are shared between all species (3,108 with the addition of the soft-core gene clusters), and 3,756 genes comprising the cloud genome, or genes specific to one to two strains (Figure 3.9A). The addition of the four -KS genomes and CSM2 resulted in approximately 2,000 genes added to the pangenome identified by Lira and colleagues (Lira, Berg, & Martínez, 2017). A majority of these genes were added to the shell and cloud genomes, with the core genome remaining similar in size. Since the addition of these five genomes did not increase the size of the core genome, it appears that nearly all the

core genes in *S. maltophilia* have been identified and the core genome should contain all genes essential for *S. maltophilia* growth and survival. However, the substantial increase in shell and cloud genomes emphasizes the diversity of the *S. maltophilia* genome, as many genes are not conserved across strains.

Core, shell, and cloud genomes were further analyzed in the -KS strains and CSM2 (Figure 3.9B). Interestingly, JCMS and K279a have larger cloud genomes than other strains, suggesting they have more strain-specific functions (Figure 3.9B). However, over 75% of all proteins found in the cloud genome encode hypothetical proteins, suggesting that many of these important strain-specific functions have yet to be determined.

Conclusions

This study analyzed genomic content of *S. maltophilia* strains of varying pathogenicity to elucidate virulence mechanisms that could account for these phenotypic differences. Although genomic differences between strains do not clearly explain differences in observed virulence, analysis of the pangenome revealed that a majority of the genome is not shared between strains. In fact, 41% of the pangenome is characterized as cloud genome, being specific to 1-2 strains (Figure 3.9A). This suggests that many non-essential genes, including those involved in virulence, are found in the cloud and shell genomes. In fact, previous analysis has identified a variety of virulence factors on genomic islands, which contain genomic regions of past or present mobility (Adamek et al., 2014; Juhas et al., 2008). In addition, we determined that many of these strain-specific genes have yet to be functionally characterized, emphasizing how much there is to learn about *S. maltophilia*. By determining phenotypic virulence of the additional strains used for the pangenome analysis, a more comprehensive analysis could help explain the relationship between genomic variability and observed virulence. In doing so, these studies could provide

more concrete conclusions about whether virulence factor differences are phenotypically important.

In conclusion, many virulence mechanisms identified in genome sequences were common to all *S. maltophilia* strains, regardless of phenotypic behavior. Because evolutionary origin and sampling origin are poor predictors of phenotypic pathogenicity (Adamek et al., 2011; Pompilio et al., 2011), understanding key virulence mechanisms in *S. maltophilia* is important for determining pathogenic potential. However, the large number of strain specific genes makes identification of virulence mechanisms across *S. maltophilia* strains difficult. Therefore, although sequencing of more strains will provide information on important genetic features, these features need to be identified within each *S. maltophilia* strain independently. In addition, future research should focus efforts on uncharacterized genes, as they potentially encode important virulence determinants.

References

Abda, E. M., Krysciak, D., Krohn-Molt, I., Mamat, U., Schmeisser, C., Förstner, K. U., Schaible, U. E., Kohl., T. A., Nieman, S., & Streit, W. R. (2015). Phenotypic heterogeneity affects *Stenotrophomonas maltophilia* K279a colony morphotypes and β-lactamase expression. *Frontiers in Microbiology, 6*(1373). https://doi.org/10.3389/fmicb.2015.01373

Adamek, M., Linke, B., & Schwartz, T. (2014). Virulence genes in clinical and environmental *Stenotrophomas maltophilia* isolates: A genome sequencing and gene expression approach. *Microbial Pathogenesis, 67–68*, 20–30. https://doi.org/10.1016/j.micpath.2014.02.001

Adamek, M., Overhage, J., Bathe, S., Winter, J., Fischer, R., & Schwartz, T. (2011). Genotyping of environmental and clinical *Stenotrophomonas maltophilia* isolates and their pathogenic potential. *PLoS ONE, 6*(11), e27615. https://doi.org/10.1371/journal.pone.0027615

Alavi, P., Starcher, M. R., Thallinger, G. G., Zachow, C., Müller, H., & Berg, G. (2014). *Stenotrophomonas* comparative genomics reveals genes and functions that differentiate beneficial and pathogenic bacteria. *BMC Genomics, 15*(482). https://doi.org/10.1186/1471-2164-15-482

Alikhan, N. F., Petty, N. K., Ben Zakour, N. L., & Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): Simple prokaryote genome comparisons. *BMC Genomics, 12*(402). https://doi.org/10.1186/1471-2164-12-402

Alonso, A., & Martínez, J. L. (1997). Multiple antibiotic resistance in *Stenotrophomonas maltophilia*. *Antimicrobial Agents and Chemotherapy, 41*(5), 1140–1142.

Aziz, R. K., Bartels, D., Best, A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G. D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A., & Zagnitko, O. (2008). The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics, 9*(75). https://doi.org/10.1186/1471-2164-9-75

Berg, G., Marten, P., Ballin, G. (1996). *Stenotrophomonas maltophilia* in the rhizosphere of oilseed rape – occurrence, characterization and interaction with phytopathogenic fungi. *Microbiological Research, 151*(1): 19-27.

Berg, G., & Martinez, J. L. (2015). Friends or foes: Can we make a distinction between beneficial and harmful strains of the *Stenotrophomonas maltophilia* complex? *Frontiers in Microbiology, 6*(241). https://doi.org/10.3389/fmicb.2015.00241

Berg, G., Roskot, N., & Smalla, K. (1999). Genotypic and phenotypic relationships between clinical and environmental isolates of *Stenotrophomonas maltophilia. Journal of Clinical Microbiology, 37*(11), 3594–600.

Berg, M., Stenuit, B., Ho, J., Wang, A., Parke, C., Knight, M., Alverez-Cohen, L., & Shapira, M. (2016). Assembly of the *Caenorhabditis elegans* gut microbiota from diverse soil microbial environments. *The ISME Journal, 10*, 1998–2009. https://doi.org/10.1038/ismej.2015.253

Bittar, F., Richet, H., Dubus, J. C., Reynaud-Gaubert, M., Stremler, N., Sarles, J., Raoult, D., & Rolain, J. M. (2008). Molecular detection of multiple emerging pathogens in sputa from cystic fibrosis patients. *PLoS ONE, 3*(8). https://doi.org/10.1371/journal.pone.0002908

Brooke, J. S. (2012). *Stenotrophomonas maltophilia*: An emerging global opportunistic pathogen. *Clinical Microbiology Reviews, 25*(1), 2–41. https://doi.org/10.1128/CMR.00019-11

Castro-Jaimes, S., Lopez-Leal, G., Barberena Jonas, C., Bustos, P., Perez-Oseguera, A. and Cevallos, M.A. (2017). Complete Genome Sequence of *Stenotrophomonas maltophilia* CSM2, accession: NZ_CP025298. RefSeq.

Chang, Y.-T., Lin, C.-Y., Chen, Y.-H., & Hsueh, P.-R. (2015). Update on infections caused by *Stenotrophomonas maltophilia* with particular attention to resistance mechanisms and therapeutic options. *Frontiers in Microbiology, 6*(893), 1–20. https://doi.org/10.3389/fmicb.2015.00893

Contreras-Moreira, B., & Vinuesa, P. (2013). GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Applied and Environmental Microbiology, 79*(24), 7696–7701. https://doi.org/10.1128/AEM.02411-13

Crossman, L. C., Gould, V. C., Dow, J. M., Vernikos, G. S., Okazaki, A., Sebaihia, M., Saunders, D., Arrowsmith, C., Carver, T., Peters, N., Adlem, E., Kerhornou, A., Lord, A., Murphy, L., Seeger, K., Squares, R., Rutter, S., Quail, M., Rajandream, M., Harris, D., Churcher, C., Bentley, S. D., Parkhill, J., Thomson, N. R., & Avison, M. B. (2008). The complete genome, comparative and functional analysis of *Stenotrophomonas maltophilia* reveals an organism heavily shielded. *Genome Biology*, *9*(4), R74.1-R74.13. https://doi.org/10.1186/gb-2008-9-4-r74

Darling, A. C. E., Mau, B., Blattner, F. R., & Perna, N. T. (2004). Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Research, 14*(7), 1394–1403. https://doi.org/10.1101/gr.2289704

De Boer, W., Klein Gunnewiek, P. J. A., Kowalchuk, G. A., & Van Veen, J. A. (2001). Growth of chitinolytic dune soil β-subclass proteobacteria in response to invading fungal hyphae. *Applied and Environmental Microbiology, 67(*8), 3358–3362. https://doi.org/10.1128/AEM.67.8.3358-3362.2001

Denton, M., & Kerr, K. G. (1998). Microbiological and clinical aspects of infection associated with *Stenotrophomonas maltophilia. Clinical Microbiology Reviews, 11*(1), 57–80. https://doi.org/10.1074/jbc.M007003200

Di Bonaventura, G., Pompilio, A., Zappacosta, R., Petrucci, F., Fiscarelli, E., Rossi, C., & Piccolomini, R. (2010). Role of excessive inflammatory response to *Stenotrophomonas maltophilia* lung infection in DBA/2 mice and implications for cystic fibrosis. *Infection and Immunity, 78*(6), 2466–2476. https://doi.org/10.1128/IAI.01391-09

Dirksen, P., Marsh, S. A., Braker, I., Heitland, N., Wagner, S., Nakad, R., Mader, S., Peterson, C., Kowallik, V., Rosenstiel, P., Felix, M-A., & Schulenburg, H. (2016). The native microbiome of the nematode *Caenorhabditis elegans*: gateway to a new host-microbiome model. *BMC Biology, 14*(38). https://doi.org/10.1186/s12915-016-0258-1

Elting, L. S., & Bodey, G. P. (1990). Septicemia due to *Xanthomonas* species and non-aeruginosa *Pseudomonas* species: Increasing incidence of catheter-related infections. *Medicine (Baltimore), 69*(5), 296–306.

Ferrer-Navarro, M., Planell, R., Yero, D., Mongiardini, E., Torrent, G., Huedo, P., Martinez, P., Roher, N., Mackenzie, S., Gibert, I., & Daura, X. (2013). Abundance of the quorum-sensing factor Ax21 in four strains of *Stenotrophomonas maltophilia* correlates with mortality rate in a new zebrafish model of infection. *PLoS ONE, 8*(6), e67207. https://doi.org/10.1371/journal.pone.0067207

Fouhy, Y., Scanlon, K., Schouest, K., Spillane, C., Crossman, L., Avison, M. B., Ryan, R. P., Dow, J. M. (2007). Diffusible signal factor-dependent cell-cell signaling and virulence in the nosocomial pathogen *Stenotrophomonas maltophilia. Journal of Bacteriology, 189*(13), 4964–4968. https://doi.org/10.1128/JB.00310-07

Fronzes, R., Christie, P. J., & Waksmas, G. (2009). The structural biology of type IV secretion systems. *Nature Review Microbiology, 7*(10). https://doi.org/10.1038/nrmicro2218.The

García, C. A., Alcaraz, E. S., Franco, M. A., & De Rossi, B. N. P. (2015). Iron is a signal for *Stenotrophomonas maltophilia* biofilm formation, oxidative stress response, OMPs expression, and virulence. *Frontiers in Microbiology, 6*(926). https://doi.org/10.3389/fmicb.2015.00926

Huedo, P., Yero, D., Martínez-Servat, S., Estibariz, I., Planell, R., Martínez, P., Ruyra, A., Roher, N., Roca, I., Vila, J., Caura, X., & Gibert, I. (2014). Two different *rpf* clusters distributed among a population of *Stenotrophomonas maltophilia* clinical strains display differential diffusible signal factor production and virulence regulation. *Journal of Bacteriology, 196*(13), 2431–2442. https://doi.org/10.1128/JB.01540-14

Jägevall, S., Rabe, L., & Pedersen, K. (2011). Abundance and diversity of biofilms in natural and artificial aquifers of the Äspö Hard Rock Laboratory, Sweden. *Microbial Ecology, 61*(2), 410–422. https://doi.org/10.1007/s00248-010-9761-z

Juhas, M., van der Meer, J. R., Gaillard, M., Harding, R. M., Hood, D. W., & Crook, D. W. (2009). Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiology Review, 33,* 376–393. https://doi.org/10.1111/j.1574-6976.2008.00136.x

Kaiser, S., Biehler, K., & Jonas, D. (2009). A *Stenotrophomonas maltophilia* multilocus sequence typing scheme for inferring population structured. *Journal of Bacteriology, 191*(9), 2934–2943. https://doi.org/10.1128/JB.00892-08

Kristensen, D. M., Kannan, L., Coleman, M. K., Wolf, Y. I., Sorokin, A., Koonin, E. V., & Mushegian, A. (2010). A low-polynomial algorithm for assembling clusters of orthologous groups from intergenomic symmetric best matches. *Bioinformatics, 26*(12), 1481–1487. https://doi.org/10.1093/bioinformatics/btq229

Lewis, K. (2007). Persister cells, dormancy and infectious disease. *Nature Reviews Microbiology, 5*(1), 48–56. https://doi.org/10.1038/nrmicro1557

Li, L., Stoeckert, C. J. J., & Roos, D. S. (2003). OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Research, 13*(9), 2178–2189. https://doi.org/10.1101/gr.1224503.candidates

Lira, F., Berg, G., & Martínez, J. L. (2017). Double-face meets the bacterial world: The opportunistic pathogen *Stenotrophomonas maltophilia. Frontiers in Microbiology, 8*, 2190. https://doi.org/10.3389/fmicb.2017.02190

Lira, F., Hernández, A., Belda, E., Sánchez, M. B., Moya, A., Silva, F. J., & Martínez, J. L. (2012). Whole-genome sequence of *Stenotrophomonas maltophilia* D457, A clinical isolate and a model strain. *Journal of Bacteriology, 194*(13), 3563–3564. https://doi.org/10.1128/JB.00602-12

Lucas, S., Copeland,A., Lapidus,A., Glavina del Rio,T., Dalin,E., Tice,H., Pitluck,S., Chain,P., Malfatti,S., Shin,M., Vergez,L., Lang,D., Schmutz,J., Larimer,F., Land,M., Hauser,L., Kyrpides,N., Mikhailova,N., Taghavi,S., Monchy,S., Newman,L., Vangronsveld,J., van der Lelie,D. and Richardson,P. (2008). Complete sequence of *Stenotrophomonas maltophilia* R551-3, accession: NC_011071.1. RefSeq.

Lucas, S., Han, J., Lapidus,A., Cheng,J.-F., Goodwin,L., Pitluck,S. Peters,L., Ovchinnikova,G., Teshima,H., Detter,J.C., Han,C., Tapia,R., Land,M., Hauser,L., Kyrpides,N., Ivanova,N., Pagani,I., Medau,R., Furlan,B., Mui Tsai,S., Rodriques,J., Tiedje,J. and Woyke,T. (2011) Complete sequence of *Stenotrophomonas maltophilia* JV3, accession: NC_015947.1. RefSeq.

Minkwitz, A., & Berg, G. (2001). Comparison of antifungal activities and 16S ribosomal DNA sequences of clinical and environmental isolates of *Stenotrophomonas maltophilia. Journal of Clinical Microbiology, 39*(1), 139–145. https://doi.org/10.1128/JCM.39.1.139

Nicoletti, M., Iacobino, A., Prosseda, G., Fiscarelli, E., Zarrilli, R., De Carolis, E., Petrucca, A., Memcioni, L., Colonna, B., & Casalino, M. (2011). *Stenotrophomonas maltophilia* strains from cystic fibrosis patients: Genomic variability and molecular characterization of some virulence determinants. *International Journal of Medical Microbiology, 301*(1), 34–43. https://doi.org/10.1016/j.ijmm.2010.07.003

Park, M., Kim, C., Yang, J., Lee, H., Shin, W., Kim, S., & Sa, T. (2005). Isolation and characterization of diazotrophic growth promoting bacteria from rhizosphere of agricultural crops of Korea. *Microbiological Research, 160*(2), 127–133. https://doi.org/10.1016/j.micres.2004.10.003

Pompilio, A., Ciavardelli, D., Crocetta, V., Consalvo, A., Zappacosta, R., Di Ilio, C., & Di Bonaventura, G. (2014). *Stenotrophomonas maltophilia* virulence and specific variations in trace elements during acute lung infection: Implications in cystic fibrosis. *PLoS ONE, 9*(2), e88769. https://doi.org/10.1371/journal.pone.0088769

Pompilio, A., Crocetta, V., Confalone, P., Nicoletti, M., Petrucca, A., Guarnieri, S., Fiscarelli, E., Savini, V., Piccolomini, & Di Bonaventura, G. (2010). Adhesion to and biofilm formation on IB3-1 bronchial cells by *Stenotrophomonas maltophilia* isolates from cystic fibrosis patients. *BMC Microbiology, 10*(102), 1–15. https://doi.org/10.1186/1471-2180-10-102

Pompilio, A., Crocetta, V., Ghosh, D., Chakrabarti, M., Gherardi, G., Vitali, L. A., Fiscarelli, E., & Di Bonaventura, G. (2016). *Stenotrophomonas maltophilia* phenotypic and genotypic diversity during a 10-year colonization in the lungs of a cystic fibrosis patient. *Frontiers in Microbiology, 7*(1551). https://doi.org/10.3389/fmicb.2016.01551

Pompilio, A., Pomponio, S., Crocetta, V., Gherardi, G., Verginelli, F., Fiscarelli, E., Dicuonzo, G., Savini, V., D'Antonio, D., & Di Bonaventura, G. (2011). Phenotypic and genotypic characterization of *Stenotrophomonas maltophilia* isolates from patients with cystic fibrosis: Genome diversity, biofilm formation, and virulence. *BMC Biology, 11*(159), 1–16.

Ribbeck-Busch, K., Roder, A., Hasse, D., De Boer, W., Martínez, J. L., Hagemann, M., & Berg, G. (2005). A molecular biological protocol to distinguish potentially human pathogenic *Stenotrophomonas maltophilia* from plant-associated *Stenotrophomonas rhizophila*. *Environmental Microbiology, 7*(11), 1853–1858. https://doi.org/10.1111/j.1462-2920.2005.00928.x

Rocco, F., De Gregorio, E., Colonna, B., & Di Nocera, P. P. (2009). *Stenotrophomonas maltophilia* genomes: A start-up comparison. *International Journal of Medical Microbiology, 299*(8), 535–546. https://doi.org/10.1016/j.ijmm.2009.05.004

Rotem, E., Loinger, A., Ronin, I., Levin-Reisman, I., Gabay, C., Shoresh, N., Biham, O., & Balaban, N. Q. (2010). Regulation of phenotypic variability by a threshold-based mechanism underlies bacterial persistence. *Proceedings of the National Academy of Sciences, 107*(28), 12541–12546. https://doi.org/10.1073/pnas.1004333107

Ryan, R. P., Monchy, S., Cardinale, M., Taghavi, S., Crossman, L., Avison, M. B., Berg, G., van der Lelie, D., & Dow, J. M. (2009). The versatility and adaptation of bacteria from the genus *Stenotrophomonas. Nature Reviews Microbiology, 7*(7), 514–525. https://doi.org/10.1038/nrmicro2163

Samuel, B. S., Rowedder, H., Braendle, C., Félix, M., & Ruvkun, G. (2016). *Caenorhabditis elegans* responses to bacteria from its natural habitats. *Proceedings of the National Academy of Sciences, 113*(27), E3941–E3949. https://doi.org/10.1073/pnas.1607183113

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J. D., & Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology, 7(*539). https://doi.org/10.1038/msb.2011.75

Souza, D. P., Oka, G. U., Alvarez-Martinez, C. E., Bisson-Filho, A. W., Dunger, G., Hobeika, L., Cavalcante, N. S., Alegria, M. C., Barbosa, L. R.S., Salinas, R. K., Guzzo, C. R., & Farah, C. S. (2015). Bacterial killing via a type IV secretion system. *Nature Communications, 6*(6453). https://doi.org/10.1038/ncomms7453

Suckstorff, I., & Berg, G. (2003). Evidence for dose-dependent effects on plant growth by *Stenotrophomonas* strains from different origins. *Journal of Applied Microbiology, 95*(4), 656–663. https://doi.org/10.1046/j.1365-2672.2003.02021.x

Turrientes, M. C., Baquero, M. R., Sánchez, M. B., Valdezate, S., Escudero, E., Berg, G., Canton, R., Baquero, F., Galan, J. C., & Martínez, J. L. (2010). Polymorphic mutation frequencies of clinical and environmental *Stenotrophomonas maltophilia* populations. *Applied and Environmental Microbiology*, *76*(6), 1746–1758. https://doi.org/10.1128/AEM.02817-09

Valdezate, S., Vindel, A., Martín-Dávila, P., Del Saz, B. S., Baquero, F., & Cantón, R. (2004). High genetic diversity among *Stenotrophomonas maltophilia* strains despite their originating at a single hospital. *Journal of Clinical Microbiology*, *42*(2), 693–699. https://doi.org/10.1128/JCM.42.2.693-699.2003

Vieira, A., Ramesh, A., Seddon, A. M., & Karlyshev, A. V. (2017). CmeABC multidrug efflux pump contributes to antibiotic resistance and promotes *Campylobacter jejuni* survival and multiplication in *Acanthamoeba polyphaga*. *Applied and Environmental Microbiology, 83*(22), e01600-17. https://doi.org/10.1128/AEM.01600-17

White, C. V, Darby, B. J., Bredden, R. J., & Herman, M. A. (2016). A *Stenotrophomonas maltophilia* strain evades a major *Caenorhabditis elegans* defense pathway. *Infection and Immunity, 84*(2), 524–536.

Yao, H., Shen, Z., Wang, Y., Deng, F., Liu, D., Naren, G., Dai, L., Su. C.-C., Wang, B., Wang, S., Wu, C., Yu, E. W., & Shen, J. (2016). Emergence of a potent multidrug efflux pump variant that enhances *Campylobacte*r resistance to multiple antibiotics. *mBio, 7*(5), 1–11. https://doi.org/10.1128/mBio.01543-16

Zhang, F., Berg, M., Dierking, K., Félix, M., Shapira, M., Samuel, B. S., & Schulenburg, H. (2017). *Caenorhabditis elegans* as a model for microbiome research. *Frontiers in Microbiology, 8*(485), 1–10. https://doi.org/10.3389/fmicb.2017.00485

Figures

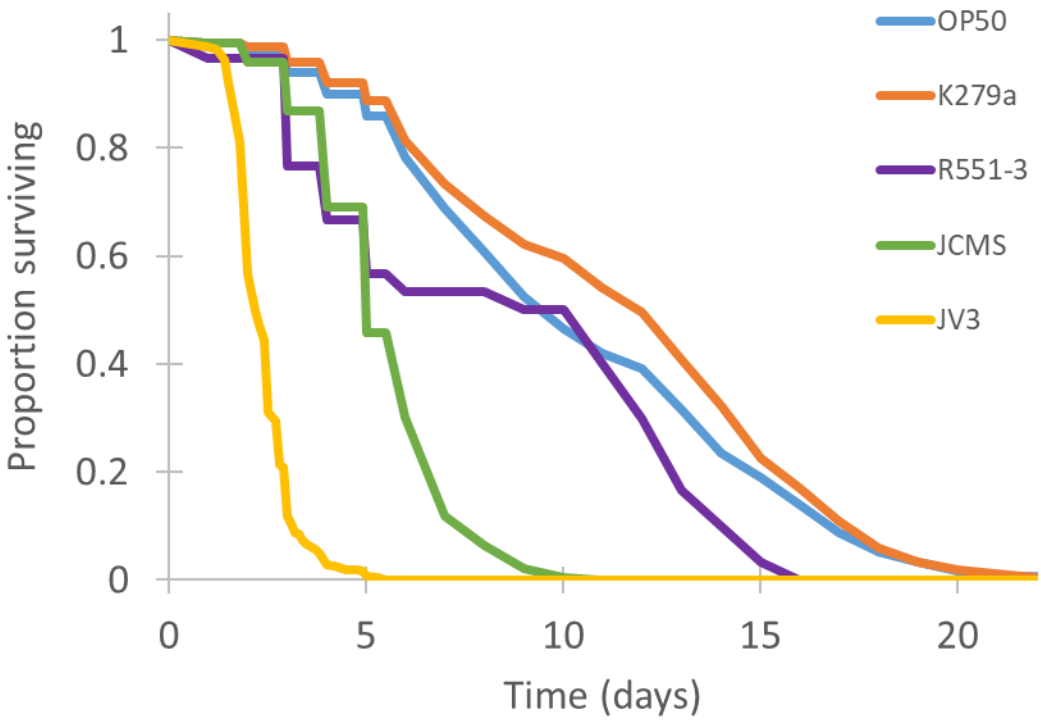

Figure 3.1 *S. maltophilia* strains display varying pathogenicity to *C. elegans*

Survivorship of *wild-type* nematodes on *S. maltophilia* K279a, R551-3, JCMS, JV3, and *E. coli* OP50. Survival estimates were determined using Kaplan-Meier estimates generated in R. This data represents 22 individual experiments, with n= 456-575 for each bacterial treatment (n=30 for R551-3). Restricted mean survival values (and standard error) generated by Kaplan-Meier estimates for each treatment are: OP50= 10.28 (0.20), K279a= 11.18 (0.18), R551-3= 8.53 (0.87), JCMS= 5.48 (0.08), and JV3= 2.41 (0.03). All hazard ratios were significant ($p<0.05$) when compared to *E. coli* OP50.

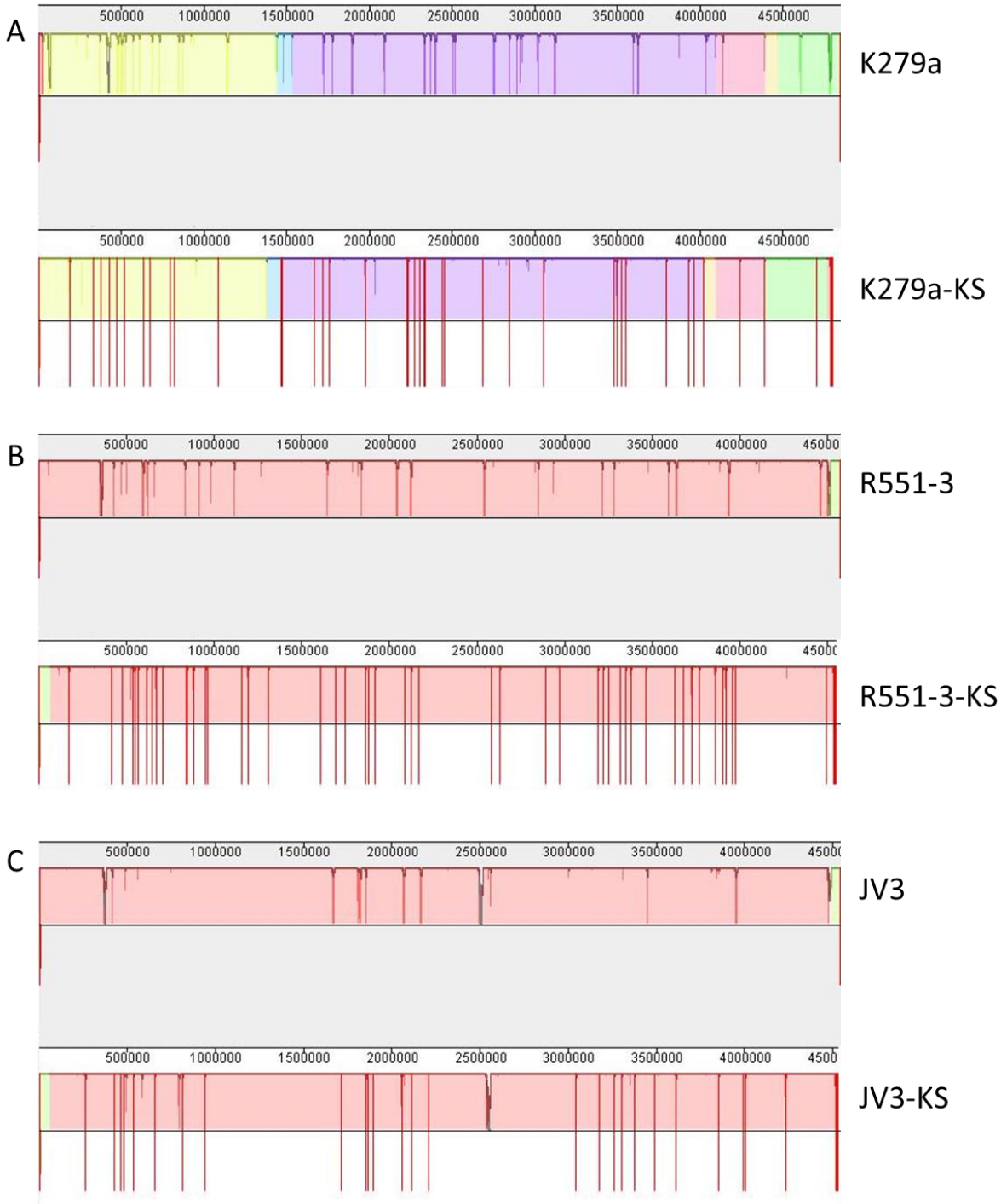

Figure 3.2 *S. maltophilia* reference and draft genomes display a high degree of synteny

References are shown on top with respective draft assemblies below, resulting in ordered contigs. Plots represent percent similarity between sequences, and similar colors between reference and draft represent regions of synteny. Contig boundaries are displayed as red lines. **A** K279a draft aligned to K279a reference. **B** R551-3 draft aligned to R551-3 reference. **C** JV3 draft aligned to JV3 reference.

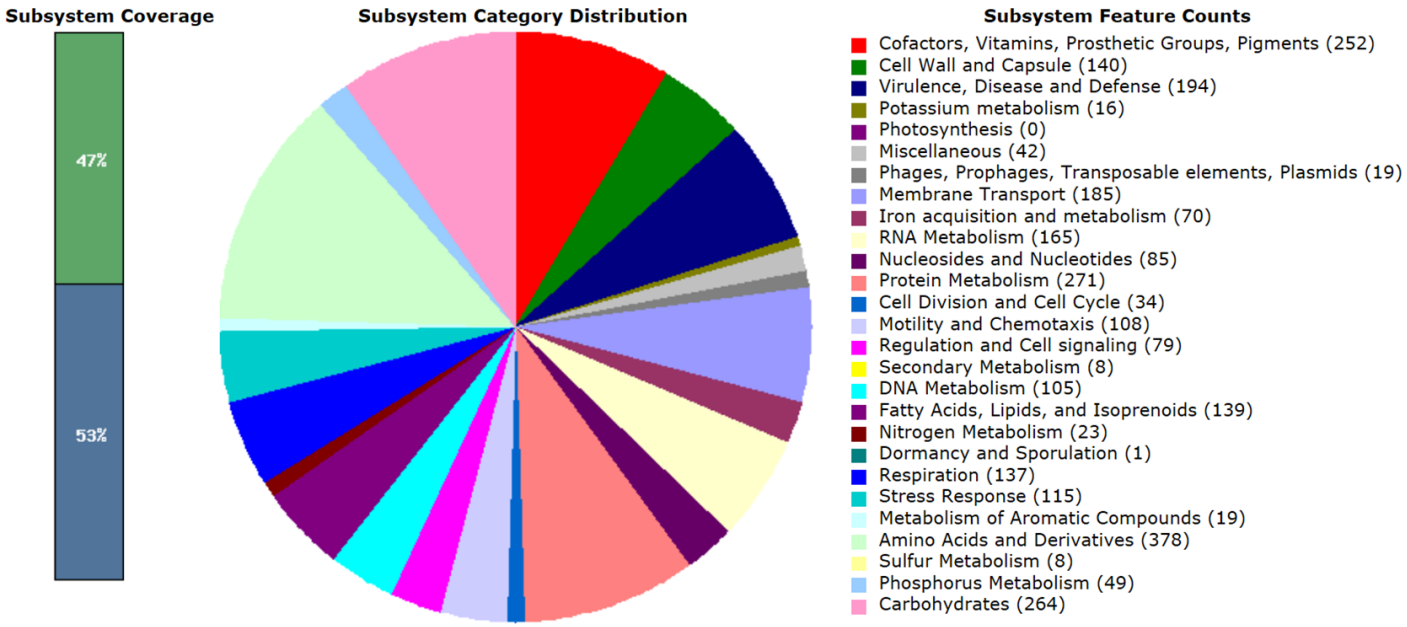

Figure 3.3 *S. maltophilia* JCMS RAST subsystem features

Summary of RAST subsystem features in the JCMS genome. Subsystem coverage indicates the percent of protein-coding sequences that can be classified into subsystem categories. Subsystem feature counts indicate the number of protein-coding genes that can be classified into each category, with distribution of each category represented in the pie-chart.

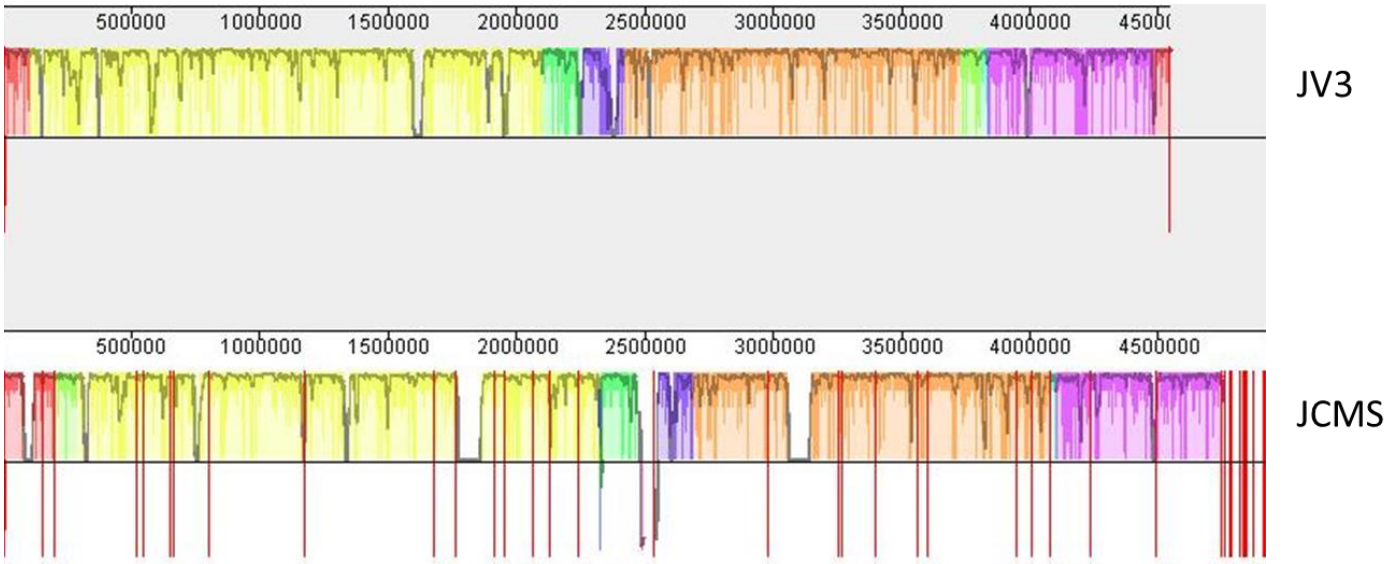

Figure 3.4 *S. maltophilia* JCMS is similar to *S. maltophilia* JV3

JCMS draft (bottom) aligned to JV3 reference (top), resulting in ordered contigs. Plots represent percent similarity between sequences, and different colors represent regions of synteny. Contig boundaries are displayed as red lines.

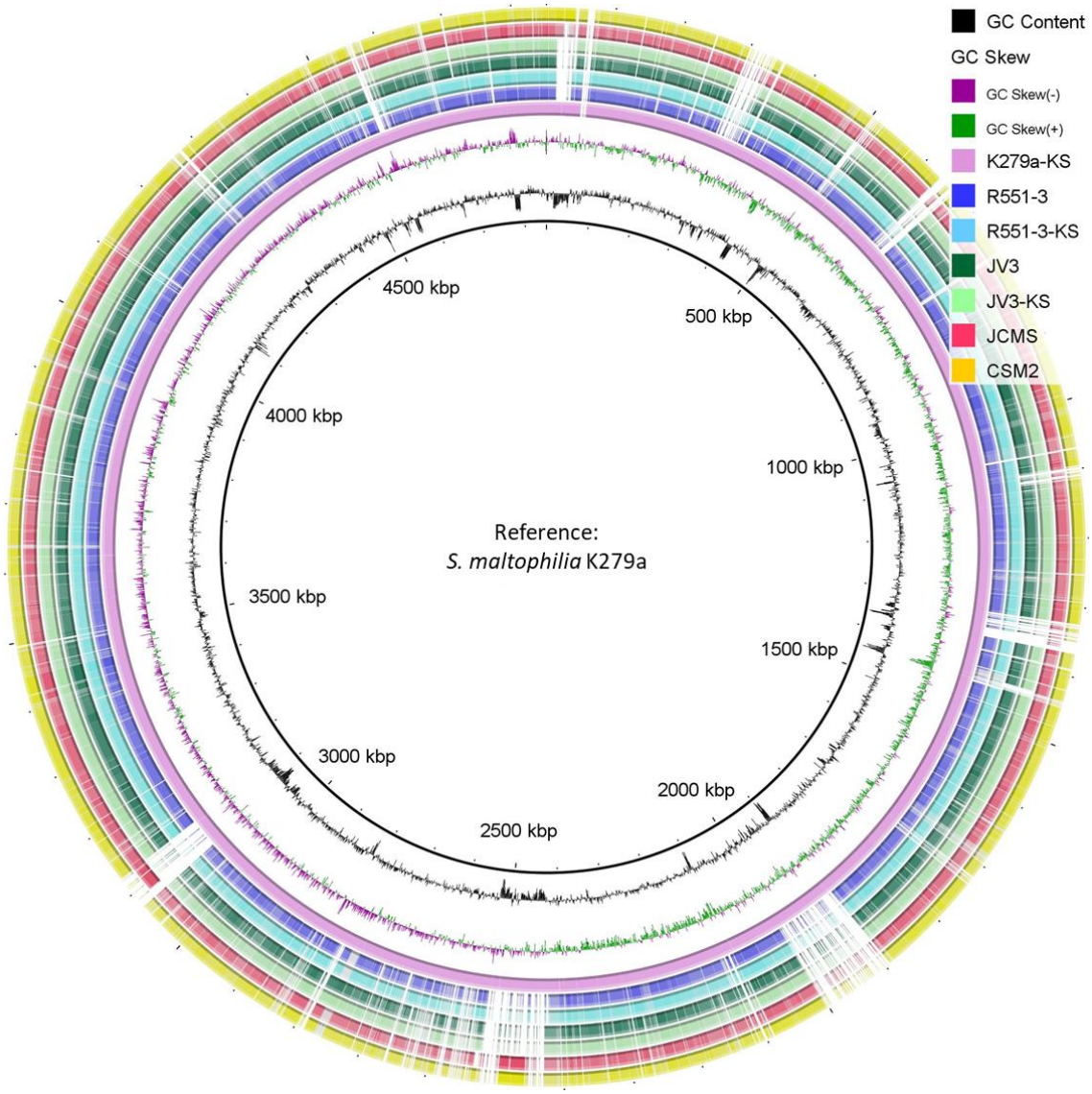

Figure 3.5 Genomic similarity between *S. maltophilia* strains

BLAST was used to align coding sequences, with K279a as a reference. Legend indicates features from inside to outside. Brightness of color for each strain represents similarity to reference sequence, with brighter shades representing higher similarity.

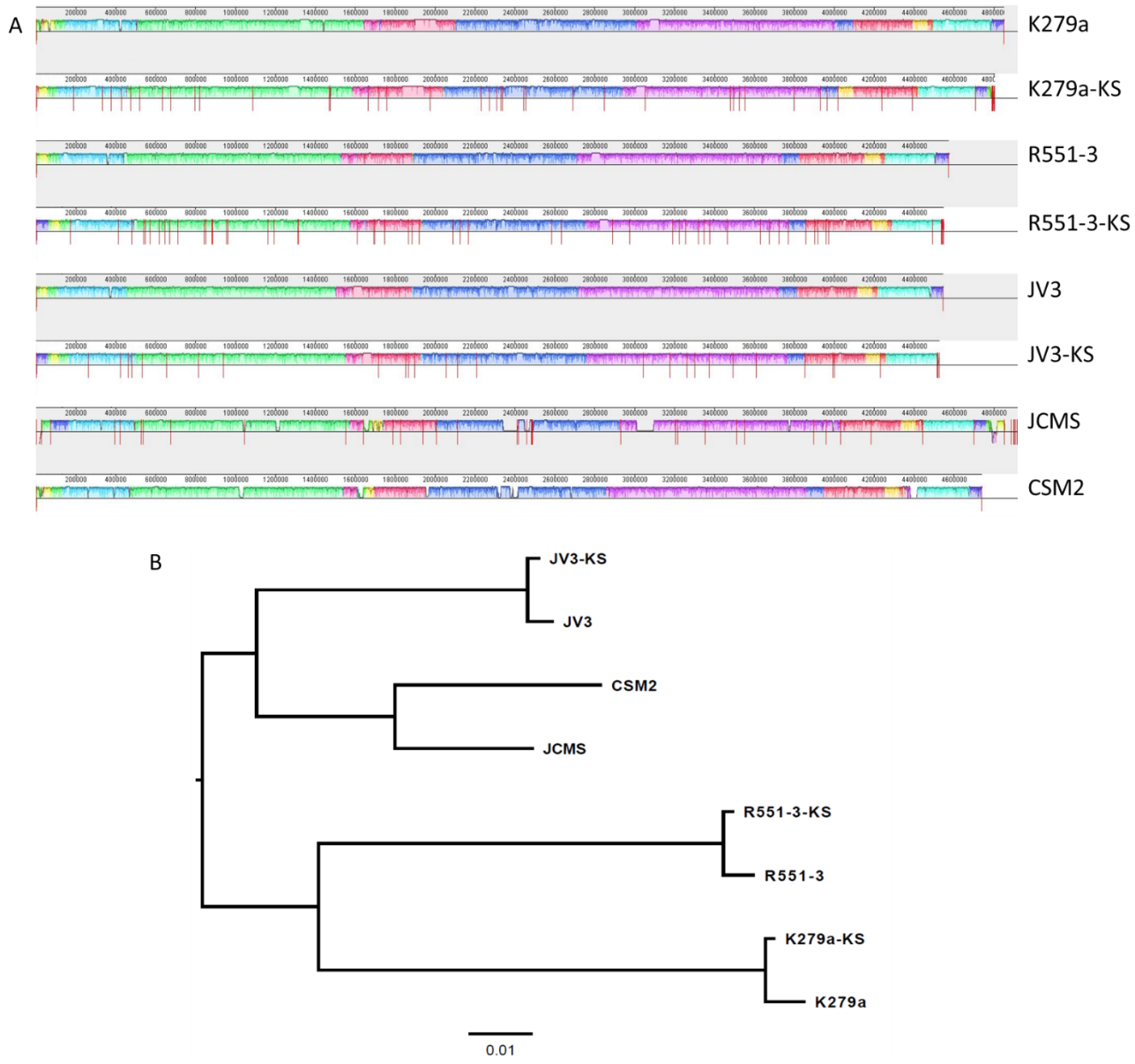

Figure 3.6 Synteny between *S. maltophilia* strains

A Mauve alignment of genome sequences using K279a as a reference. Plots represent percent similarity between sequences, and different colors represent regions of synteny. Contig boundaries are displayed as red lines. **B** Phylogenetic tree representing relationship between *S. maltophilia* strains. Evolutionary distances calculated from Mauve multiple sequence alignments of whole-genome sequences using Neighbor joining method. Distance of the x axis represents amount of genetic change based on number of substitutions between strains. Tree was visualized using FigTree v1.4.3.

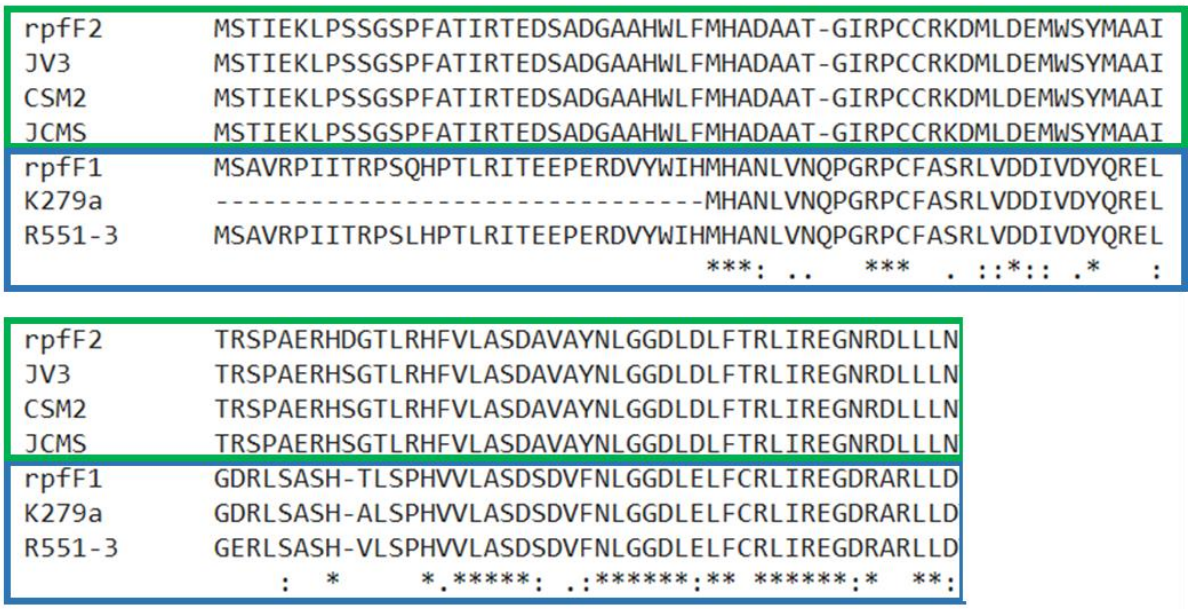

Figure 3.7 RfpF alignment shows segregation of strains based on RfpF variant

RfpF protein sequences of –KS strains were aligned to protein sequences of RfpF-1 (GenBank: KJ149475) and RfpF-2 (GenBank: KJ149552) variants. Clustal Omega v 1.2.4 was used for multiple sequence alignment. Blue outline indicates variant 1 and green outline indicates variant 2.

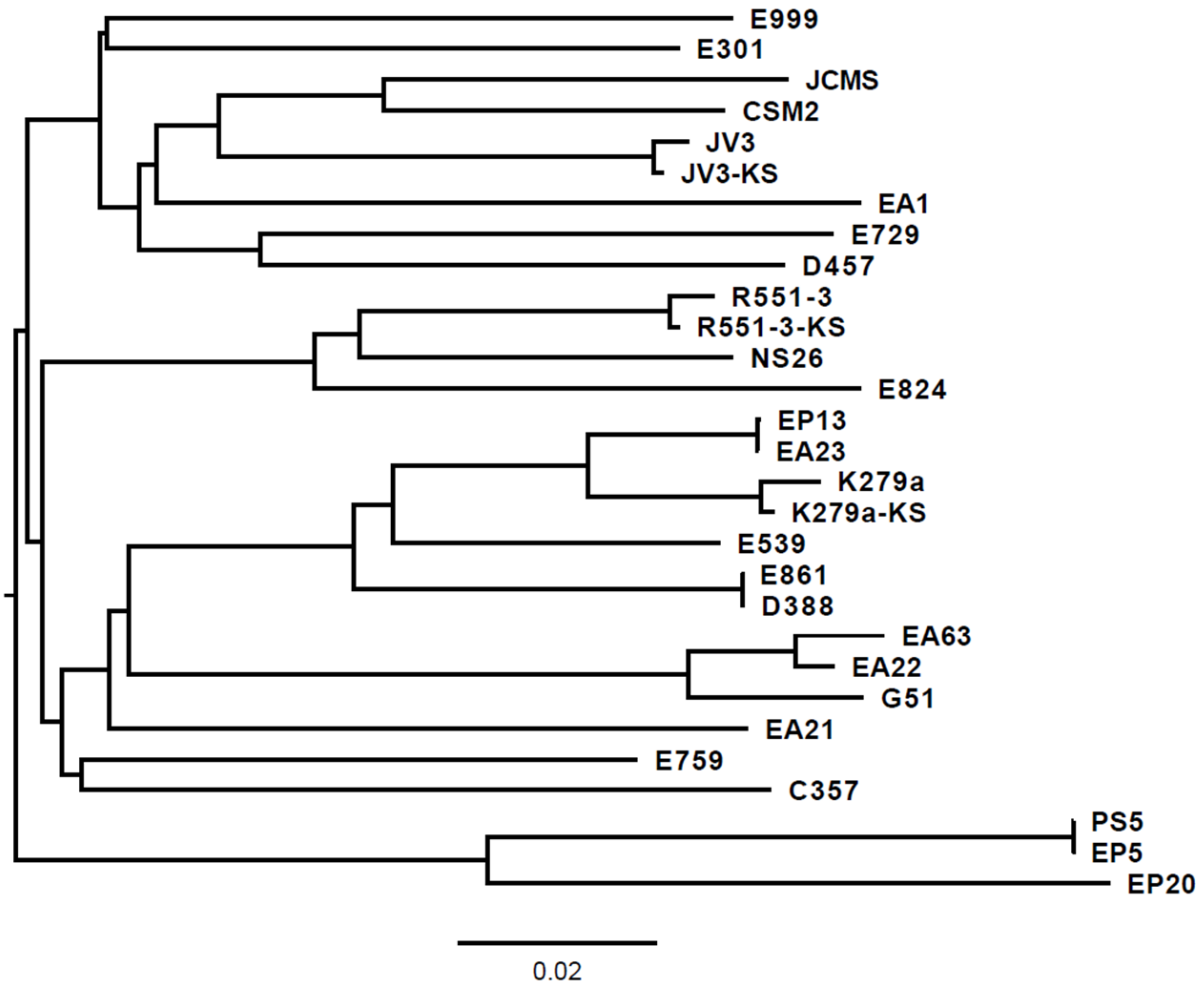

Figure 3.8 Whole genome alignment of strains used for pangenome analysis

Phylogenetic tree representing relationship between 29 *S. maltophilia* strains. Evolutionary distances calculated from Mauve multiple sequence alignments of whole-genome sequences using Neighbor joining method. Distance of the x axis represents amount of genetic change based on number of substitutions between strains. Tree was visualized using FigTree v1.4.3.

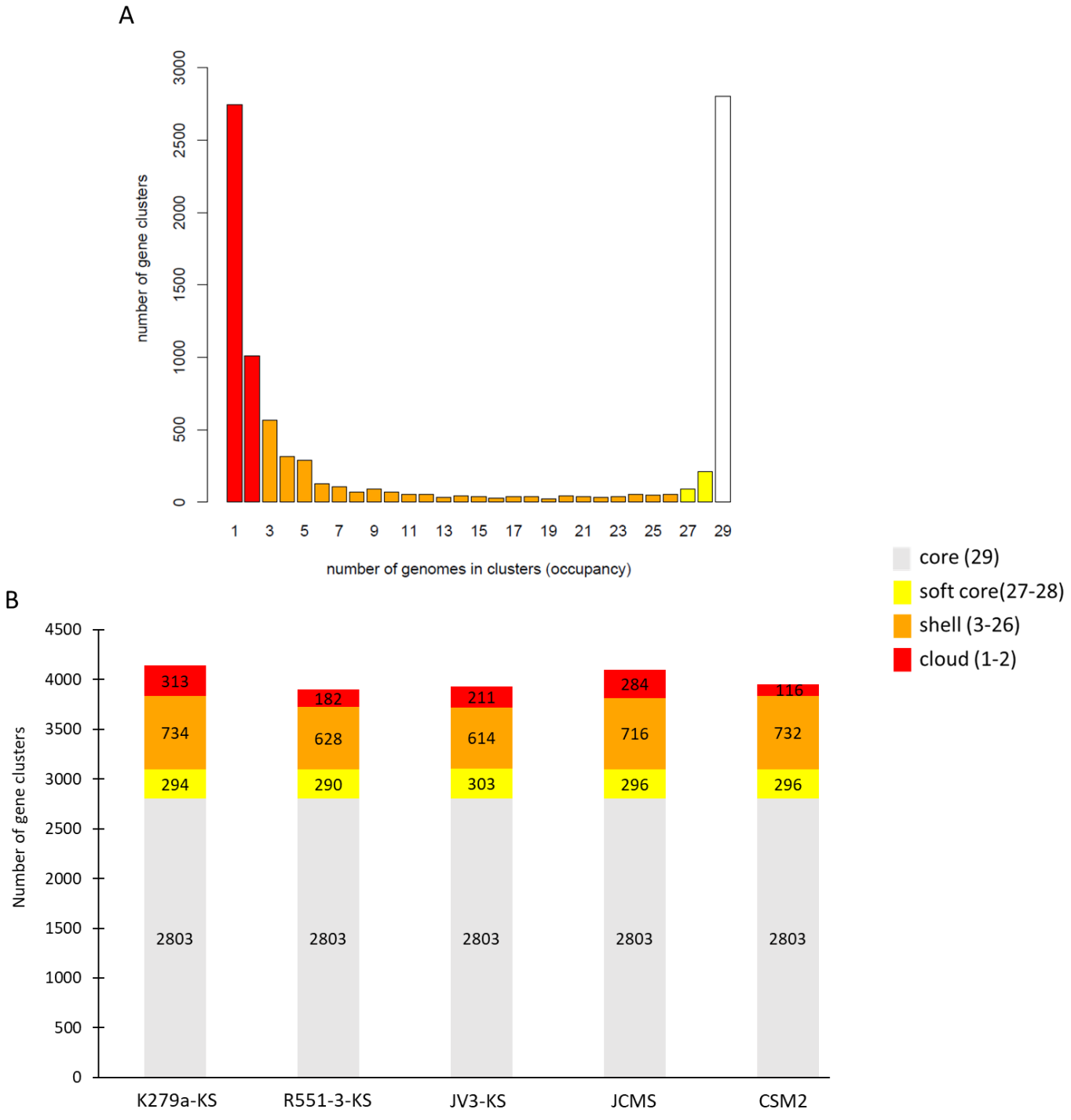

Figure 3.9 Core, soft core, shell, and cloud genomes of 29 *S. maltophilia* strains

A Representation of genes in the core, soft core, shell, and cloud genomes of all 29 strains in the pangenome analysis **B** Distribution of core, soft core, shell, and cloud genes within –KS strains and CSM2.

Tables

Table 3.1 *S. maltophilia* reference and draft genome characteristics.

Draft assembly information (-KS strains and JCMS) is from RAST annotation database. Reference, complete assembly data (indicated by asterisks) is from NCBI genbank information.

Strain	Number of contigs	N50	Size (bp)	GC content (%)	Number of CDS
K279a-KS	48	218290	4,803,885	66.4	4,392
K279a*			4,851,126	66.3	4,386
R551-3-KS	58	172947	4,547,979	66.4	4,109
R551-3*			4,573,969	66.3	4,039
JV3-KS	35	244505	4,526,743	66.9	4,126
JV3*			4,544,477	66.9	4,056
JCMS	42	274796	4,919,076	66.4	4,371

Table 3.2 Subsystem annotation differences between strains

All RAST subsystem annotations that differ between at least one comparison. Category and subsystem are shown for each functional role assigned to protein-coding sequences. Presence/absence matrix is also shown for each strain (**K**=K279a, **R**=R551-3, **V**=JV3, **J**=JCMS), with comparison category shown on the far left.

	Category	Subsystem	Role	K	R	V	J
K279a vs all	Cofactors, Vitamins, Prosthetic Groups, Pigments	Folate biosynthesis cluster	transmembrane protein, distant homology with ydbS	+	-	-	-
	Miscellaneous	Phosphoglycerate mutase protein family	Phosphoglycerate mutase family	+	-	-	-
	Phages, Prophages, TEs, Plasmids	Phage tail proteins	Phage tail completion protein	+	-	-	-
	Phages, Prophages, TEs, Plasmids	Phage packaging machinery	Phage terminase small subunit	+	-	-	-
	Phages, Prophages, TEs, Plasmids	Phage packaging machinery	Phage terminase, ATPase subunit	+	-	-	-
	Phages, Prophages, TEs, Plasmids	Phage packaging machinery	Phage terminase, endonuclease subunit	+	-	-	-
	Phages, Prophages, TEs, Plasmids	Phage capsid proteins	Phage head completion-stabilization protein	+	-	-	-
	Phages, Prophages, TEs, Plasmids	Phage capsid proteins	Phage capsid scaffolding protein	+	-	-	-
	Membrane Transport	Choline Transport	Sodium-Choline Symporter	+	-	-	-
	Membrane Transport	pVir Plasmid of Campylobacter	Type IV secretion system protein VirD4	+	-	-	-
	Membrane Transport	Conjugative transfer	Conjugative transfer protein TrbG	+	-	-	-
	Membrane Transport	Conjugative transfer	IncF plasmid conjugative transfer protein TraD	+	-	-	-
	Membrane Transport	Conjugative transfer	Conjugative transfer protein TrbD	+	-	-	-
	Membrane Transport	Conjugative transfer	Conjugative transfer protein TrbL	+	-	-	-
	Membrane Transport	Conjugative transfer	Conjugative transfer protein TrbF	+	-	-	-
	Membrane Transport	Conjugative transfer	Conjugative transfer protein TrbC	+	-	-	-
	Membrane Transport	Conjugative transfer	Conjugative transfer protein TrbI	+	-	-	-
	Membrane Transport	Conjugative transfer	Conjugative transfer protein TrbE	+	-	-	-
	Membrane Transport	Conjugative transfer	Conjugative transfer protein TrbJ	+	-	-	-

Membrane Transport	Conjugative transfer	Conjugative transfer protein TrbB	+	-	-	-
Iron acquisition and metabolism	Encapsulating protein for DyP-type peroxidase and ferritin-like protein oligomers	Predicted dye-decolorizing peroxidase (DyP), encapsulated subgroup	+	-	-	-
Regulation and Cell signaling	Trans-envelope signaling system VreARI in Pseudomonas	Filamentous haemagglutinin family outer membrane protein associated with VreARI signalling system	+	-	-	-
Regulation and Cell signaling	Trans-envelope signaling system VreARI in Pseudomonas	Sigma factor regulator VreR (cytoplasmic membrane-localized) of trans-envelope signaling system	+	-	-	-
Regulation and Cell signaling	Trans-envelope signaling system VreARI in Pseudomonas	Putative outer membrane TonB-dependent receptor associated with haemagglutinin family outer membrane protein	+	-	-	-
Regulation and Cell signaling	Trans-envelope signaling system VreARI in Pseudomonas	Extracytoplasmic function (ECF) sigma factor VreI	+	-	-	-
Regulation and Cell signaling	Trans-envelope signaling system VreARI in Pseudomonas	Outer membrane TonB-dependent transducer VreA of trans-envelope signaling system	+	-	-	-
Regulation and Cell signaling	Trans-envelope signaling system VreARI in Pseudomonas	Hemolysin activation/secretion protein associated with VreARI signalling system	+	-	-	-
DNA Metabolism	DNA repair, bacterial	Methyl-directed repair DNA adenine methylase (EC 2.1.1.72)	+	-	-	-
Fatty Acids, Lipids, and Isoprenoids	Glycerolipid and Glycerophospholipid Metabolism in Bacteria	Glycerate kinase (EC 2.7.1.31)	+	-	-	-
Stress Response	Oxidative stress	Alkyl hydroperoxide reductase subunit C-like protein	+	-	-	-
Stress Response	Cold shock, CspA family of proteins	Cold shock protein CspC	+	-	-	-
Amino Acids and Derivatives	Threonine degradation	Threonine dehydrogenase and related Zn-dependent dehydrogenases	+	-	-	-
Phosphorus Metabolism	Phosphate-binding DING proteins	RecA/RadA recombinase	+	-	-	-
Phosphorus Metabolism	Phosphate-binding DING proteins	Phosphate-binding DING protein (related to PstS)	+	-	-	-
Carbohydrates	Pyruvate Alanine Serine Interconversions	Omega-amino acid--pyruvate aminotransferase (EC 2.6.1.18)	+	-	-	-
Carbohydrates	Trehalose Biosynthesis	Malto-oligosyltrehalose synthase (EC 5.4.99.15)	+	-	-	-
Carbohydrates	Trehalose Biosynthesis	1,4-alpha-glucan (glycogen) branching enzyme, GH-13-type (EC 2.4.1.18)	+	-	-	-
Carbohydrates	Trehalose Biosynthesis	Trehalose synthase (EC 5.4.99.16)	+	-	-	-
Carbohydrates	Trehalose Biosynthesis	Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141)	+	-	-	-
Carbohydrates	Trehalose Biosynthesis	Glycogen debranching enzyme (EC 3.2.1.-)	+	-	-	-
Carbohydrates	Maltose and Maltodextrin Utilization	Transcriptional regulator of maltose utilization, LacI family	+	-	-	-

	Carbohydrates	Maltose and Maltodextrin Utilization	Predicted maltose transporter MalT	+	-	-	-
	Carbohydrates	Maltose and Maltodextrin Utilization	4-alpha-glucanotransferase (amylomaltase) (EC 2.4.1.25)	+	-	-	-
	Carbohydrates	Maltose and Maltodextrin Utilization	Maltodextrin glucosidase (EC 3.2.1.20)	+	-	-	-
	Carbohydrates	Maltose and Maltodextrin Utilization	Predicted maltose-specific TonB-dependent receptor	+	-	-	-
	Carbohydrates	Maltose and Maltodextrin Utilization	Multiple sugar ABC transporter, ATP-binding protein	+	-	-	-
	Carbohydrates	Glycogen metabolism	Glycogen synthase, ADP-glucose transglucosylase (EC 2.4.1.21)	+	-	-	-
	Cell Wall and Capsule	Rhamnose containing glycans	Teichoic acid export ATP-binding protein TagH (EC 3.6.3.40)	-	+	+	+
	Virulence, Disease and Defense	Multidrug Resistance Efflux Pumps	RND efflux system, membrane fusion protein CmeA	-	+	+	+
	Phages, Prophages, TEs, Plasmids	Phage tail fiber proteins	Phage tail fiber protein	-	+	+	+
	Membrane Transport	Ton and Tol transport systems	Outer membrane receptor for ferric coprogen and ferric-rhodotorulic acid	-	+	+	+
	Membrane Transport	Ton and Tol transport systems	Colicin I receptor precursor	-	+	+	+
	Fatty Acids, Lipids, and Isoprenoids	Acyl-CoA thioesterase II	TesB-like acyl-CoA thioesterase 5	-	+	+	+
	Nitrogen Metabolism	Nitrate and nitrite ammonification	Nitrate/nitrite transporter	-	+	+	+
	Stress Response	Oxidative stress	Superoxide dismutase [Cu-Zn] precursor (EC 1.15.1.1)	-	+	+	+
	Metabolism of Aromatic Compounds	Salicylate ester degradation	Salicylate hydroxylase (EC 1.14.13.1)	-	+	+	+
	Nitrogen Metabolism	Nitrosative stress	NnrS protein involved in response to NO	-	+	-	-
	Dormancy and Sporulation	Persister Cells	HipA protein	-	+	-	-
	Amino Acids and Derivatives	Lysine degradation	Lysine/cadaverine antiporter membrane protein CadB	-	+	-	-
	Carbohydrates	Glycerol and Glycerol-3-phosphate Uptake and Utilization	Glycerol-3-phosphate ABC transporter, ATP-binding protein UgpC (TC 3.A.1.1.3)	-	+	-	-
R551 vs all	Cofactors, Vitamins, Prosthetic Groups, Pigments	Thiamin biosynthesis	Hydroxymethylpyrimidine ABC transporter, substrate-binding component	+	-	+	+
	Cofactors, Vitamins, Prosthetic Groups, Pigments	Thiamin biosynthesis	Hydroxymethylpyrimidine ABC transporter, transmembrane component	+	-	+	+
	Cofactors, Vitamins, Prosthetic Groups, Pigments	NAD and NADP cofactor biosynthesis global	Ribosylnicotinamide kinase (EC 2.7.1.22)	+	-	+	+

	Cofactors, Vitamins, Prosthetic Groups, Pigments	NAD and NADP cofactor biosynthesis global	Nicotinamide phosphoribosyltransferase (EC 2.4.2.12)	+	-	+	+
	Cofactors, Vitamins, Prosthetic Groups, Pigments	NAD and NADP cofactor biosynthesis global	Nicotinamide-nucleotide adenylyltransferase, NadM family (EC 2.7.7.1)	+	-	+	+
	Cofactors, Vitamins, Prosthetic Groups, Pigments	NAD and NADP cofactor biosynthesis global	Nicotinamide-nucleotide adenylyltransferase, NadR family (EC 2.7.7.1)	+	-	+	+
	Virulence, Disease and Defense	Cobalt-zinc-cadmium resistance	DNA-binding heavy metal response regulator	+	-	+	+
	Regulation and Cell signaling	LysR-family proteins in Escherichia coli	LysR family transcriptional regulator YbhD	+	-	+	+
	DNA Metabolism	Uracil-DNA glycosylase	Uracil-DNA glycosylase, putative family 6	+	-	+	+
	DNA Metabolism	Uracil-DNA glycosylase	Domain often clustered or fused with uracil-DNA glycosylase	+	-	+	+
	DNA Metabolism	DNA Repair Base Excision	DNA-3-methyladenine glycosylase II (EC 3.2.2.21)	+	-	+	+
	Fatty Acids, Lipids, and Isoprenoids	Glycerolipid and Glycerophospholipid Metabolism in Bacteria	Phosphatidylglycerophosphatase B (EC 3.1.3.27)	+	-	+	+
	Stress Response	Osmotic stress cluster	Heat shock (predicted periplasmic) protein YciM, precursor	+	-	+	+
	Amino Acids and Derivatives	Polyamine Metabolism	Arginine/agmatine antiporter	+	-	+	+
	Amino Acids and Derivatives	Isoleucine degradation	Acyl-CoA dehydrogenase, short-chain specific (EC 1.3.99.2)	+	-	+	+
	Phosphorus Metabolism	Phosphate metabolism	Phosphate-specific outer membrane porin OprP	+	-	+	+
	Cofactors, Vitamins, Prosthetic Groups, Pigments	Molybdenum cofactor biosynthesis	Xanthine and CO dehydrogenases maturation factor, XdhC/CoxF family	+	+	+	-
	Phages, Prophages, TEs, Plasmids	Phage packaging machinery	Phage terminase, large subunit	+	+	+	-
	Nucleosides and Nucleotides	Purine Utilization	Periplasmic aromatic aldehyde oxidoreductase, FAD binding subunit YagS	+	+	+	-
JCMS vs all	Cell Wall and Capsule	dTDP-rhamnose synthesis	dTDP-rhamnosyl transferase RfbF (EC 2.-.-.-)	-	-	-	+
	RNA Metabolism	Ribonucleases in Bacillus	Metallo-beta-lactamase family protein, RNA-specific	-	-	-	+
	Regulation and Cell signaling	cAMP signaling in bacteria	cyclolysin secretion ATP-binding protein	-	-	-	+
	Regulation and Cell signaling	Orphan regulatory proteins	Glycine cleavage system transcriptional activator GcvA	-	-	-	+

	DNA Metabolism	DNA topoisomerases, Type I, ATP-independent	DNA topoisomerase III (EC 5.99.1.2) in PFGI-1-like cluster	-	-	-	+
	Stress Response	Glutathione-dependent pathway of formaldehyde detoxification	Glutathione-dependent formaldehyde-activating enzyme (EC 4.4.1.22)	-	-	-	+
	Stress Response	Universal stress protein family	Universal stress protein family 4	-	-	-	+
JV3 vs all	Cofactors, Vitamins, Prosthetic Groups, Pigments	Riboflavin, FMN and FAD metabolism in plants	3,4-dihydroxy-2-butanone 4-phosphate synthase (EC 4.1.99.12)	+	+	-	+
	Cell Wall and Capsule	Rhamnose containing glycans	Alpha-L-Rha alpha-1,3-L-rhamnosyltransferase (EC 2.4.1.-)	+	+	-	+
	Membrane Transport	Two partner secretion pathway (TPS)	Fimbrial adhesin	+	+	-	+
	Membrane Transport	Two partner secretion pathway (TPS)	Channel-forming transporter/cytolysins activator of TpsB family	+	+	-	+
	Membrane Transport	Two partner secretion pathway (TPS)	Putative large exoprotein involved in heme utilization or adhesion of ShlA/HecA/FhaA family	+	+	-	+
	Nucleosides and Nucleotides	GMP synthase	GMP synthase [glutamine-hydrolyzing], ATP pyrophosphatase subunit (EC 6.3.5.2)	+	+	-	+
	Nucleosides and Nucleotides	GMP synthase	GMP synthase [glutamine-hydrolyzing], amidotransferase subunit (EC 6.3.5.2)	+	+	-	+
	Nucleosides and Nucleotides	GMP synthase	GMP synthase (EC 6.3.5.2)	+	+	-	+
	DNA Metabolism	DNA topoisomerases, Type I, ATP-independent	Hypothetical protein PA2244 (similar to DNA topoisomerase IB, but possibly involved in glycosyl-transfer)	+	+	-	+
	Cofactors, Vitamins, Prosthetic Groups, Pigments	Molybdenum cofactor biosynthesis	DNA-binding domain of ModE	-	-	+	-
	Cofactors, Vitamins, Prosthetic Groups, Pigments	Molybdenum cofactor biosynthesis	Molybdate-binding domain of ModE	-	-	+	-
	Nucleosides and Nucleotides	Purine conversions	GMP synthase [glutamine-hydrolyzing] (EC 6.3.5.2)	-	-	+	-
	Secondary Metabolism	Lanthionine Synthetases	Lanthionine biosynthesis protein LanL	-	-	+	-
	DNA Metabolism	DNA topoisomerases, Type I, ATP-independent	DNA topoisomerase IB (poxvirus type) (EC 5.99.1.2)	-	-	+	-
	Carbohydrates	Chitin and N-acetylglucosamine utilization	N-Acetyl-D-glucosamine ABC transport system, permease protein 2	-	-	+	-

K279a/ R551-3 vs JCMS/ JV3	Cofactors, Vitamins, Prosthetic Groups, Pigments	Molybdenum cofactor biosynthesis	Protein co-occuring with molybdenum cofactor biosynthesis protein B	-	-	+	+
	Virulence, Disease and Defense	Copper homeostasis	Cu(I)-responsive transcriptional regulator	-	-	+	+
	Virulence, Disease and Defense	Cobalt-zinc-cadmium resistance	Heavy metal sensor histidine kinase	-	-	+	+
	Phages, Prophages, TEs, Plasmids	Phage tail proteins	Phage tail assembly protein I	-	-	+	+
	Phages, Prophages, TEs, Plasmids	Phage tail proteins	Phage tail length tape-measure protein 1	-	-	+	+
	Phages, Prophages, TEs, Plasmids	Phage tail proteins	Phage minor tail protein	-	-	+	+
	Phages, Prophages, TEs, Plasmids	Phage tail proteins	Phage tail assembly protein	-	-	+	+
	Phages, Prophages, TEs, Plasmids	Phage lysis modules	Phage lysin, 1,4-beta-N-acetylmuramidase (EC 3.2.1.17) or lysozyme	-	-	+	+
	Membrane Transport	Widespread colonization island	Type II/IV secretion system protein TadC, associated with Flp pilus assembly	-	-	+	+
	Motility and Chemotaxis	Bacterial Chemotaxis	Maltose/maltodextrin ABC transporter, substrate binding periplasmic protein MalE	-	-	+	+
	Regulation and Cell signaling	LysR-family proteins in Salmonella enterica Typhimurium	LysR family transcriptional regulator STM2281	-	-	+	+
	Regulation and Cell signaling	DNA-binding regulatory proteins, strays	Aromatic hydrocarbon utilization transcriptional regulator CatR (LysR family)	-	-	+	+
	Regulation and Cell signaling	Toxin-antitoxin replicon stabilization systems	HigB toxin protein	-	-	+	+
	Metabolism of Aromatic Compounds	Biphenyl Degradation	biphenyl-2,3-diol 1,2-dioxygenase III-related protein	-	-	+	+
	Carbohydrates	Alpha-Amylase locus in Streptococcus	putative esterase	-	-	+	+
	Cofactors, Vitamins, Prosthetic Groups, Pigments	Thiamin biosynthesis	Phosphomethylpyrimidine kinase (EC 2.7.4.7)	+	+	-	-
	Cofactors, Vitamins, Prosthetic Groups, Pigments	Thiamin biosynthesis	Thiamin biosynthesis protein ThiC	+	+	-	-
	Cofactors, Vitamins, Prosthetic Groups, Pigments	Ubiquinone Biosynthesis - gjo	Uncharacterized hydroxylase PA0655	+	+	-	-

Cell Wall and Capsule	Alginate metabolism	Alginate lyase precursor (EC 4.2.2.3)	+	+	-	-
Cell Wall and Capsule	Alginate metabolism	Mannose-6-phosphate isomerase (EC 5.3.1.8)	+	+	-	-
Cell Wall and Capsule	Alginate metabolism	Poly (beta-D-mannuronate) lyase (EC 4.2.2.3)	+	+	-	-
Cell Wall and Capsule	Lipopolysaccharide assembly	Outer membrane lipoprotein LolB	+	+	-	-
Cell Wall and Capsule	Lipid A-Ara4N pathway (Polymyxin resistance)	Polymyxin resistance protein PmrL, sucrose-6 phosphate hydrolase	+	+	-	-
Cell Wall and Capsule	Lipid A-Ara4N pathway (Polymyxin resistance)	Polymyxin resistance protein ArnT, undecaprenyl phosphate-alpha-L-Ara4N transferase	+	+	-	-
Cell Wall and Capsule	Lipid A-Ara4N pathway (Polymyxin resistance)	UDP-glucose 6-dehydrogenase (EC 1.1.1.22)	+	+	-	-
Cell Wall and Capsule	Peptidoglycan Biosynthesis	D-alanine--D-alanine ligase A (EC 6.3.2.4)	+	+	-	-
Cell Wall and Capsule	YjeE	Phosphotransferase involved in threonylcarbamoyladenosine t(6)A37 formation in tRNA	+	+	-	-
Virulence, Disease and Defense	Copper homeostasis: copper tolerance	Cytoplasmic copper homeostasis protein CutC	+	+	-	-
Potassium metabolism	Potassium homeostasis	Glutathione-regulated potassium-efflux system protein KefC	+	+	-	-
Phages, Prophages, TEs, Plasmids	Phage tail proteins	Phage major tail tube protein	+	+	-	-
Phages, Prophages, TEs, Plasmids	Phage tail proteins	Phage tail sheath monomer	+	+	-	-
Phages, Prophages, TEs, Plasmids	Phage tail proteins	Phage tail protein	+	+	-	-
Phages, Prophages, TEs, Plasmids	Phage packaging machinery	Phage portal protein	+	+	-	-
Phages, Prophages, TEs, Plasmids	Phage tail fiber proteins	Phage tail fibers	+	+	-	-
Phages, Prophages, TEs, Plasmids	Phage capsid proteins	Phage major capsid protein	+	+	-	-
Membrane Transport	Fap amyloid fiber secretion system	Fap amyloid fiber secretin	+	+	-	-
Membrane Transport	Fap amyloid fiber secretion system	Sigma-54 dependent transcriptional regulator	+	+	-	-
Membrane Transport	Fap amyloid fiber secretion system	Fap amyloid fibril major component	+	+	-	-
Membrane Transport	Fap amyloid fiber secretion system	Fap unknown function protein-Stenotrophomonas type	+	+	-	-
Membrane Transport	Fap amyloid fiber secretion system	Fap protein with C39 domain	+	+	-	-
Membrane Transport	Fap amyloid fiber secretion system	Fap system putative outer membrane protein	+	+	-	-
Membrane Transport	Fap amyloid fiber secretion system	Fap amyloid fibril minor component	+	+	-	-

Membrane Transport	Vir-like type 4 secretion system	ATPase required for both assembly of type IV secretion complex and secretion of T-DNA complex, VirB11	+	+	-	-
Membrane Transport	Vir-like type 4 secretion system	Minor pilin of type IV secretion complex, VirB5	+	+	-	-
Membrane Transport	Vir-like type 4 secretion system	ATPase required for both assembly of type IV secretion complex and secretion of T-DNA complex, VirB4	+	+	-	-
Membrane Transport	Vir-like type 4 secretion system	Inner membrane protein forms channel for type IV secretion of T-DNA complex, VirB8	+	+	-	-
Membrane Transport	Vir-like type 4 secretion system	Peptidoglycan hydrolase VirB1, involved in T-DNA transfer	+	+	-	-
Membrane Transport	Vir-like type 4 secretion system	Inner membrane protein forms channel for type IV secretion of T-DNA complex, VirB3	+	+	-	-
Membrane Transport	Vir-like type 4 secretion system	Coupling protein VirD4, ATPase required for T-DNA transfer	+	+	-	-
Membrane Transport	Vir-like type 4 secretion system	Inner membrane protein of type IV secretion of T-DNA complex, TonB-like, VirB10	+	+	-	-
Membrane Transport	Vir-like type 4 secretion system	Major pilus subunit of type IV secretion complex, VirB2	+	+	-	-
Membrane Transport	Vir-like type 4 secretion system	Outer membrane and periplasm component of type IV secretion of T-DNA complex, has secretin-like domain, VirB9	+	+	-	-
Membrane Transport	Vir-like type 4 secretion system	Inner membrane protein of type IV secretion of T-DNA complex, VirB6	+	+	-	-
RNA Metabolism	Possible RNA modification and stress response cluster	Organic solvent tolerance protein precursor	+	+	-	-
RNA Metabolism	Possible RNA modification and stress response cluster	Bis(5'-nucleosyl)-tetraphosphatase, symmetrical (EC 3.6.1.41)	+	+	-	-
RNA Metabolism	Possible RNA modification and stress response cluster	ApaG protein	+	+	-	-
RNA Metabolism	RNA methylation	16S rRNA (cytosine(967)-C(5))-methyltransferase (EC 2.1.1.176)	+	+	-	-
DNA Metabolism	DNA repair, bacterial	DNA polymerase IV-like protein ImuB	+	+	-	-
DNA Metabolism	DNA replication strays	Error-prone repair homolog of DNA polymerase III alpha subunit (EC 2.7.7.7)	+	+	-	-
Fatty Acids, Lipids, and Isoprenoids	Isoprenoinds for Quinones	Geranylgeranyl diphosphate synthase (EC 2.5.1.29)	+	+	-	-
Stress Response	Periplasmic Stress Response	Intramembrane protease RasP/YluC, implicated in cell division based on FtsL cleavage	+	+	-	-
Amino Acids and Derivatives	Arginine and Ornithine Degradation	Delta-1-pyrroline-5-carboxylate dehydrogenase (EC 1.2.1.88)	+	+	-	-
Amino Acids and Derivatives	Lysine degradation	L-lysine permease	+	+	-	-

	Carbohydrates	Entner-Doudoroff Pathway	2-dehydro-3-deoxygluconate kinase (EC 2.7.1.45)	+	+	-	-
	Carbohydrates	D-gluconate and ketogluconates metabolism	Broad-specificity glycerol dehydrogenase (EC 1.1.99.22), subunit SldB	+	+	-	-
K279a/ JV3 vs R551-3/ JCMS	Phages, Prophages, TEs, Plasmids	Phage tail proteins	Phage tail length tape-measure protein	+	-	+	-
	Virulence, Disease and Defense	Resistance to chromium compounds	Chromate transport protein ChrA	-	+	-	+
	Regulation and Cell signaling	Phd-Doc, YdcE-YdcD toxin-antitoxin (programmed cell death) systems	Death on curing protein, Doc toxin	-	+	-	+
K279a/ JCMS vs R551-3/ JV3	Virulence, Disease and Defense	Copper homeostasis	Copper resistance protein CopC	+	-	-	+
	Virulence, Disease and Defense	Copper homeostasis	Copper resistance protein CopD	+	-	-	+
	Virulence, Disease and Defense	Cobalt-zinc-cadmium resistance	Heavy metal resistance transcriptional regulator HmrR	+	-	-	+
	Virulence, Disease and Defense	Mercuric reductase	Mercuric ion reductase (EC 1.16.1.1)	+	-	-	+
	Virulence, Disease and Defense	Mercury resistance operon	Mercuric transport protein, MerT	+	-	-	+
	Virulence, Disease and Defense	Mercury resistance operon	Periplasmic mercury(+2) binding protein	+	-	-	+
	Virulence, Disease and Defense	Mercury resistance operon	Mercuric resistance operon regulatory protein	+	-	-	+
	Phages, Prophages, TEs, Plasmids	Phage replication	DNA polymerase III alpha subunit (EC 2.7.7.7)	+	-	-	+
	Phages, Prophages, TEs, Plasmids	Phage replication	DNA primase/helicase, phage-associated	+	-	-	+
	Amino Acids and Derivatives	Histidine Degradation	Formiminoglutamase (EC 3.5.3.8)	+	-	-	+

Chapter 4 - Summary and Future Directions

In the natural environment, *C. elegans* interact with a variety of organisms, and these relationships come in a variety of forms. Snails and isopods act as a vector for transportation of nematodes between locations, while nematodes serve as prey for insects and fungi. In addition, *C. elegans* encounter thousands of taxa of bacteria that can form commensalistic, mutualistic, or pathogenic interactions with *C. elegans*. Our lab focuses on bacteria that are pathogenic to *C. elegans* and the genetic mechanisms behind these interactions. We use *S. maltophilia* to study these host-pathogen interactions as members of *Stenotrophomonas* are found in relatively high abundance in the microbiome of *C. elegans* and most strains are either detrimental or neutral to the health of *C. elegans* (Dirksen et al, 2016; Samuel et al., 2016). In addition to this natural interaction, *S. maltophilia* is an emerging nosocomial pathogen known for causing infection in patients with respiratory diseases such as cystic fibrosis. Therefore, understanding the genetic basis of *S. maltophilia* pathogenesis using *C. elegans* has both ecological and medical implications. Throughout this study, we utilized several strains of *S. maltophilia* that display varying pathogenicity to *C. elegans* in order to relate phenotypic virulence to genomic differences in *S. maltophilia* and genotypic responses in *C. elegans*.

Host gene expression reveals common and strain-specific responses to *S. maltophilia*

One goal of this thesis was to elucidate host responses upon *S. maltophilia* infection. We found that *S. maltophilia* induces both common and strain-specific responses in *C. elegans*. Specifically, it seems that the degree of differential expression correlates with level of virulence, as the more virulent strain, JV3, resulted in higher overall expression of genes induced by both pathogenic *S. maltophilia* strains JV3 and JCMS. In addition, JV3 resulted in downregulation of

a variety of genes in *C. elegans* that were not downregulated in response to other strains. Many of these genes appear to be involved in oxidoreductase and glycosyltransferase activities. These enzymes play a role in a variety of metabolic processes; thus, one explanation for the increased virulence of JV3 might be its ability to repress overall metabolism. To provide more insight into this hypothesis, functional analyses of individual genes could elucidate their role in pathogen response. In addition, metabolomic analyses of *C. elegans* upon exposure to different *S. maltophilia* strains would determine if and what metabolic pathways are affected.

Functional analyses of candidate genes differentially expressed in response to one or both strains of pathogenic *S. maltophilia* revealed that while a majority of these genes are important for survival upon *S. maltophilia* infection, differences in survival were not dependent on the strain that resulted in differential expression of that gene. Therefore, although gene expression is strain specific, these genes may be important for survival on a broader scale. Utilizing more *S. maltophilia* strains for transcriptomic and functional analysis of more genes could uncover more concrete patterns in the complex dynamics of *C. elegans* responses to *S. maltophilia*.

Elucidating the role of the DAF-2/16 insulin-like pathway in response to *S. maltophilia*

The DAF-2/16 insulin-like pathway has been shown to play a role in longevity and defense responses (Kenyon et al., 1993; McElwee, Bubb, & Thomas, 2003; Murphy et al., 2003). Previously, we determined that this pathway does not seem to be effective in response to *S. maltophilia* JCMS (White et al., 2016), and this is one of the only reported environmental conditions in which *daf-2* mutants do not result in increased lifespan in *C. elegans*. However, further analysis is needed to identify whether this is a strain-specific phenomenon or a common characteristic of *S. maltophilia*. To do this, we obtained a variety of *S. maltophilia* strains that

were identified as part of the *C. elegans* native microbiome. Survival analyses comparing *wild-type* and *daf-2* mutant *C. elegans* upon exposure to these strains will reveal whether the DAF-2/16 pathway is an effective defense mechanism in response to each strain.

In addition, to further explore this phenomenon, our transcriptomic experiment described in Chapter 2 also included *daf-2* mutants exposed to each bacterial strain. Comparison of differentially expressed genes in *wild-type* and *daf-2* mutant *C. elegans* upon exposure to each bacterial strain could identify genes that are independent and dependent on DAF-2. Mutants of these genes can be used in survival assays to determine if they are important for response to *S. maltophilia* strains. This could provide mechanisms for the ineffectiveness of the DAF-2/16 pathway in response to *S. maltophilia* JCMS.

Genomic differences: the answer to phenotypic differences?

The other goal of this thesis was to determine genomic differences in *S. maltophilia* strains that could account for differences in phenotypic virulence. Although phenotypic differences have been observed between our strains and the same strains in other labs, strains that we used for genome assembly and their respective references were overall very similar. This means that genomic analyses performed using these reference strains should also be relevant to the strains used for our analyses.

In addition, very few genes classified as being involved in "Virulence, Disease, and Defense" differed between strains. Functional analyses of several identified differences, such as the CmeABC efflux pump component, Type IV secretion system, and metal resistance genes, would determine if they are important for virulence. For example, knock-down of these genes in the strains in which they are found paired with exogenous expression in strains lacking these genes followed by survival assays in *C. elegans* would reveal their importance in pathogenesis.

However, identifying genomic differences between strains was dependent on presence and absence of coding sequences. Therefore, another genomic explanation for differences in phenotypic virulence could be single nucleotide polymorphisms within regulatory or coding regions. Therefore, looking for smaller genetic variations within virulence factors and regulatory regions could identify important genomic differences.

Further analysis of these strains also revealed that many genes within *S. maltophilia* genomes are strain-specific or only shared between a small fraction of genomes. Most of the genes identified in the shell and cloud genomes, and therefore more unique to individual strains, encode "hypothetical proteins." Therefore, it is possible that many virulence genes have not yet been characterized. Determining phenotypic virulence of more strains would allow us to narrow down genes important for virulence. For example, determining phenotypic virulence of all strains used for the pangenome analysis and comparing genes between pathogenic and nonpathogenic strains would result in candidate genes that could be used in functional analyses. In addition, transcriptomic approaches could be used to determine differential expression of genes upon infection. This experiment could involve several comparisons, in which gene expression of bacteria under normal growth conditions could be compared to bacteria upon infection in *C. elegans*, or gene expression of pathogenic and nonpathogenic strains could be compared. Both approaches would identify candidate genes involved in virulence, some of which may be novel virulence factors.

Overall, it appears that the *S. maltophilia-C. elegans* interaction is complex and the genetic basis of this interaction is strain-dependent. However, understanding these dynamics will begin to unravel the complex pathogen-host interactions in the native environment of *C. elegans* in addition to providing insight into clinical aspects of *S. maltophilia* pathogenesis.

References

Dirksen, P., Marsh, S. A., Braker, I., Heitland, N., Wagner, S., Nakad, R., Mader, S., Petersen, C., Kowallik, Rosenstiel, P., Felix, M.-A., & Schulenburg, H. (2016). The native microbiome of the nematode *Caenorhabditis elegans*: gateway to a new host-microbiome model. *BMC Biology, 14*(38). https://doi.org/10.1186/s12915-016-0258-1

Kenyon, C., Chang, J., Gensch, E., Rudner, A., & Tabtiang, R. (1993). A *C. elegans* mutant that lives twice as long as wild type. *Nature, 366,* 461–464.

McElwee, J., Bubb, K., & Thomas, J. H. (2003). Transcriptional outputs of the *Caenorhabditis elegans* forkhead protein DAF-16. *Aging Cell, 2*(2), 111–121. https://doi.org/10.1046/j.1474-9728.2003.00043.x

Murphy, C. T., McCarroll, S. A., Bargmann, C. I., Fraser, A., Kamath, R. S., Ahringer, J., Li, H., & Kenyon, C. (2003). Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans. Nature*, *424*(6946), 277–283. https://doi.org/10.1038/nature01789

Samuel, B. S., Rowedder, H., Braendle, C., Félix, M.-A., & Ruvkun, G. (2016). *Caenorhabditis elegans* responses to bacteria from its natural habitats. *Proceedings of the National Academy of Sciences, 113*(27), E3941–E3949. https://doi.org/10.1073/pnas.1607183113

White, C. V, Darby, B. J., Bredden, R. J., & Herman, M. A. (2016). A *Stenotrophomonas maltophilia* strain evades a major *Caenorhabditis elegans* defense pathway. *Infection and Immunity, 84*(2), 524–536.