Chapter I

Microbiology, Genomics and Clinical Significance of the *Pseudomonas fluorescens* Species Complex, an Unappreciated Colonizer of Humans.

Modified from:

Microbiology, genomics, and clinical significance of the *Pseudomonas fluorescens* species complex, an unappreciated colonizer of humans. Scales BS., *et al*. Clinical Microbiology Reviews. 2014 Oct;27(4):927-48.

INTRODUCTION

Over the past 15 years, the application of culture-independent methods for microbial identification has revealed a previously unappreciated complexity within human-microbe interactions. One interesting feature is that a number of these studies have identified the bacteria *Pseudomonas fluorescens* as a low abundance member of the indigenous microbiota of various body sites, including the mouth, stomach and lungs (1-5). *P. fluorescens* has generally been considered non-pathogenic for humans, an assessment dating back to its earliest descriptions:

"*The bacillus [P. fluorescens] itself is not pathogenic. A culture applied to animals subcutaneously or injected into the peritoneum does not elicit a reaction. Even when introduced many times into fresh wounds it does not irritate healing by primary intention. Also, ingestion of cultures caused no harm to my stomach or intestines*."

- A. Baader, "*Über Antagonisten unter den Bacterien*" (1887) (6).

However, while far less virulent than *P. aeruginosa, P. fluorescens* can cause acute infections in humans and has been reported in samples from the mouth, stomach and lungs (**Figure I.1**). The most common site of *P. fluorescens* infection is the bloodstream. Most reported cases have been iatrogenic, with bacteremia attributable either to transfusion of contaminated blood products (7-12) or contaminated equipment associated with intravenous infusions (13-17). While not suspected of being an etiologic agent of pulmonary disease, we have recently reported that *P. fluorescens* is routinely cultured at a low frequency from clinically-indicated respiratory samples (3) (**Figure I.2**). Perhaps the most intriguing association between *P. fluorescens* and human disease is that approximately 50% of Crohn's Disease patients develop serum antibodies to the I2 antigen encoded by *P. fluorescens* and, in some studies, this seroreactivty has correlated with success of therapies aimed at the microbiome, rather than the immune system (18-22). Altogether, these reports and others are beginning to highlight a far more common, and potentially complex, interaction between humans and *P. fluorescens* during health and disease.

The extremely versatile metabolic capabilities of *Pseudomonas fluorescens* impart these bacteria with the ability to persist in a wide range of environments beyond mammalian hosts (**Figure I.3**), including soil, the rhizosphere and surfaces of plants, non-sterile pharmaceuticals,

showerheads, and even indoor wall surfaces (23, 24). *P. fluorescens* has been most widely studied as an environmental microbe, most notably for its role in promoting plant health via a number of encoded antimicrobial mechanisms (25-38). However, *P. fluorescens* also possesses a number of functional traits that provide it with the capability to grow and thrive in mammalian hosts, including production of bioactive secondary metabolites (26-30, 33, 39-42) and siderophores (43-45), a type-III secretion system (46-51), the ability to form biofilms (20, 52-56), and the plasticity of some strains to adapt to growth at higher temperature (53, 57-59).

With recent rapid advancements in taxonomy and comparative genomics, many *Pseudomonas* isolates originally identified as the "species" *fluorescens* are now being re-classified as novel *Pseudomonas* species within the *P. fluorescens* "species complex" (23, 60, 61). There are at least 52 species within this group (**Figure I.4**) and they share many phenotypic characteristics (**Figure I.5**). Since the taxonomic reclassifications within *P. fluorescens* are relatively new and ongoing, and beyond the scope of this review, we will use the term "*P. fluorescens* species complex" or simply "*P. fluorescens"* in this review for studies on any isolates within this *Pseudomonas* species complex (**Figure I.4**).

PHENOTYPIC TRAITS AND CULTIVATION OF *P. FLUORESCENS*

The bacteria in the *P. fluorescens* species complex are Gram negative, motile rods that are primarily aerobic, unable to ferment glucose, chemoorganotrophic, and grow at a pH between 4-8 (62) (**Figure I.5**). Isolates of *P. fluorescens* derived from non-mammalian samples have a permissive growth range between 4-32°C (62) while isolates from humans and other mammals have an elevated upper range extending to 37°C (53, 57-59). As of the end of 2013, there were 16 fully sequenced strains from the *P. fluorescens* species complex and all but one

originated from plant surfaces, roots or the surrounding soil (**Figure I.6**). *P. fluorescens* can also be found in an antagonistic relationship with eukaryotic microbes, including oomyceytes and amoeba (35, 36, 48, 51, 59, 63-65), the latter relationship potentially reflecting conserved mechanisms that are also used with macrophages, as has been hypothesized for other bacteria (66).

Like most members of the Pseudomonas genus, *P*. *fluorescens* species complex strains grow best in a rich, peptide containing media with 0.1-1.0% (w/v) energy source (62). Examples of such basic media include nutrient broth/agar and trypic soy broth/agar (62). Selective media that is deficient in iron allows for the detection of the natural fluorescence produced by these bacteria, which is enhanced due to increased production of fluorescent siderophores. King's A and B media (67); Pseudosel agar media (BBL Microbiology Systems) and *Pseudomonas* agar F medium (Difco Laboratories, Detroit, MI) are all examples of pigment-enhancing media. These media also contain additional compounds, such as potassium, magnesium and/or Cetrimide, that further enable selection growth of *P. fluorescens* species complex bacteria. Cetrimide in particular helps inhibit the growth of non-*Pseudomonas* microbial flora and allows for adequate pigment production from *P. aeruginosa* (68). One of the difficulties in isolation of particular species in the *Pseudomonas* genus is that they share of many of the same phenotypic traits and grow under the same cultivation conditions. However, it is possible to use pigment production, which varies by species group, to visibly distinguish isolates from different groups. The blue-green pigment pyocyanin readily produced by *P. aeruginosa* strain is typically not produced by strains of the *P. fluorescens* species complex (62). Therefore, a mix culture of *P. fluorescens* species complex bacteria and *P. aeruginosa* bacteria grown on Cetrimide agar will produce blue, fluorescent colonies of *P. aeruginosa* and non-blue, fluorescent colonies of *P. fluorescens* complex bacteria.

Environmental isolates of *P. fluorescens* are readily cultivated in the laboratory using standard culturing techniques at a lower temperature range (5°C-32°C) but, in samples from higher temperatures or in clinical material, cultivation of *P. fluorescens* may be more difficult. *P. fluorescens* can be cultivated from environmental samples using a simple medium with a carbon source and aerobic incubation for 24-48 hours at 27°C-32°C (62). However, incubation of environmental samples at temperatures between 35°C-37°C can cause *P. fluorescens* to enter a viable-but-not-culturable (VBNC) state (69), complicating cultivation. During the VBNC state, the bacteria are still metabolically active but unable to undergo cellular division and replication (70). Bacteria in a VBNC state often will not grow when immediately transferred to standard culture conditions. *Escherichia coli, Listeria monocytogenes, Salmonella enteritidis* and *Shigella dysenteriae* are all examples of bacteria that can enter into a VBNC state (71). Vibrio species also undergo a switch to a VBNC state that, similar to *P. fluorescens*, is prompted by a switch in temperature (72). The VBNC state is hypothesized to be a survival strategy that allows bacteria to persist in harsh environments (73). The ability of *P. fluorescens* to become VBNC could explain the phenomenon in which *P. fluorescens* can be found more frequently in human lung metagenomic DNA than is reported by standard hospital culture methods (3). However, some isolates of *P. fluorescens* from human samples have adapted well to a higher permissive temperature range than those from environmental samples. For example, we have a collection of over 30 *P. fluorescens* strains from cystic fibrosis patients that grow well at 37°C. Another study reported a series of *P. fluorescens* isolates from surface abscess, septicemia and respiratory or urinary tract infections that were able to grow at 37°C (57). All seven were also able to grow at 4°C, often considered the lower limit of the optimal temperature range of *Pseudomonas spp.* (62), suggesting that these strains did not shift their temperature range, rather the range expanded upwards.

GENOMICS

Taxonomy and genomics of the *Pseudomonas* genus and P. *fluorescens* species complex

Of the many species within the *Pseudomonas* genus, the *P. fluorescens* species complex contains ~20% of all *Pseudomonas* species (74). As of January 2014, the List of Prokaryotic Names with Standing in Nomenclature (LPSN) recognized 211 species and 18 subspecies in the *Pseudomonas* genus (www.bacterion.net/*Pseudomonas*.html). This reflects a 40% increase in newly defined *Pseudomonas* species compare to 2006 (74). In the last few decades, isolates classified as *P. fluorescens* have undergone extensive renaming and reorganization, consistent with the high degree of genomic diversity within this species complex (75). Historically, any bacteria that was a Gram negative, strictly aerobic, non-sporulating, motile bacillus was classified as belong to within the *Pseudomonas* genus (76). The name *Pseudomonas* derives from the Greek word for 'false' (pseudes) and 'single unit' (monas), so it is ironic that taxonomy within the *Pseudomonas* genus is undergoing reorganization in the genomics era (77). Molecular methods including analysis of 16S rRNA gene sequences, other highly conserved "housekeeping" genes and, more recently, full length genomes have accelerated the pace of taxonomic reorganization, especially within the *P. fluorescens* species complex (76, 78-80). Multi-locus sequence typing (MLST) and multi-locus sequence analysis (MLSA) utilize the concept of genetic evolution at multiple conserved genes to measure evolutionary distances between species or strains (81-83). These conserved ("housekeeping") genes are required for the basic functions of the cell and must be found in all bacteria in the comparison (84-86). Examples of housekeeping genes used in classifying *Pseudomonas* species include *rpoD* (sigma subunit of RNA polymerase), *rpoB* (β-subunit of RNA polymerase), and *gyrB* (β-subunit of gyrase, responsible for negative supercoiling of DNA during replication)

(80, 87, 88). The combination of these three housekeeping genes and the 16S rRNA gene was used to identify members of the *Pseudomonas* genus, create a phylogenetic tree and divide them into different groups (74, 76) (**Figure I.4**). One of the key findings of these analyses is that the widest range of genomic diversity in the *Pseudomonas* genus is found in the *P. fluorescens* species complex (74) (**Figure I.4**).

The *P. fluorescens* species complex includes at least 52 separately named species, including *P. poae*, *P. synxatha*, *P. tolaasii*, *P. brassicacearum*, *P. chlororaphis* and *P. fluorescens* (23, 60, 61). The *P. fluorescens* species complex can also be divided into three smaller taxonomic clades, based on 16S rRNA gene and MLSA analyses (23, 49, 74, 89, 90) (**Figure I.7**). Comparative genomics of full- length genomes confirms the high genetic diversity of this group of bacteria. Between the *P. fluorescens* complex strains in **Figures I.7** and **I.8**, there are 2789 shared protein coding genes, representing only 45%- 52% of the proteome from any one strain (49). This is in contrast to *P. aeruginosa*, where a comparison of five sequenced genomes found that they shared a total of 5021 protein-encoding genes and that no more than 8.2% was unique to any one strain's proteome (23). Within the *P. fluorescens* species complex, there is a "core genome" of 2789 genes; only 20 are unique to the *P. fluorescens* species complex itself and encode proteins involved in regulation, biofilm formation or of hypothetical functions (49). Another way of looking at the level of genetic diversity among *P. fluorescens* isolates is to look at the pan-genome, which is the total number of genes found across all strains. The pan-genome of *P. fluorescens* bacteria is 13,782 protein-coding genes. In comparison, the pan-genome of *P. aeruginosa* is only 7,824 CDs. Within each clade, the level of genetic similarity between strains is higher, with 4188, 3729 and 3893 shared conserved domains between members of clade 1, 2 and 3 respectively (49). The clade designation also offers some potential insights into functional differences between clusters of *P. fluorescens*, including the presence/absence and type of Type-III Secretion System (T3SS), a molecular "needle" complex

utilized by bacteria to inject bacterial proteins in host cells (49, 91). Genes for a T3SS is only found in clades 2 and 3, not clade 1. The biosynthesis gene cluster to produce hydrogen cyanide, a volatile molecule used to kill off competing bacteria, is only found in clades 1 and 2 (29, 33, 49). Genes found in every clade, such as for the siderophore pyoverdine (92, 93), reflect functional categories that are generally preserved across the *Pseudomonas* genus as a whole (49). Comparative genomics of *P. fluorescens* is only beginning, but since there are marked differences in the presence/absence of numerous genes between strains, this approach holds significant promise as a step in organizing the *P. fluorescens* species complex according to putative functional differences.

Identifying *P. fluorescens* in samples by high-throughput sequencing

The coupling of high-throughput sequencing with the generation of 16S rRNA gene amplicon libraries from metagenomic samples has fueled the explosion in information about the microbiome and environmental microbial communities. Databases for the subsequent bioinformatic analysis have continued to expand at a staggering pace. Historically, taxonomic assignment of a short read sequence from this type of analysis was limited to the family or genus level. However, as additional fully sequenced genomes become available to build validated phylogenetic trees of short read sequences, some genera can be resolved at a species level. This is turning out to be the case for some of the species in the *Pseudomonas* genus; the groups identified by MLST and MLSA can also be identified using the V3-V5 region of the 16S rRNA gene. As illustrated in **Figure I.9**, a phylogenetic tree can be generated using the16S rRNA gene sequences corresponding to the V3-V5 regions of the gene and a progressive tree alignment strategy (94-97). The bootstrap values for separating *P. aeruginosa* from the other *Pseudomonas* genera are very high. While the bootstrap values are much lower for distinguishing the non-aeruginosa *Pseudomonas* species based on the V3-V5 region alone,

the, short-read high-throughput sequencing technologies that target the V3-V5 variable region of the 16S rRNA gene can offer a first pass analysis that discriminates between members of the *P. putida* and *P. fluorescens*/*P. syringae* clusters.

We have used this type of analysis, combined with other data, to demonstrate that both *P. aeruginosa* and *P. fluorescens* are prominent members of the respiratory microbiota of lung transplant recipients but increases in their relative proportions are associated with widely divergent clinical associations (3). Multiple independent studies had identified the presence of *P. aeruginosa* in respiratory cultures as a positive risk factor for the subsequent development of bronchiolitis obliterans syndrome (BOS) (98-100). However, in the largest published study to date of lung transplant subjects utilizing high-throughput sequencing for microbial identification, a negative association was reported between the presence of *Pseudomonas* species and the diagnosis of BOS (101). In our study (3), we similarly observed high levels of *Pseudomonas* in lung transplant recipients (as determined by high-throughput sequencing of V3-V5 16S rRNA gene amplicon libraries). However, after applying the analysis described above and adding our *Pseudomonas* operational taxonomic units to the phylogenetic tree to delineate *P. aeruginosa* vs. *P. fluorescens*, subjects with abundant *P. aeruginosa* had other clinical symptoms consistent with an acute infection, including positive *P. aeruginosa* bacterial cultures. In contrast, for the numerous subjects with abundant *P. fluorescens*, they exhibited little evidence of acute infection and no *Pseudomonas* species was detected via standard clinical laboratory bacterial culture. We alluded earlier in this review to the gap in knowledge about the factors that control culturability of *P. fluorescens* from clinical samples, which was underscored in our study. The surprising stark difference in culture positivity between these pseudomonads may explain the difference between prior culture-based studies (98-100) and the culture-independent study (101). Of note is the fact that healthy controls in our study had very little signal for either *P. aeruginosa* or *P. fluorescens* in their bronchoalveolar lavage fluid (3). We provide this as an

example of the potential power of high-throughput sequencing to provide new insights into the association of *P. fluorescens* with humans during health and disease.

FACTORS AFFECTING HOST COLONIZATION AND PERSISTENCE

Antibiotics and secondary metabolites

P. fluorescens produces a long list of secondary metabolites that allow it to successfully vie with competing microorganisms. Examples include phenazine (26-28), hydrogen cyanide (HCN) (29), 2,4-diacetylphloroglucinol (DAPG) (30, 31), rhizoxin (32-34) and pyoluteorin (35, 36). Phenazines can be produced by Gram negative bacteria found in soil and marine environments, with *Pseudomonas spp.* being one of the major producers (28). Phenazines are pigmented compounds that have antitumor, antimalarial, antiparasitic and antimicrobial activity (26). *P. fluorescens* produces the yellow phenazine, phenazine-1-carboxylic acid (PCA) (28)**.** Hydrogen cyanide is a volatile, colorless compound that inhibits cytochrome c oxidases and other metalloproteins in competing bacteria (33). The production of HCN by rhizosphere-inhabiting *P. fluorescens* suppresses plant disease (29). While it has not been studied for *P. fluorescens,* other *Pseudomonas spp.* are capable of producing HCN during human disease, such as cystic fibrosis (102). The anaerobic regulator protein ANR regulates the *hcnABC* gene cluster that encodes hydrogen cyanide synthase and, due to the oxygen sensitivity of the synthase, ensures the genes are only expressed under low-oxygen conditions (103). DAPG production plays a significant role in the plant-disease control activity of many *P. fluorescens* strains (30). Despite its importance, the DAPG biosynthetic cluster (*phl*) has been lost from all but a subset of *P. fluorescens* strains through evolution (31). Members of clades 1 and 2 of the *P. fluorescens* species complex (such as *P. protegens* Pf-5; *P. fluorescens* Q8r1-96 and *P. fluorescens* Q2-87) have retained the DAPG biosynthesis cluster while all members of clade 3

do not (49). Intragenomic recombination and rearrangement occurs frequently at this locus, such that DAPG-producing strains often have multiple versions of the *phl* gene cluster. In the phylogenetic lineage that retained DAPG synthesis, the gene cluster has maintained is structure even though it has been relocated multiple times in the various *P. fluorescens* genomes (31). Rhizoxins are 16-membered macrocyclic lactones that interfere with microtubulin dynamics during mitosis by binding β-tubulin (32) and show inhibitory activity against fungi, bacteria and tumors (33, 34). The rhizoxin-producing gene cluster in *P. fluorescens* is shared with another gammaproteobacteria, *Burkholderia* (104). Pyoluteorin was first isolated from a *P. aeruginosa* strain (105) but is now known to be produced by multiple *Pseudomonas spp.*, including *P. fluorescens* (35). It has been studied in *P. fluorescens* strains Pf-5 and CHA0 for its anti-bacterial activity and ability to improve plant health (35, 36). While the activities of these secondary metabolites on human hosts remain to be determined, they benefit the survival of *P. fluorescens* in polymicrobial environments, opening the possibility of a role for these metabolites in survival of *P. fluorescens* in the human microbiome.

Other secondary metabolites produced by *P. fluorescens*, notably pyrrolnitrin and the pseudomonic acids, have been formulated for medical and agricultural uses. Pyrrolnitrin, a chlorinated molecule with antifungal activity, was developed into both a topical antimycotic for clinical use (39) and a fungicide for agricultural use (106). Pseudomonic acids are perhaps the most clinically important antibacterials produced by *P. fluorescens*. There are multiple pseudomonic acids (107) and each exhibits some level of antibacterial activity. Pseudomonic acid A has the highest activity and is the major pseudomonic acid (90%) in mupirocin, a topical antibiotic (40, 41, 108, 109). Topical mupirocin (2% concentration) is effective for treatment of superficial skin infections, such as impetigo, caused by the Gram positive bacteria *Staphylococcus spp.* and *Streptococcus spp.* and the Gram negative bacteria *Hemophilus influenzae* and *Neisseria gonorrhoeae* (41, 109-111). Pseudomonic acid A interacts with the

amino acid binding site of isoleucyl-tRNA synthase and the respective ATP binding site, inhibiting the ability of bacteria to produce isoleucyl-tRNA synthetase (112-114). This inhibits protein synthesis primarily and, to a less extent, RNA and bacterial cell wall synthesis, possibly due to auxotrophy of amino acids that are important for these processes. The end result is death of the bacterial cell. *P. fluorescens* is protected from pseudomonic acid because the *P. fluorescens* isoleucyl-tRNA target synthetase is structurally different and binds pseudomonic acid with much lower affinity (115). Overall, the production of pyrrolnitrin and the pseudomonic acids provide *P. fluorescens* with significant growth advantages in polymicrobial environments.

Siderophores and pigments

The secretion of a fluorescent pigment, pyoverdine (formally called fluorescein), is what imparts *P. fluorescens* with its fluorescent properties under UV light. Pyoverdine is a siderophore (116), a high-affinity iron chelating compound that is essential for acquisition of iron from the environment, bacterial growth and survival (43). Pyoverdine is the main siderophore of *P. fluorescens* (92), but some strains of *P. fluorescens* contain additional secondary siderophores for iron acquisition. *P. protegens* CHA0, a *P. fluorescens* species complex strain, produces the secondary siderophore enantio-pyochelin (44) and *P. fluorescens* strain ATCC 17400 produces the secondary siderophores quinolobactin, pseudomonine, acinetobactin and anguibactin (117). The latter three secondary siderophores from this strain are synthesized through a single pathway, with different primary substrates determining which final siderophore molecule is synthesized (45). Strains BG33R (*P. synxantha*) and A506 (*P. fluorescens*) also have the gene clusters necessary for the biosynthesis and uptake of a pseudomine-like molecule similar to that found in ATCC 17400, but the functionality of this molecule is not yet demonstrated (49, 118). The full complement of genes necessary for the biosynthesis and efflux of a hemophore, which allows for the chelation and transportion of heme through a specific

outer membrane (49), are also present in multiple *P. fluorescens* strains, but it is not known how and when hemophores are utilized by *P. fluorescens*. Much work remains on identifying the spectrum of siderophores produced by *P. fluorescens*, as well as determining their role in the physiology of these organisms under different conditions, including polymicrobial competition.

Like many other members within its genus, *P. fluorescens* produces a range of pigments, with and without siderophore ability. Due to the ability of *P. fluorescens* to grow at temperatures as low as 4°C, contamination of food products can be a problem (119). In 2010, European consumers noticed that some mozzarella products were blue versus white in color and extremely high levels of *P. fluorescens,* up to 10^6 CFU/g, were identified on the "blue" cheese samples (120). Beyond being blue, little is known about this particular pigment produced by *P. fluorescens*. Pyocyanin, another blue pigment, is produced by *P. aeruginosa*, but this secondary metabolite has not yet been identified in *P. fluorescens* (121). This incident either indicated the emergence of a new strain of *P. fluorescens* that had acquired the biosynthesis machinery for a new blue-colored pigment or horizontal acquisition of the biosynthesis machinery from another, closely related *Pseudomonas* strain.

Two-component gene systems

P. fluorescens also contains a two-component GacS/GacA system that plays a role in environmental sensing. This system controls the expression of multiple secondary metabolites and enzymes in *P. fluorescens*, including DAPG, pyoluteorin, HCN, phospholipase C and exoprotease (122-125). In *P. aeruginosa*, GacA controls gene expression through acylated homoserine lactone (AHL) signaling (126, 127). However, GacA can also function independently of AHL signaling (126) and this AHL-independent GacA cascade has been reported in *P. fluorescens* strain CHA0 (128). The diffusable non-AHL bacterial signal, whose chemical nature

is still under investigation, turns on and regulates a two-component GacS/GacA system that activates the transcription of a novel, small non-coding RNA, RsmY (128). RsmY then combines with a riboregulator (RsmA), which is a small, untranslated RNA that can regulate cellular processes (129-132), to positively regulate the expression of downstream genes at a post-transcriptional level (65).

Quorum sensing and biofilms

Bacteria are able to regulate their population density through the release and sensing of signal molecules, i. e. quorum sensing (133, 134). Quorum sensing regulates genes that control motility (swimming, swarming), antibiotic synthesis and biofilm formation. Genes involved in biofilm formation and quorum sensing are found in the core genome of the *P. fluorescens* species complex (49). Quorum sensing and biofilm formation are integral to the many environmental niches occupied by *P. fluorescens* and allow it to colonize surfaces such as hospital equipment and food-grade stainless steel surfaces (52, 135), as well as the surfaces of plants, showerheads and even indoor wall surfaces (23, 24, 136). *P. fluorescens* readily forms biofilms with highly complex, 3-D structures (**Figure I.10**) (20, 52-56) and strains that form plant-associated biofilms are often important biocontrol agents that protect plants against pathogenic fungi (54, 137). Less is known about *P. fluorescens* biofilm formation on mammalian surfaces, though the adaption to a 37°C permissive growth range is linked to biofilm formation on human cells (53). Thus, whether on plants or human cells, biofilm formation is likely important for the successful long-term colonization by *P. fluorescens*.

Two types of quorum sensing systems have been described in *P. fluorescens*: the AHL/lux and *hdtS* systems. In Gram negative bacteria, AHL molecules are produced by LuxI-like proteins and interact with LuxR-like proteins to form dual AHL-LuxR complexes. This AHL-

LuxR complex then binds *lux*-boxes of quorum-sensing regulated genes to either turn on/up or off/down their expression (138). A *luxI/luxR*-like system in *P. fluorescens* was first discovered in the strain NCIMB 10586 and was termed the *mpuI/mpuR* system due to its regulation of the antimicrobial mupirocin biosynthesis pathway (139). Another quorum sensing system was later discovered in *P. fluorescens* strain F113; the *hdtS* system (140). The hdtsS gene encodes a novel AHL synthase that produces separate signaling molecules: a N-(3-hydroxy-7-cis-tetradecenoyl) homoserine lactone, 3-OH-$C_{14:1}$-AHL, a *N*-decanoylhomoserine lactone C_{10}-AHL and a C_6-AHL. Though the signaling molecules and synthase have been elucidated, the genes regulated by the *hdtS* system are still unknown and no detectable phenotype in F113 has yet been linked to the signaling molecules (140).

The second messenger cyclic di-GMP is essential for regulation of steps involved in biofilm formation, including the production of LapA, an adhesive protein necessary for *P. fluorescens* attachment to surfaces (141). LapA is negatively regulated by the periplasmic protease LapG and positively regulated by the inner membrane protein LapD (142). LapG typically cleaves LapA from the bacterial surface, but when LapD is bound by c-di-GMP, LapD undergoes a conformation change that allows it to bind to LapG, inhibiting LapA cleavage. Diguanaylate cyclases catalyze c-di-GMP synthase activity and, in *P. fluorescens* Pf0-1, there are a total of 43 potential diguanaylate cyclases encoded in the genome, each potentially connected to a different aspect of biofilm formation (143).

Type-III secretion systems

Type III secretion systems (T3SSs) are molecular needle-like complexes that act like syringes to deliver bacteria proteins, called effectors, from the bacterial cytoplasm directly into host cells (91) (**Figure I.11**). T3SSs are highly conserved genomic clusters typically found in

bacteria that have close interactions with eukaryotic hosts (often transferred horizontally between phylogenetically unrelated bacteria) and the type of T3SS usually mirrors the type of interaction a bacteria has with the eukaryotes in its environment. The first T3SS was described in *Yersinia*, which delivers Yop (*Yersinia* outer protein) effector proteins into human host cells (144, 145). A total of five different T33S groups have since been described: the Ysc group (which includes the *Yersinia* Ysc, the *P. aeruginosa* Psc, the *Bordetella* Bsc, the *Rhizobium* Rsc and the *Chlamydia spp.* T3SSs); the Hrp1 group (found in *non-aeruginosa Pseudomonas spp.* and *Erwinia spp.*); the Hrp2 group (found in *Xanthomonas spp.* and *Ralstonia spp.*); the Inv/Mxi/Spa group (which includes the *Salmonella* SPI-I, the *Shigella spp.* and *Y.enterocolitica* Ysa T3SSs and the T3SS2 of EHEC); and the Esa/Ssa group (including *Salmonella* SPI-2, EPEC T3SS and EHEC T3SS1) (146). The most common T3SS found *P. fluorescens* strains is the Hrp1 system (46-50).

The Hrp1 family is the most common T3SS found among *P. fluorescens* strains (46-50). The Hrp system (hypersensitivity response and pathogenicity) triggers the hypersensitivity defense response in resistant plants, while leading to disease in susceptible plants, and was first described in *P. syringae* (147). Like the T3SS found in *Yersinia,* the Hrp1 system is involved in delivering bacterial proteins directly into host cells (148-151) (**Figure I.11**). While the fully sequenced *P. fluorescens* strains SBW25, BG33R, A506, SS101, Q8r1-96 and Q2-87 have at least one copy of the Hrp1 family T3SS, Pf0-1 and Pf-5 do not encode the gene cluster at all (49, 152). The activity and functionality of the Hrp1 system has been worked out in only a couple of the strains where it has been found. The Hrp1 T3SS of *P. fluorescens* Pf29Arp, a strain known for its ability to reduce the severity of wheat take-all, shows activity during the colonization of wheat rhizosphere (46). The homologous Hrp1 T3SS in strain SBW25 is induced during sugar beet rhizosphere colonization (153) and can induce a hypersensitive response in tobacco (47, 154). Interestingly, in addition to Hrp1-system effectors, SBW25 also contains the T3SS effector

ExoY (155), which in *P. aeruginosa* targets the actin cytoskeleton of eukaryotic cells (156). Since most of the work on the functionality of the Hrp1 T3SS in *P. fluorescens* has been done *in vitro*, many of the target host cells are still unknown, but the presence of the ExoY effector protein in some strains suggests that there might be an additional, non-plant use of this T3SS in SBW25 and genetically related strains.

Additional evidence that *P. fluorescens* may target their T3SS against eukaryotic cells was provided in 2013, when a SPI-I-like T3SS gene cluster was discovered in strain F113 (51) (**Figure I.11**). The F113 strain was originally isolated from sugar-beet rhizosphere in Ireland (157) and can inhibit the growth of plant-pathogenic bacteria, oomycotes, fungi and wide-range of nematodes (158-160). Predation against protozoa in both terrestrial and aquatic environments is an important factor influencing bacterial community makeup and behavior (37, 38, 161). In F113, the SPI-I T3SS promoter *hilA* shows increased expression during close contact with the amoeba, *Acanthamoeba castellani*, suggesting that this T3SS is directly involved in protecting the bacteria from amoeba predation. Interestingly, both the Hrp1 and SPI-I systems in F113 appear to be involved in protection against predation by the worm *Caenorhabditis elegans* (51). A similar result was found with the Hrp1 system of *P. fluorescens* CHA0 (162). Additional SPI-I T3SSs have also been found in *P. fluorescens* strains HK44 (163) and Q2-87 (49), providing further evidence of T3SS action outside of the plant ecosphere. Thus, the identification of multiple T3SS across the *P. fluorescens* species complex that target plant and non-plant eukaryotic cells supports the model of a wider interaction of *P. fluorescens* with eukaryotic hosts.

INTERACTION OF *P. FLUORESCENS* WITH HUMAN CELLS

Environmental isolates of *P. fluorescens* have an optimal temperature growth range between 25 - 30 °C and are not virulent to human cells, but certain strains *P. fluorescens* isolated from clinical samples have a higher permissive growth range, up to 37°C, and show increased virulence against human cells (53, 57-59). Two *P. fluorescens* strains, MFY162 and MFN1032, can adhere to human glial cells in culture, and MFN1032 can induce apoptosis. Originally isolated from an individual with a lung infection (57), MFN1032 not only exhibits cytotoxicity on human intestinal epithelial cells *in vitro* but also triggers a proinflammatory response (164). Human airway epithelial cells exposed to a different strain of *P. fluorescens* have been shown to trigger both antiapoptotic responses via the epidermal growth factor receptor (EGFR) and interleukin-8 production via TLR4-independent NF-κB signaling pathways (165). Exposure to a strain of *P. fluorescens* isolated from a moldy building decreased viability of mouse macrophages (RAW cells) while inducing production of nitric oxide, TNF and IL-6 (166).

On red blood cells, *P. fluorescens* MFN1032 displays both cell-associated and secretion-dependent hemolytic activity. The secretion-dependent pathway is positively regulated by the GacS/GacA two-component system (58), the same two-component system that regulates phase variation in this strain (167). This hemolytic activity involves the production of phospholipase C and biosurfactants, similar to that seen for pathogenic *P. aeruginosa* (168). Similarities between *P. aeruginosa* and *P. fluorescens* also exist with the cell-associated hemolytic activities of MFN1032. The cell-associated hemolytic activity is independent of the secretion-association hemolytic activity, is active at 37°C, occurs without the secretion of phospholipase C and biosurfactants and does not depend on the GacS/GacA two-component system (169). In *P. aeruginosa*, cell-associated hemolytic activity occurs alongside type-III secretion of the PcrV, PopB and PopD effectors (170). MFN1032 also harbors the genes necessary to produce a

T3SS (169), the *hrcRST* gene cluster, which shares a high level of homology to the *hrcRST* genes of the hrpU operon in *P. syringae* DC3000. When this operon is mutated, MFN1032 no longer produces cell-associated hemolytic activity (169). In *P. aeruginosa*, similar mutations in its T3SS also abolish its cell-associated hemolytic activity. Thus, adaptation of *P. fluorescens* MFN1032 results in an increased temperature permissivity along with hemolytic activity against human cells that are similar to that found in *P. aeruginosa.*

The production of cyclolipopeptides (CLPs) by *P. fluorescens* MFN1032 is another functional characteristic that is altered during a shift to higher temperatures. Cyclolipopeptides are the most widely studied biosurfactants produced by *P. fluorescens* and are involved in swarming motility, biofilm formation and colonization of host surfaces (171). If MFN1032 is grown for multiple generations at 37°C, CLP functionality is lost, with ~ $4x10^{-3}$ CLP-deficient mutants found per generation (58). High mutation rates, inversions of DNA segments, DNA methylation and epigenetic switches are all mechanisms that bacteria use to alter their genome in the process of adaptation, which allows survival in changing environments and an increase in overall fitness with time (172). In the case of *P. aeruginosa* T3SS, there is an epigenetic switch between a non-inducible and an inducible state (167). Using a Boolean modeling system, a similar epigenetic switch has been shown to be the likely mechanism by which *P. fluorescens* regulates its CLP production (167). In much the same way that chronic *P. aeruginosa* strains lose their ability to produce biofilms after long-term growth in a cystic fibrosis lung (173), *P. fluorescens* also has a mechanism to turn off energy expensive surfactant production after long-term growth at physiologically-relevant temperatures.

CLINICAL SIGNIFICANCE

P. fluorescens as a disease-causing agent

By far, the bloodstream is the most common site reported for *P. fluorescens* infection in humans. Most reported cases have been iatrogenic, with bacteremia attributable either to transfusion of contaminated blood products (7-12) or contaminated equipment associated with intravenous infusions (13-17). *P. fluorescens* bacteremia has occurred in outbreaks (8, 13-16); the largest affecting at least 80 patients in 6 states after indirect exposure to contaminated heparinized saline flushes prepared at a common compounding pharmacy (16). Of these patients, 41% were bacteremic more than 84 days after exposure; all of these delayed-onset patients had indwelling ports for venous access, indicating that *P. fluorescens* can persist endovascularly when an indwelling catheter is in place. The ability to grow at refrigerated temperatures and form biofilms on fomite surfaces makes *P. fluorescens* contamination a particular problem for blood-infusion related infections and outbreaks.

Confounding the diagnosis of *P. fluorescens* bacteremia is the well-described phenomenon of “pseudobacteremia" due to environmental contamination of blood culture collection bottles and equipment by the organism (174-181). Indeed, in a systematic review of the medical literature, more positive *P. fluorescens* blood culture results were attributable to pseudobacteremia (174-181) than to true bacteremia (8-17, 182-188). Sources have included blood culture bottles cleaned with contaminated disinfectant (178) and, most commonly, contaminated blood collection tubes used prior to culture bottle inoculation (175, 177, 179-181). Despite not reflecting “true” human pathology, pseudobacteremia is a legitimate clinical problem, resulting in diagnostic confusion for clinicians and inappropriate antibiotic exposure for patients (180). The diagnosis of pseudobacteremia should be considered when patient symptoms are discordant with disseminated bacterial infection and bacteria that are uncommon infectious agents (such as *P. fluorescens*) are isolated, especially in a geographic or temporal cluster.

Identification of *P. fluorescens* as an acute cause of infection in sites other than the blood has been rare and sporadic (Table 1). Two reports have identified *P. fluorescens* in skin wounds and abscesses following dog bites (189, 190) and, in one instance, the patient subsequently developed disseminated *P. fluorescens* bacteremia (190). *P. fluorescens* has been implicated as a cause of acute bacterial cystitis (191-193), both with (191) and without (192) the presence of an indwelling urinary catheter. In a study comparing the oral microbiome of 20 solid organ transplant recipients and 19 non-immunosuppressed control subjects, *P. fluorescens* was abundant in the saliva of nearly 50% of transplant subjects while nearly absent from non-transplant controls (1). In another study of 258 stomach wall biopsies acquired from patients with various upper gastrointestinal disorders, 93% had evidence of *P. fluorescens* presence (identified both via culture-dependent and -independent methods) (2). Thus, *P. fluorescens* can clearly establish itself in diseased humans, but questions remain about the pathogenicity of such interactions and whether the strains are all restricted to a specific clade.

P. fluorescens in respiratory diseases

While *P. fluorescens* has been repeatedly cultured from respiratory specimens, its role in pneumonia or other respiratory infections is unclear. *P. fluorescens* has been cultured from the tracheal aspirates of patients receiving mechanical ventilation and subsequently identified as an organism in the humidifier water used in the ventilator circuit (194), however, it is unclear if the tracheal aspirate culture results reflected acute infection or benign colonization. In another case study, during recovery from a recent polymicrobial peritonitis, a patient developed clinical evidence of pneumonia with sputum cultures positive for *P. fluorescens* (195). The patient improved after treatment with a third-generation cephalosporin and subsequent sputum cultures did not grow *P. fluorescens*. In another report, *P. fluorescens* is mentioned in the etiology of community-acquired pneumonia in a single patient, but clinical details are lacking (196). Using

amplification of bacterial 16S rRNA genes, another study detected *P. fluorescens* and other bacteria in the bronchoalveolar lavage fluid acquired from a single patient with clinically-diagnosed ventilator-associated pneumonia (197). Most notably, in a survey of over 1000 respiratory cultures acquired from subjects with cystic fibrosis, Klinger et al. identified the organism in roughly 2% of specimens (198) and considered the organism a colonizer rather than an acute pathogen. We have reported, using bronchoalveolar lavage fluid acquired from lung transplant recipients, that *P. fluorescens* is frequently identified in this patient population in the absence of evidence of acute infection (3).

In a survey of bacterial culture isolates at the University of Michigan Hospital, *P. fluorescens* was cultured from respiratory specimens with relative frequency (3) (Table 2). Over a eleven-year period, *P. fluorescens* was cultured from over 240 distinct respiratory specimens, roughly 2 specimens per month. Among patients with positive *P. fluorescens* respiratory cultures, the most common underlying pulmonary condition was cystic fibrosis (38.8% of all isolates), followed by other chronic airway diseases (COPD, asthma and non-CF bronchiectasis: 16.1%). *P. fluorescens* was often co-isolated with other organisms, most often (85.1%) species designated as “oral flora” by the clinical microbiology laboratory, followed by *P. aeruginosa* (25.6%), *Staphylococcus aureus* (15.7%) and *Stenotrophomonas maltophilia* (11.6%). In no cases was *P. fluorescens* the unambiguous causative agent in a monomicrobial pneumonia. This survey highlights the fact that *P. fluorescens* is commonly isolated from human clinical samples in cases where it is not the cause of active acute infection. This contrasts with much of the literature, which states that *P. fluorescens* is only found in human hosts in extreme cases of outbreak or contamination. In addition, no reports were created in response to these *P. fluorescens* cultures, revealing that the number of reports in the literature also likely do not reflect the consistency by which *P. fluorescens* is cultured from clinical samples.

P. fluorescens and inflammatory bowel disease

P. fluorescens has also been speculated to have a possible role in the pathogenesis of Crohn's Disease and other inflammatory conditions. I2, a peptide encoded by *P. fluorescens*, was found to be more frequently detected in gut wall biopsies of patients with Crohn's Disease than those of patients with other bowel diseases and a similar difference was noted in detection of circulating anti-I2 antibodies (199). Interestingly, there was no evidence of *P. fluorescens* in the stool of subjects with Crohn's Disease, either by culture or microbe-specific PCR. The same I2 sequence was also found in the proximal colon, cecum and distal small intestine of C57BL/6J mice, suggesting that *P. fluorescens* could exist in the intestinal microbiota of multiple mammalian species (200). In TLR4- and MyD88-knockout mice that were treated with dextran sodium sulfate, the resulting colitis and impaired immune response led to systemically detectable *P. fluorescens*, such that it could be cultured from their mesenteric lymph nodes (201). Approximately 50% of Crohn's Disease patients develop serum anti-I2 antibodies and, in some studies, this seroreactivity has correlated with success of therapies aimed at the microbiome, rather than the immune system (18-22). I2 is encoded within the *P. fluorescens* *pfiT* gene and has T-cell superantigen activity (202). The presence of anti-I2 serum antibodies in Crohn's Disease has subsequently been shown to be positively associated with prognosis (19). Anti-I2 antibodies have also been associated with the diagnosis of celiac disease, including a decrease in titer after a gluten-free diet is initiated (203, 204), ankylosing spondylitis (205) and chronic granulomatous disease (206). Whether *P. fluorescens* is directly contributing to these chronic inflammatory conditions, or whether anti-I2 antibodies are only indirect biomarkers of disease, is undetermined.

FUTURE PERSPECTIVES

Despite being identified in the last half of the 1800's and more recent associations with human disease, the role of the *P. fluorescens* species complex in human health and disease remains largely unexplored. The research in the last two decades on the genetic, molecular, environmental, and immunological aspects of *P. fluorescens* has begun to expand our understanding of these bacteria overall and lay the groundwork for investigating their role in human health. Full genome sequencing and comparison have led to the discovery of potential pathogenic traits (such as T3SS and T cell super antigens) and further revealed the high level of genetic diversity within the *P. fluorescens* species complex. The discovery of human-adapted *P. fluorescens* with higher permissive temperature ranges has revealed that these bacteria can readily exist outside of plant and soil niches and even potentially change their functional phenotypes in response to a new mammalian-based niche. Clinical surveys have also found that *P. fluorescens* is regularly cultured from clinical samples even in the absence of acute infection or outbreak. Studies are beginning to identify *P. fluorescens* via high throughput sequencing in multiple sites of the human body, suggesting that the human-*P. fluorescens* connection will only grow as more studies are reported.

However, there is still much more that is unknown about the role of the *P. fluorescens* species complex in human disease. Taxonomic classifications within the *P. fluorescens* species complex are still in flux; a general consensus on what constitutes a '*P. fluorescens*' would codify classification and greatly assist functional microbiology research, as well as the clinical microbiology lab and clinician. Almost nothing is known about the host response to *P. fluorescens,* and while correlations have been found between *P. fluorescens*-specific antibodies and Crohn's Disease, the mechanisms underlying this connection have not been identified. Finally, there is a glaring disparity between reports in the medical literature that only find *P. fluorescens* infections during outbreaks/extreme situations and clinical surveys that readily identify *P. fluorescens* in human samples in the absence of acute disease. The former suggest

that *P. fluorescens* are accidentally associated with human hosts through contamination or when the host is immunocompromised; the latter suggest that there are strains of *P. fluorescens* that can colonize and thrive in a human host. Additional work on the genomics, molecular microbiology and host immune response to *P. fluorescens* will provide insight into the role these bacteria play in human health and disease.

RATIONALE FOR THESIS PROJECT

The over-arching goal of this dissertation is to use comparative genomics to gain insight into potential mechanisms involved in *P. fluorescens*-human interactions in health and disease. A study published from our research group in 2011 found that *Pseudomonas spp.* bacteria dominated the lower respiratory microbiota of individuals with severe chronic obstructive pulmonary disease (COPD) (207). I performed additional bioinformatic analysis on the dataset and identified that the *Pseudomonas* OTUs found in high relative abundance were not *P. aeruginosa* but were very likely members of the *P. fluorescens* species complex. This finding was highly unexpected. At the onset of this project, no literature existed connecting *P. fluorescens* bacteria and chronic respiratory disease in humans. The taxonomy of the *P. fluorescens* species complex was limited to only 10 sequenced strains AND all 10 strains had been collected from environmental sources. There were no clinical strains of *P. fluorescens* in the sequence archives. Thus, one unanswered question is whether clinical *P. fluorescens* strains are the same as environmental strains. In addition, with so few sequenced strains, the heterogeneity in the taxonomy also remains to be determined. Some early preliminary analyses of the lung microbiota from genetically identical (inbred) mice obtained from two different vendors/colonies revealed that mice from certain colonies contained *P. fluorescens* in their lungs while others did not, suggesting that *P. fluorescens* can colonize mammalian lungs (Chapter 6). The data presented in this thesis will address the gaps in knowledge of the

taxonomy and genetics of clinical *P. fluorescens* strains, as well as provide an analysis of the relationship between *P. fluorescens* outgrowth in the lungs and pulmonary inflammation.

The central hypothesis of my thesis is that members of the *Pseudomonas fluorescens* species complex are unappreciated colonizers of the human host, particularly in the content of pulmonary inflammation. To test this hypothesis, I formed the following specific aims:

1. Expand the limited taxonomy of the *P. fluorescens* species complex, including the inclusion of the first sequenced strains isolated from clinical human samples.

2. Provide the first genomic characterization of *P. fluorescens* strains isolated from clinical human samples, with a focus on genomic aspects that predict functional attributes related to survival in the human host environment.

3. Investigate the association between lung inflammation and *P. fluorescens* by addressing these two sub-aims:

    a. Analyze the effect of inflammation on *P. fluorescens* growth in the lung microbiota of mice.

    b. Taxonomically classify *Pseudomonas* OTUs from multiple lung microbiota datasets, including patients with COPD, idiopathic pulmonary fibrosis, those who had received a lung transplant, and healthy controls.

By addressing these specific aims, the research presented in this thesis will provide the first data on the genomic attributes of *P. fluorescens* strains isolated from human samples and the effect of chronic inflammation on indigenous *P. fluorescens* bacteria in the lower airways of a mammalian host.

Figure I.1: Reported *P. fluorescens* infections**.**

Organ/Tissue	Cases	References
Blood	110	(8-17, 182-188)
Bone	2	(207, 208)
Cerebral Spinal Fluid	1	(209)
Eye	3	(210-212)
Lung	3	(194-197)
Sinus	3	(213)
Skin/Wound	5	(189, 190, 193)
Urinary Tract	5	(191-193)
Uterus	1	(214)

Total cases reported in the medical literature. MEDLINE search performed with search terms "*Pseudomonas fluorescens*" and filtered for Human studies with no date or language restriction. All abstracts read and reviewed by authors; relevant references were read in entirety.

Figure I.2: *P. fluorescens* isolates cultured over a eleven-year period by the University of Michigan Hospital Microbiology Lab (3).

Breakdown of 242 *P. fluorescens* isolates cultured between 01-1-2002 & 12-13-2012	
Culture Method Used	
Cultured using routine laboratory protocols	59.50%
Cultured using CF modified protocols	40.10%
Sample Site	
Sputum samples	53.70%
Throat Swabs	21.10%
Bronchoscopically-obtained samples (BALs or brushings)	13.20%
Other (tracheal aspirates, sinus aspirates)	12.00%
Underlying Disease/Cause	
Cystic Fibrosis	38.80%
Other chronic airway disease (COPD, asthma, non-CF bronchiectasis)	16.10%
Lung transplantation	7.40%
Acute pneumonia (chronically immunosuppressed or hospital-acquired)	9.90%
Acute pneumonia (not chronically immunosuppressed or hospital-acquired)	1.60%
Other (chronic tracheostomy, sinusitis, acute respiratory distress syndrome and bone marrow transplantation).	26.20%
Co-cultured Bacteria	
'Oral flora' species	85.10%
Pseudomonas aeruginosa	25.60%
Staphylococcus aureus	15.70%
Stenotrophomonas maltophillia	11.60%

Figure I.3: The functional range and environmental niches of the *Pseudomonas* genus, highlighting the broad distribution of the *P. fluorescens* species complex.

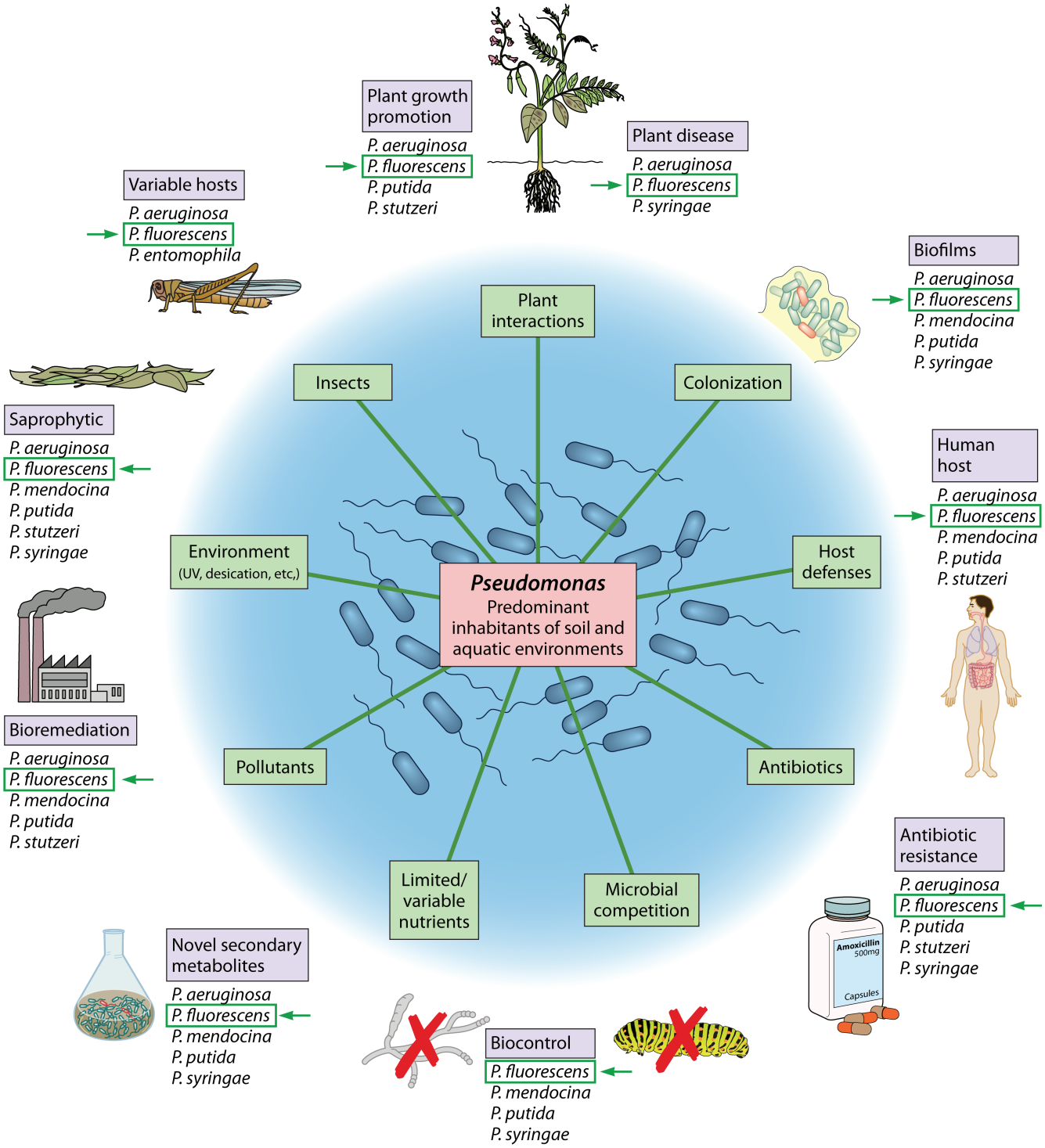


Figure I.3: The functional range and environmental niches of the *Pseudomonas* genus, highlighting the broad distribution of the *P. fluorescens* species complex. Members of the *P. fluorescens* species complex are successful colonizers in a wide range of environments and habitats due to diverse functional abilities. As discussed in this review, *P. fluorescens* has been associated with disease in humans but has not been demonstrated to be pathogen. Figure adapted and redrawn from (230), with permission.

Figure I.4: Species diversity within the *P. fluorescens* species complex.

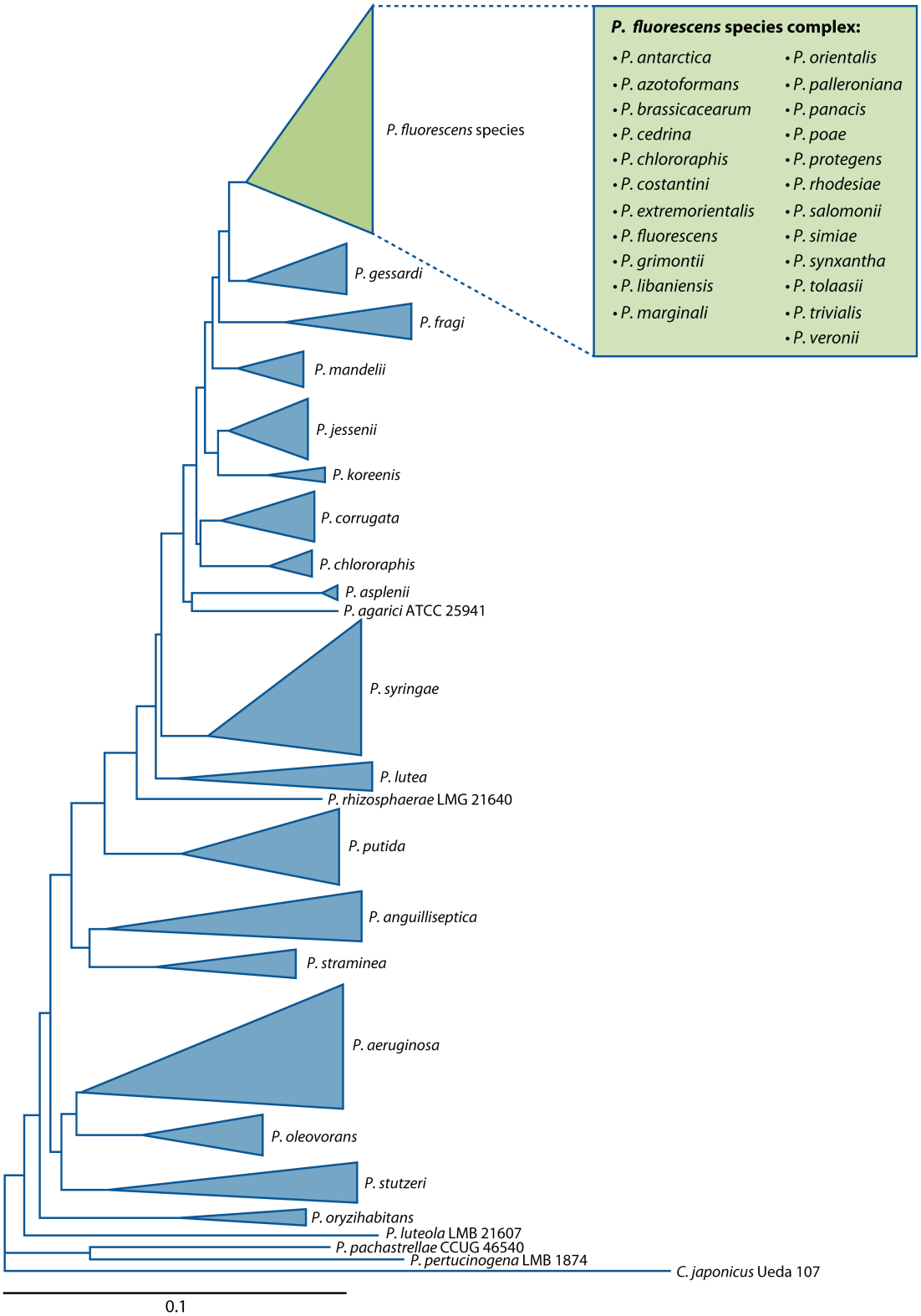


Figure I.4: Species diversity within the *P. fluorescens* species complex.
Mulet *et. al* generated a phylogenetic tree from 107 *Pseudomonas* type strains, based on the concatenated analysis of the 16S rRNA, gyrB, rpoB and rpoD genes, with Cellvibrio japonicum Ueda107 used as the outgroup (74). The bar indicates sequence divergence. Figure adapted from (74), with permission.

Figure I.5: Scanning electron micrograph of *P. fluorescens* (courtesy of Science Source).

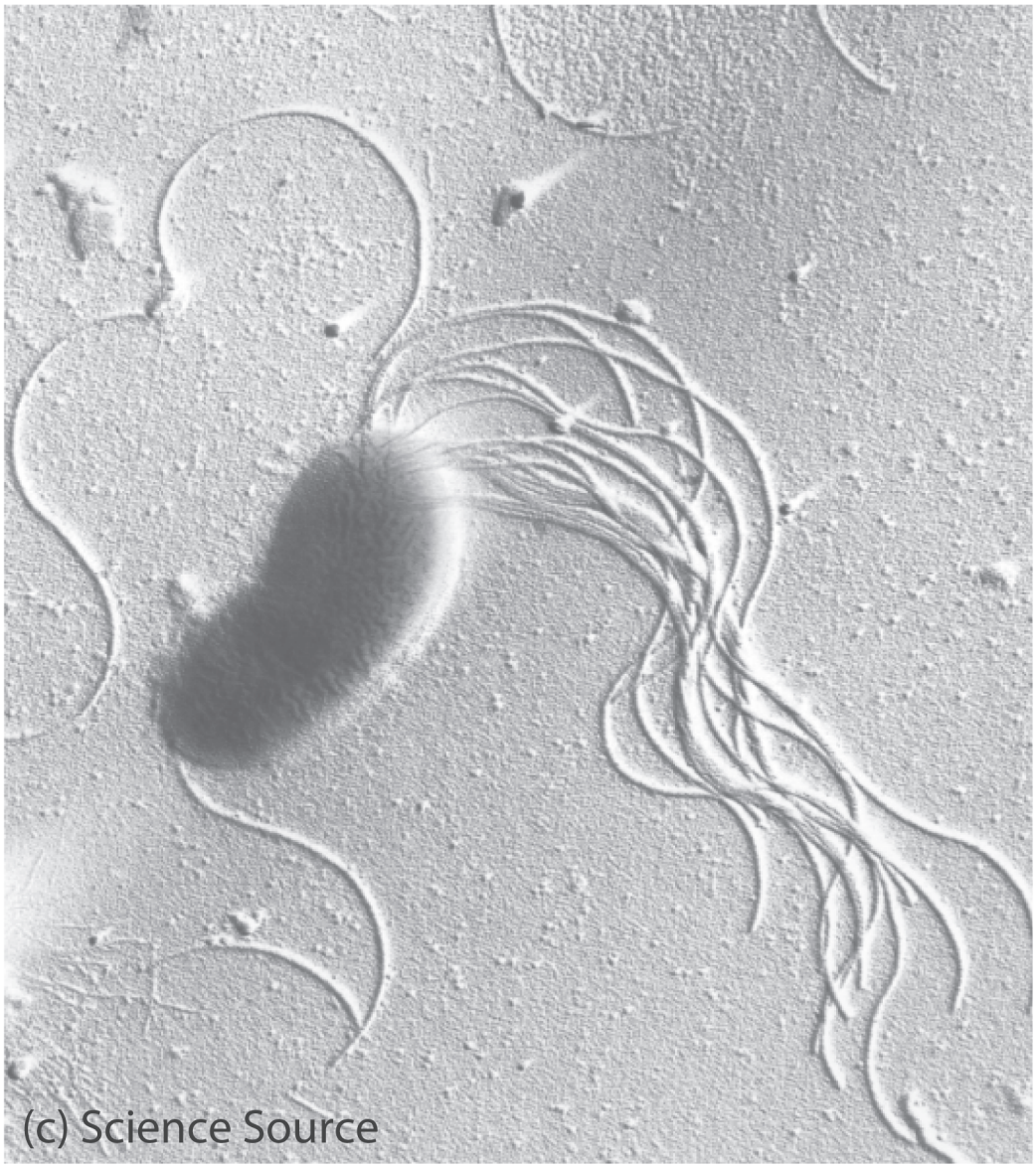


Figure I.6: Characteristics of *P. fluorescens* complex bacteria**.**

Taxonomy:
Bacteria; Proteobacteria; Gammeoproteobacteria; Pseudomonadaels; Pseudomonadaceae, Pseudomonas
Physical Characteristics:
Gram-negative, rod-shaped (bacilli)
Motile via motile polar flagella
Non-spore forming
Produces a fluorescent pigment (pyocyanin), from which it gets its name
Produces an exopolysaccharides and readily forms biofilms
Growth Characteristics:
Obligate aerobe but capable of using nitrate instead of oxygen as a final electron acceptor during cellular respiration
Optimal temperatures for growth:
-environmental isolates: 25-30°C
-mammalian isolates: 34-37°C
Oxidase positive
Catalase positive
Grows well on trypticase soy agar (TSA) and luria agar (LA)
Hemolytic activity on red blood cells
-environmental isolates: No
-certain mammalian isolates (ex: MFN1032): Yes
Forms small, white convex colonies

Figure I.7: Phylogenetic tree of 38 *Pseudomonas* type strains, based on a concatenated nine gene MLST analysis.

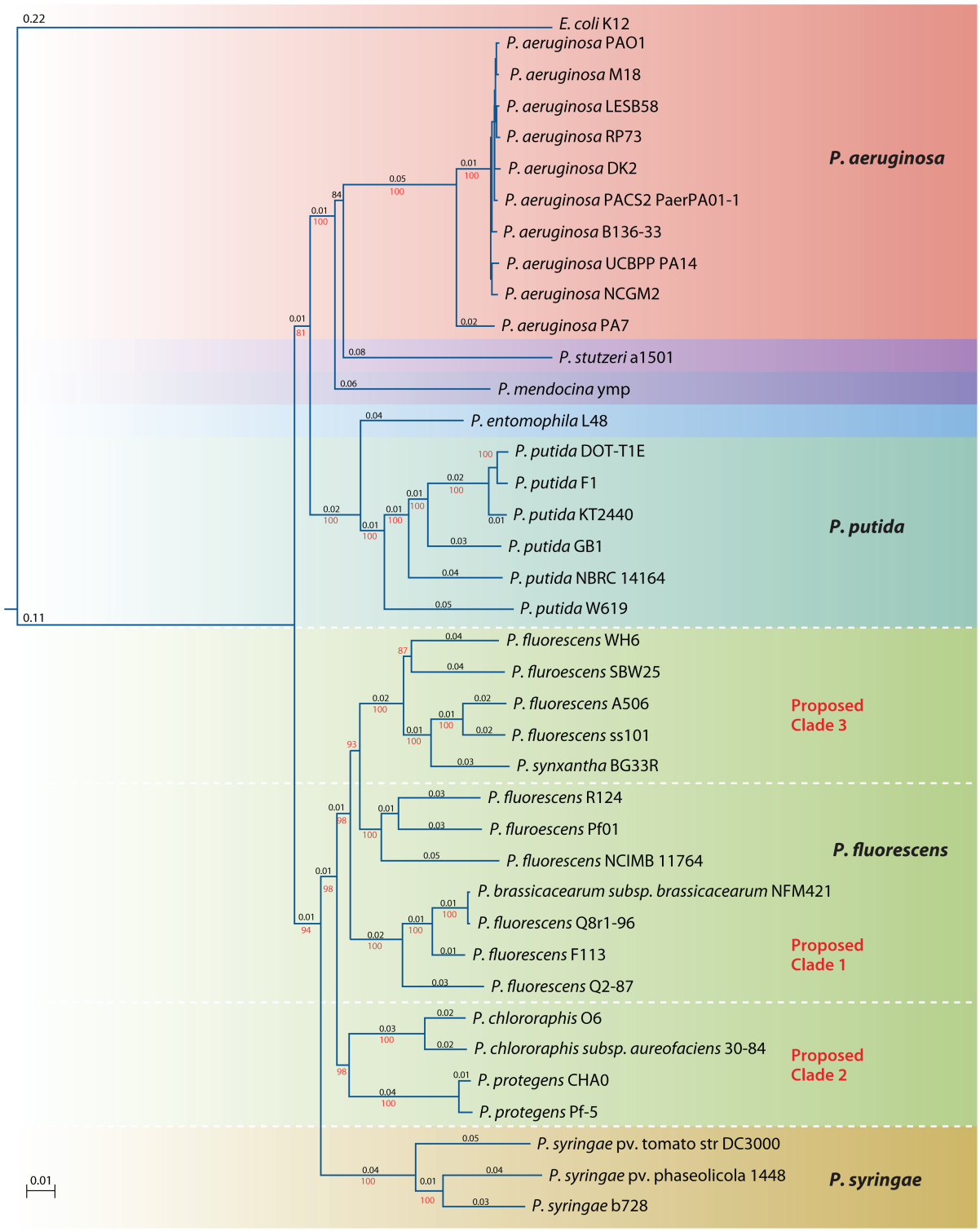


Figure I.7: Phylogenetic tree of 38 *Pseudomonas* type strains, based on a concatenated nine gene MLST analysis.
The strains selected have full genome sequences available through public databases. The MLSA analysis was performed using nine housekeeping genes (DnaE, PpsA, RecA, RpoB, GyrB, GuaA, MutL, PyrC, and AcsA), with *E.coli* strain K12 used as an outgroup. A maximum-likelihood tree was calculated in the online version of MAFFT (231, 232) and visualized with the software program Archaeopteryx (233). The confidence intervals after 1,000 bootstrap resamplings are indicated in red and the branch distance is indicated in black. The bar indicates sequence divergence. *P. fluorescens* clade destinations are based on those proposed in (49).

Figure I.8: Summary information of fully sequenced bacterial strains from the *P. fluorescens* species complex.

Strain	Isolation Source	Genome Size (Mb)	G+C%	Year Isolated/ Year Sequenced	NCBI Sequence ID	Ref.
P. fluorescens Pf0-1	Loam soil, Sherborn Massachusetts, USA	6.44	60.5	1988/ 2009	NC_007492.2	(23, 215)
P. fluorescens SBW25	Sugar beet phyllosphere, Oxfordshire, UK	6.72	60.5	1989/ 2009	NC_012660	(23)
P. fluorescens A506	Pear phyllosphere, California, USA	6.02	59.9	1994/2012	NC_017911	(49, 216)
P. fluorescens Q2-87	Wheat rhizosphere, Washington, USA (same field as Q8r1-96)	6.37	60.6	1987/2012	NZ_CM001558.1	(49, 217)
P. fluorescens Q8r1-96	Wheat rhizosphere, Washington, USA (same field as Q2-87)	6.6	61	1996/2012	NZ_CM001512.1	(49, 218)
P. fluorescens SS101	Wheat rhizosphere, near city of Bergen op Zoom, The Netherlands	6.18	60	2003/2012	NZ_CM001513	(49, 64)
P. fluorescens WH6	Rhizosphere of *Poa* sp. and *Triticum aestivum* at Hyslop Research Farm, Benton County, Oregon USA	NA	NA	2008/2010 (Draft)	NA	(219, 220)
P. fluorescens F113	Sugar-beet rhizosphere	6.85	60.8	1992/2012	NC_016830	(50, 157)
P. fluorescens R124	Tepui orthoquartzite sandstone cave in Guiana Shield, South America	6.3	NA	2007/2013	NC_CM001561	(221)
P. protegens Pf-5	Soil, Texas, USA	7.07	63.3	1978/2005	NC_004129.6	(222)
P. protegens CHA0	Tobacco roots, Morens, Switzerland	6.87	63.4	1983/2013	NC_021237.1	(223, 224)
P. brassicacearum subsp. brassicacearum NFM421	Plant-rhizosphere	6.84	60.8	NA/2011	NC_015379.1	(225)
P. fluorescens NCIM 11764	Culture-supplied with potassium cyanide as sole nitrogen source	6.97	59	1983/2012	NA	(226, 227)
P. chlororaphis subsp. aureofaciens 30-84	Wheat rhizosphere, Kansas, USA	6.67	62.9	NA/2012	NZ_CM001559	(49)
P. chlororaphis subsp. aureofaciens O6	Soil, Utah, USA	6.98	62.9	1996/2012	NZ_CM001490	(49, 228)
P. synxatha BG33R	Peach rhizosphere, South Carlonia, USA	6.3	59.6	1993/2012	NZ_CM001514	(49, 229)

Figure I.9: Phylogenetic tree of 38 *Pseudomonas* type strains, based on the V3-V5 region sequence of the 16S rRNA gene.

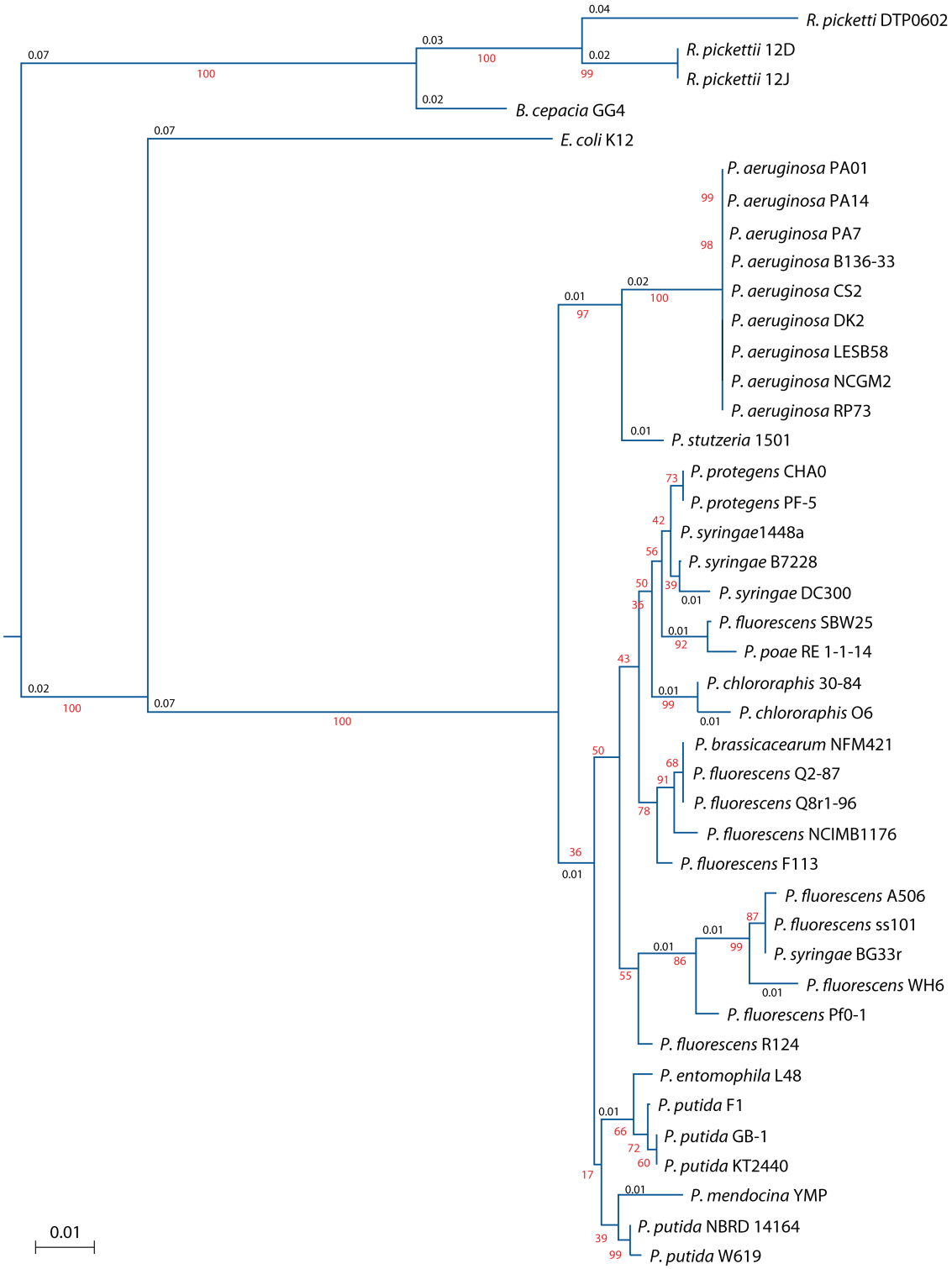


Figure I.9: Phylogenetic tree of 38 *Pseudomonas* type strains, based on the V3-V5 region sequence of the 16S rRNA gene.
The strains selected have full genome sequences available through public databases. The V3-V5 sequence primers (V3 primer: 442-492, V5 primer: 822-879; numbered according to the *E. coli* 16S rRNA gene map) (234) were aligned to each genome using DNAstar SeqBuilder Software. A maximum-likelihood tree was calculated in the online version of MAFFT (231, 232) and visualized with the software program Archaeopteryx (233). The confidence intervals after 1,000 bootstrap resamplings are indicated in red and the branch distance is indicated in black. The bar indicates sequence divergence.

Figure I.10: Scanning electron micrographs of *P. fluorescens* biofilms.

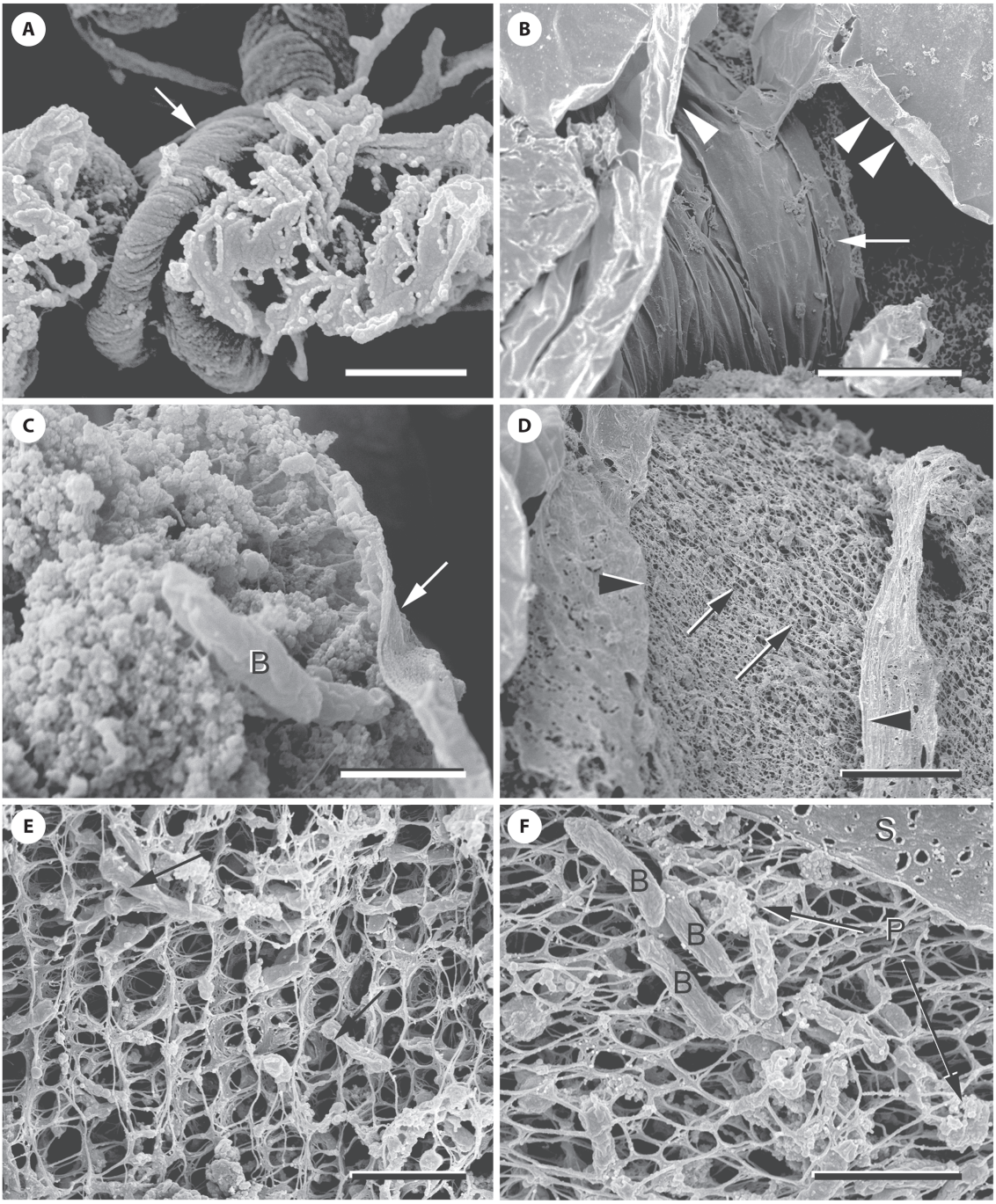


Figure I.10: Scanning electron micrographs of *P. fluorescens* biofilms.
In these photomicrographs, Baum et al. prepared and cryopreserved 14 day bioflims from *P. fluorescens* EvS4-B1 monocultures (56). "(A) Fibrillary structures made up of twisted fibers (arrow); scale bar = 1 μm. (B) Flat sheets of material (arrowhead) with some of the sheets wrapped around other structures (arrow); scale bar = 20 μm. (C) The inside core of the "wrapped" structures, consisting of bacteria [B], embedded in an extracellular matrix of particulate matter, and a thin sheet of material (arrow); scale bar = 1 μm. (D) The outer sheet (arrowheads) which envelopes an inner core consisting of fibers, forming irregular network-like structures (arrow); scale bar = 10 μm. (E) The network consisting of fibers arranged in a periodic pattern with bacteria (arrows) dispersed throughout the network; scale bar = 2 μm. (F) A sheet of material, [S], consisting of extracellular material and dead cells, covering and attaching to the fiber network that included associated bacteria [B] and particulate matter [P]; scale bar = 2 μm." Reproduced from (56), with permission.

Figure 1.11: Type-III secretion systems in *P. fluorescens*.

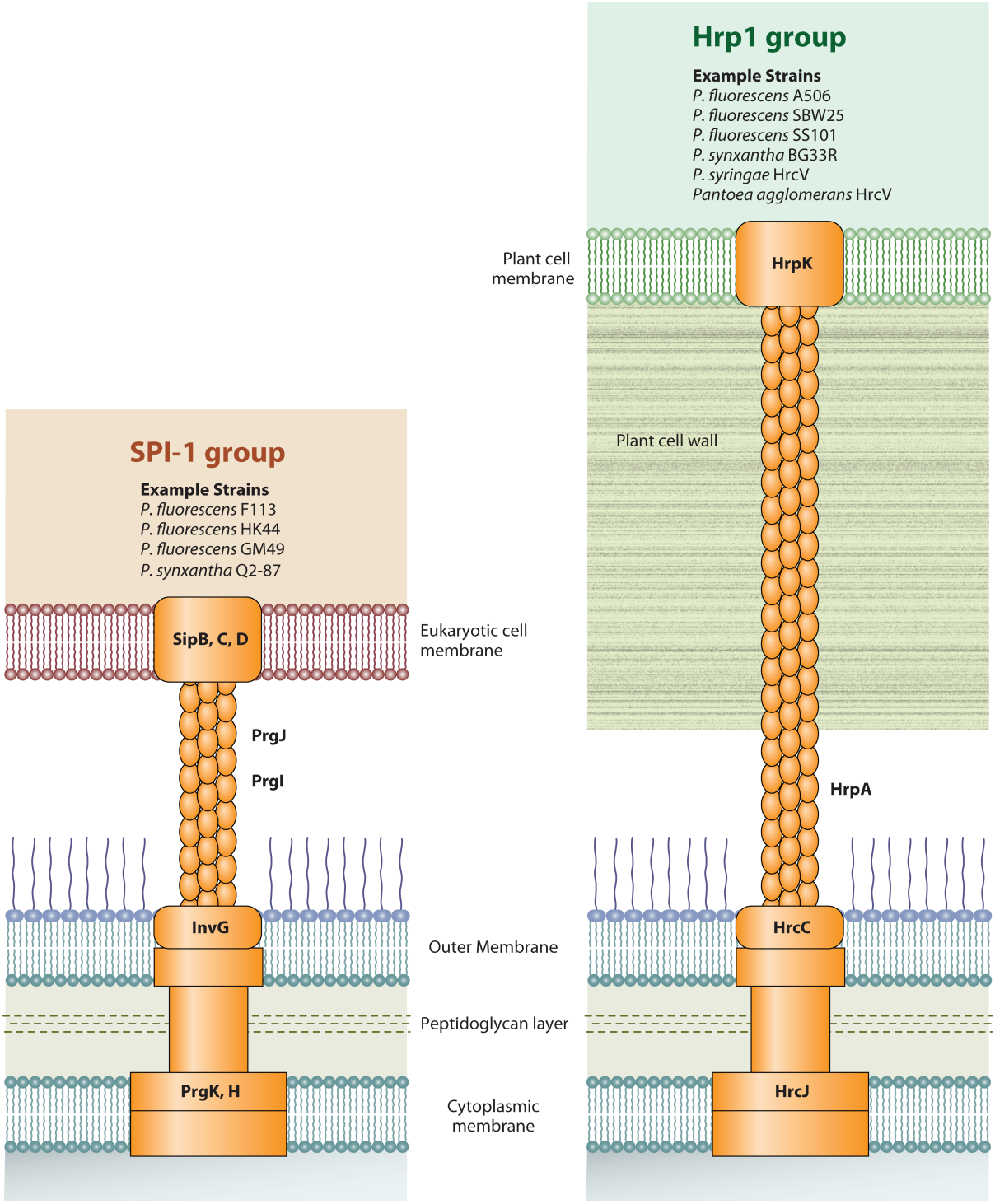


Figure I.11: Type-III secretion systems in *P. fluorescens*.
Components and structure of the SPI-I and Hrp1 systems, with the corresponding strains listed in which these systems have been identified. Sources include adaption from (146) and information from (50).

Chapter I

References

1. Diaz PI, Hong BY, Frias-Lopez J, Dupuy AK, Angeloni M, Abusleme L, Terzi E, Ioannidou E, Strausbaugh LD, Dongari-Bagtzoglou A. 2013. Transplantation-associated long-term immunosuppression promotes oral colonization by potentially opportunistic pathogens without impacting other members of the salivary bacteriome. Clin Vaccine Immunol 20:920-930.
2. Patel SK, Pratap CB, Verma AK, Jain AK, Dixit VK, Nath G. 2013. Pseudomonas fluorescens-like bacteria from the stomach: a microbiological and molecular study. World J Gastroenterol 19:1056-1067.
3. Dickson RP, Erb-Downward JR, Freeman CM, Walker N, Scales BS, Beck JM, Martinez FJ, Curtis JL, Lama VN, Huffnagle GB. 2014. Changes in the Lung Microbiome following Lung Transplantation Include the Emergence of Two Distinct Pseudomonas Species with Distinct Clinical Associations. PLoS One 9:e97214.
4. Stenhouse MA, Milner LV. 1992. A survey of cold-growing gram-negative organisms isolated from the skin of prospective blood donors. Transfus Med 2:235-237.
5. Wagner J, Short K, Catto-Smith AG, Cameron DJ, Bishop RF, Kirkwood CD. 2008. Identification and characterisation of Pseudomonas 16S ribosomal DNA from ileal biopsies of children with Crohn's disease. PLoS One 3:e3578.
6. Baader A, Garre C. 1887. Über Antagonisten unter den Bacterien. Correspondenz-Blatt für Schweizer Ärzte 13:385-392.
7. Gibaud M, Martin-Dupont P, Dominguez M, Laurentjoye P, Chassaing B, Leng B. 1984. [Pseudomonas fluorescens septicemia following transfusion of contaminated blood]. Presse Med 13:2583-2584.

8. Hsueh PR, Teng LJ, Pan HJ, Chen YC, Sun CC, Ho SW, Luh KT. 1998. Outbreak of Pseudomonas fluorescens bacteremia among oncology patients. J Clin Microbiol 36:2914-2917.

9. Khabbaz RF, Arnow PM, Highsmith AK, Herwaldt LA, Chou T, Jarvis WR, Lerche NW, Allen JR. 1984. Pseudomonas fluorescens bacteremia from blood transfusion. Am J Med 76:62-68.

10. Murray AE, Bartzokas CA, Shepherd AJ, Roberts FM. 1987. Blood transfusion-associated Pseudomonas fluorescens septicaemia: is this an increasing problem? J Hosp Infect 9:243-248.

11. Pittman M. 1953. A study of bacteria implicated in transfusion reactions and of bacteria isolated from blood products. The Journal of laboratory and clinical medicine 42:273.

12. Scott J, Boulton FE, Govan JR, Miles RS, McClelland DB, Prowse CV. 1988. A fatal transfusion reaction associated with blood contaminated with Pseudomonas fluorescens. Vox Sang 54:201-204.

13. Benito N, Mirelis B, Luz Galvez M, Vila M, Lopez-Contreras J, Cotura A, Pomar V, March F, Navarro F, Coll P, Gurgui M. 2012. Outbreak of Pseudomonas fluorescens bloodstream infection in a coronary care unit. J Hosp Infect 82:286-289.

14. 2005. Pseudomonas bloodstream infections associated with a heparin/saline flush--Missouri, New York, Texas, and Michigan, 2004-2005. MMWR Morb Mortal Wkly Rep 54:269-272.

15. 2006. Update: Delayed onset Pseudomonas fluorescens bloodstream infections after exposure to contaminated heparin flush--Michigan and South Dakota, 2005-2006. MMWR Morb Mortal Wkly Rep 55:961-963.

16. Gershman MD, Kennedy DJ, Noble-Wang J, Kim C, Gullion J, Kacica M, Jensen B, Pascoe N, Saiman L, McHale J, Wilkins M, Schoonmaker-Bopp D, Clayton J, Arduino M, Srinivasan A. 2008. Multistate outbreak of Pseudomonas fluorescens bloodstream

infection after exposure to contaminated heparinized saline flush prepared by a compounding pharmacy. Clin Infect Dis 47:1372-1379.

17. Sarubbi FA, Jr., Wilson B, Lee M, Brokopp C. 1978. Nosocomial meningitis and bacteremia due to contaminated amphotericin B. JAMA 239:416-418.

18. Landers CJ, Cohavy O, Misra R, Yang H, Lin YC, Braun J, Targan SR. 2002. Selected loss of tolerance evidenced by Crohn's disease-associated immune responses to auto- and microbial antigens. Gastroenterology 123:689-699.

19. Arnott ID, Landers CJ, Nimmo EJ, Drummond HE, Smith BK, Targan SR, Satsangi J. 2004. Sero-reactivity to microbial components in Crohn's disease is associated with disease severity and progression, but not NOD2/CARD15 genotype. Am J Gastroenterol 99:2376-2384.

20. Iltanen S, Tervo L, Halttunen T, Wei B, Braun J, Rantala I, Honkanen T, Kronenberg M, Cheroutre H, Turovskaya O, Autio V, Ashorn M. 2006. Elevated serum anti-I2 and anti-OmpW antibody levels in children with IBD. Inflamm Bowel Dis 12:389-394.

21. Mow WS, Landers CJ, Steinhart AH, Feagan BG, Croitoru K, Seidman E, Greenberg GR, Targan SR. 2004. High-level serum antibodies to bacterial antigens are associated with antibiotic-induced clinical remission in Crohn's disease: a pilot study. Dig Dis Sci 49:1280-1286.

22. Spivak J, Landers CJ, Vasiliauskas EA, Abreu MT, Dubinsky MC, Papadakis KA, Ippoliti A, Targan SR, Fleshner PR. 2006. Antibodies to I2 predict clinical response to fecal diversion in Crohn's disease. Inflamm Bowel Dis 12:1122-1130.

23. Silby MW, Cerdeno-Tarraga AM, Vernikos GS, Giddens SR, Jackson RW, Preston GM, Zhang XX, Moon CD, Gehrig SM, Godfrey SA, Knight CG, Malone JG, Robinson Z, Spiers AJ, Harris S, Challis GL, Yaxley AM, Harris D, Seeger K, Murphy L, Rutter S, Squares R, Quail MA, Saunders E, Mavromatis K, Brettin TS, Bentley SD, Hothersall J, Stephens E, Thomas CM, Parkhill J, Levy SB, Rainey PB, Thomson NR. 2009. Genomic

and genetic analyses of diversity and plant interactions of Pseudomonas fluorescens. Genome Biol 10:R51.

24. Feazel LM, Baumgartner LK, Peterson KL, Frank DN, Harris JK, Pace NR. 2009. Opportunistic pathogens enriched in showerhead biofilms. Proc Natl Acad Sci U S A 106:16393-16399.

25. Haas D, Defago G. 2005. Biological control of soil-borne pathogens by fluorescent pseudomonads. Nat Rev Microbiol 3:307-319.

26. Laursen JB, Nielsen J. 2004. Phenazine natural products: biosynthesis, synthetic analogues, and biological activity. Chem Rev 104:1663-1686.

27. Weller DM, Landa BB, Mavrodi OV, Schroeder KL, De La Fuente L, Blouin Bankhead S, Allende Molar R, Bonsall RF, Mavrodi DV, Thomashow LS. 2007. Role of 2,4-diacetylphloroglucinol-producing fluorescent Pseudomonas spp. in the defense of plant roots. Plant Biol (Stuttg) 9:4-20.

28. Mavrodi DV, Blankenfeldt W, Thomashow LS. 2006. Phenazine compounds in fluorescent Pseudomonas spp. biosynthesis and regulation. Annu Rev Phytopathol 44:417-445.

29. Ramette A, Moenne-Loccoz Y, Defago G. 2003. Prevalence of fluorescent pseudomonads producing antifungal phloroglucinols and/or hydrogen cyanide in soils naturally suppressive or conducive to tobacco black root rot. FEMS Microbiol Ecol 44:35-43.

30. Keel C, Weller DM, Natsch A, Defago G, Cook RJ, Thomashow LS. 1996. Conservation of the 2,4-diacetylphloroglucinol biosynthesis locus among fluorescent Pseudomonas strains from diverse geographic locations. Appl Environ Microbiol 62:552-563.

31. Moynihan JA, Morrissey JP, Coppoolse ER, Stiekema WJ, O'Gara F, Boyd EF. 2009. Evolutionary history of the phl gene cluster in the plant-associated bacterium Pseudomonas fluorescens. Appl Environ Microbiol 75:2122-2131.

32. Takahashi M, Matsumoto S, Iwasaki S, Yahara I. 1990. Molecular basis for determining the sensitivity of eucaryotes to the antimitotic drug rhizoxin. Mol Gen Genet 222:169-175.

33. Gross H, Loper JE. 2009. Genomics of secondary metabolite production by Pseudomonas spp. Nat Prod Rep 26:1408-1446.

34. Tsuruo T, Oh-hara T, Iida H, Tsukagoshi S, Sato Z, Matsuda I, Iwasaki S, Okuda S, Shimizu F, Sasagawa K, et al. 1986. Rhizoxin, a macrocyclic lactone antibiotic, as a new antitumor agent against human and murine tumor cells and their vincristine-resistant sublines. Cancer Res 46:381-385.

35. Schnider U, Keel C, Blumer C, Troxler J, Defago G, Haas D. 1995. Amplification of the housekeeping sigma factor in Pseudomonas fluorescens CHA0 enhances antibiotic production and improves biocontrol abilities. J Bacteriol 177:5387-5392.

36. Sarniguet A, Kraus J, Henkels MD, Muehlchen AM, Loper JE. 1995. The sigma factor sigma s affects antibiotic production and biological control activity of Pseudomonas fluorescens Pf-5. Proc Natl Acad Sci U S A 92:12255-12259.

37. Rosenberg K, Bertaux J, Krome K, Hartmann A, Scheu S, Bonkowski M. 2009. Soil amoebae rapidly change bacterial community composition in the rhizosphere of Arabidopsis thaliana. ISME J 3:675-684.

38. Ronn R, McCaig AE, Griffiths BS, Prosser JI. 2002. Impact of protozoan grazing on bacterial community structure in soil microcosms. Appl Environ Microbiol 68:6094-6105.

39. Umio S, Kawanishi T, Kamishita T, Mine Y. 1987. Antifungal Composition Employing Pyrrolnitrin in Combination wtih an Imidazole Compound. Japan.

40. Fuller AT, Mellows G, Woolford M, Banks GT, Barrow KD, Chain EB. 1971. Pseudomonic acid: an antibiotic produced by Pseudomonas fluorescens. Nature 234:416-417.

41. Bork K, Brauers J, Kresken M. 1989. Efficacy and safety of 2% mupirocin ointment in the treatment of primary and secondary skin infections--an open multicentre trial. Br J Clin Pract 43:284-288.

42. Strock LL, Lee MM, Rutan RL, Desai MH, Robson MC, Herndon DN, Heggers JP. 1990. Topical Bactroban (mupirocin): efficacy in treating burn wounds infected with methicillin-resistant staphylococci. J Burn Care Rehabil 11:454-459.

43. Neilands JB. 1981. Iron absorption and transport in microorganisms. Annu Rev Nutr 1:27-46.

44. Youard ZA, Mislin GL, Majcherczyk PA, Schalk IJ, Reimmann C. 2007. Pseudomonas fluorescens CHA0 produces enantio-pyochelin, the optical antipode of the Pseudomonas aeruginosa siderophore pyochelin. J Biol Chem 282:35546-35553.

45. Wuest WM, Sattely ES, Walsh CT. 2009. Three siderophores from one bacterial enzymatic assembly line. J Am Chem Soc 131:5056-5057.

46. Marchi M, Boutin M, Gazengel K, Rispe C, Gauthier JP, Guillerm-Erckelboudt AY, Lebreton L, Barret M, Daval S, Sarniguet A. 2013. Genomic analysis of the biocontrol strain Pseudomonas fluorescens Pf29Arp with evidence of T3SS and T6SS gene expression on plant roots. Environ Microbiol Rep 5:393-403.

47. Preston GM, Bertrand N, Rainey PB. 2001. Type III secretion in plant growth-promoting Pseudomonas fluorescens SBW25. Mol Microbiol 41:999-1014.

48. Rezzonico F, Binder C, Defago G, Moenne-Loccoz Y. 2005. The type III secretion system of biocontrol Pseudomonas fluorescens KD targets the phytopathogenic Chromista Pythium ultimum and promotes cucumber protection. Mol Plant Microbe Interact 18:991-1001.

49. Loper JE, Hassan KA, Mavrodi DV, Davis EW, 2nd, Lim CK, Shaffer BT, Elbourne LD, Stockwell VO, Hartney SL, Breakwell K, Henkels MD, Tetu SG, Rangel LI, Kidarsa TA, Wilson NL, van de Mortel JE, Song C, Blumhagen R, Radune D, Hostetler JB, Brinkac

LM, Durkin AS, Kluepfel DA, Wechter WP, Anderson AJ, Kim YC, Pierson LS, 3rd, Pierson EA, Lindow SE, Kobayashi DY, Raaijmakers JM, Weller DM, Thomashow LS, Allen AE, Paulsen IT. 2012. Comparative genomics of plant-associated Pseudomonas spp.: insights into diversity and inheritance of traits involved in multitrophic interactions. PLoS Genet 8:e1002784.

50. Redondo-Nieto M, Barret M, Morrisey JP, Germaine K, Martinez-Granero F, Barahona E, Navazo A, Sanchez-Contreras M, Moynihan JA, Giddens SR, Coppoolse ER, Muriel C, Stiekema WJ, Rainey PB, Dowling D, O'Gara F, Martin M, Rivilla R. 2012. Genome sequence of the biocontrol strain Pseudomonas fluorescens F113. J Bacteriol 194:1273-1274.

51. Barret M, Egan F, Moynihan J, Morrissey JP, Lesouhaitier O, O'Gara F. 2013. Characterization of the SPI-1 and Rsp type three secretion systems in Pseudomonas fluorescens F113. Environ Microbiol Rep 5:377-386.

52. Tuttlebee CM, O'Donnell MJ, Keane CT, Russell RJ, Sullivan DJ, Falkiner F, Coleman DC. 2002. Effective control of dental chair unit waterline biofilm and marked reduction of bacterial contamination of output water using two peroxide-based disinfectants. J Hosp Infect 52:192-205.

53. Donnarumma G, Buommino E, Fusco A, Paoletti I, Auricchio L, Tufano MA. 2010. Effect of temperature on the shift of Pseudomonas fluorescens from an environmental microorganism to a potential human pathogen. Int J Immunopathol Pharmacol 23:227-234.

54. Bianciotto V, Andreotti S, Balestrini R, Bonfante P, Perotto S. 2001. Mucoid mutants of the biocontrol strain pseudomonas fluorescens CHA0 show increased ability in biofilm formation on mycorrhizal and nonmycorrhizal carrot roots. Mol Plant Microbe Interact 14:255-260.

55. O'Toole GA, Kolter R. 1998. Initiation of biofilm formation in Pseudomonas fluorescens WCS365 proceeds via multiple, convergent signalling pathways: a genetic analysis. Mol Microbiol 28:449-461.

56. Baum MM, Kainovic A, O'Keeffe T, Pandita R, McDonald K, Wu S, Webster P. 2009. Characterization of structures in biofilms formed by a Pseudomonas fluorescens isolated from soil. BMC Microbiol 9:103.

57. Chapalain A, Rossignol G, Lesouhaitier O, Merieau A, Gruffaz C, Guerillon J, Meyer JM, Orange N, Feuilloley MG. 2008. Comparative study of 7 fluorescent pseudomonad clinical isolates. Can J Microbiol 54:19-27.

58. Rossignol G, Sperandio D, Guerillon J, Duclairoir Poc C, Soum-Soutera E, Orange N, Feuilloley MG, Merieau A. 2009. Phenotypic variation in the Pseudomonas fluorescens clinical strain MFN1032. Res Microbiol 160:337-344.

59. Sperandio D, Decoin V, Latour X, Mijouin L, Hillion M, Feuilloley MG, Orange N, Merieau A. 2012. Virulence of the Pseudomonas fluorescens clinical strain MFN1032 towards Dictyostelium discoideum and macrophages in relation with type III secretion system. BMC Microbiol 12:223.

60. Almeida LA, Araujo R. 2013. Highlights on molecular identification of closely related species. Infect Genet Evol 13:67-75.

61. Konstantinidis KT, Ramette A, Tiedje JM. 2006. Toward a more robust assessment of intraspecies diversity, using fewer genetic markers. Appl Environ Microbiol 72:7286-7293.

62. Moore ERB, Tindall BJ, Dos Santos VAPM, Pieper DH, Ramos JL, Palleron NJ. 2006. Nonmedical: Pseudomonas. Prokaryotes: A Handbook on the Biology of Bacteria, Vol 6, Third Edition:646-703.

63. Howell C, Stipanovic R. 1980. Suppression of *Pythium ultimum*-induced damping-off of cotton seedlings by *Pseudomonas fluorescens* and its antibiotic, pyoluteorin. . Phytopathology 70:712-715.

64. De Souza JT, De Boer M, De Waard P, Van Beek TA, Raaijmakers JM. 2003. Biochemical, genetic, and zoosporicidal properties of cyclic lipopeptide surfactants produced by Pseudomonas fluorescens. Appl Environ Microbiol 69:7161-7172.

65. Valverde C, Heeb S, Keel C, Haas D. 2003. RsmY, a small regulatory RNA, is required in concert with RsmZ for GacA-dependent expression of biocontrol traits in Pseudomonas fluorescens CHA0. Mol Microbiol 50:1361-1379.

66. Cosson P, Soldati T. 2008. Eat, kill or die: when amoeba meets bacteria. Curr Opin Microbiol 11:271-276.

67. King EO, Ward MK, Raney DE. 1954. Two simple media for the demonstration of pyocyanin and fluorescin. J Lab Clin Med 44:301-307.

68. Lowbury EJ, Collins AG. 1955. The use of a new cetrimide product in a selective medium for Pseudomonas pyocyanea. J Clin Pathol 8:47-48.

69. Bunker ST, Bates TC, Oliver JD. 2004. Effects of temperature on detection of plasmid or chromosomally encoded gfp- and lux-labeled Pseudomonas fluorescens in soil. Environ Biosafety Res 3:83-90.

70. Oliver J. 1993. Formation of viable but nonculturable cells., p. 239-272. *In* S IK (ed.), Starvation in bacteria. Plenum Press, New York.

71. Rowan N. 2004. Viable but non-cultruable forms of food and waterborne bacteria: Quo Vadis? Trends in Food Science & Technology 15:462-467.

72. Oliver JD. 1995. The viable but non-culturable state in the human pathogen Vibrio vulnificus. FEMS Microbiol Lett 133:203-208.

73. Oliver JD. 2005. The viable but nonculturable state in bacteria. J Microbiol 43 Spec No:93-100.

74. Mulet M, Lalucat J, Garcia-Valdes E. 2010. DNA sequence-based analysis of the Pseudomonas species. Environ Microbiol 12:1513-1530.

75. Peix A, Ramirez-Bahena MH, Velazquez E. 2009. Historical evolution and current status of the taxonomy of genus Pseudomonas. Infect Genet Evol 9:1132-1147.

76. Anzai Y, Kim H, Park JY, Wakabayashi H, Oyaizu H. 2000. Phylogenetic affiliation of the pseudomonads based on 16S rRNA sequence. Int J Syst Evol Microbiol 50 Pt 4:1563-1589.

77. Palleroni NJ. 2010. The Pseudomonas story. Environ Microbiol 12:1377-1383.

78. Palleroni NJ. 2008. The road to the taxonomy of *Pseudomonas*., p. 1-18. *In* Cornelis P (ed.), *Pseudomonas. Genomics and Molecular Biology.* Caister Academic Press, Norfolk, UK.

79. Moore ERB, Mau M, Arnscheidt A, Bottger EC, Hutson RA, Collins MD, Peer YVD, Wachter RD, Timmis KN. 1996. The determination and comparison of the 16S rRNA gene sequences of species of the genus pseudomonas (sensu stricto) and estimation of the natural intragereric relationships. Syst Appl Microbiol 1:478-492.

80. Yamamoto S, Kasai H, Arnold DL, Jackson RW, Vivian A, Harayama S. 2000. Phylogeny of the genus Pseudomonas: intrageneric structure reconstructed from the nucleotide sequences of gyrB and rpoD genes. Microbiology 146 (Pt 10):2385-2394.

81. Urwin R, Maiden MC. 2003. Multi-locus sequence typing: a tool for global epidemiology. Trends Microbiol 11:479-487.

82. Chan MS, Maiden MC, Spratt BG. 2001. Database-driven multi locus sequence typing (MLST) of bacterial pathogens. Bioinformatics 17:1077-1083.

83. Sullivan CB, Diggle MA, Clarke SC. 2005. Multilocus sequence typing: Data analysis in clinical microbiology and public health. Mol Biotechnol 29:245-254.

84. Byun R, Elbourne LD, Lan R, Reeves PR. 1999. Evolutionary relationships of pathogenic clones of Vibrio cholerae by sequence analysis of four housekeeping genes. Infect Immun 67:1116-1124.

85. Savli H, Karadenizli A, Kolayli F, Gundes S, Ozbek U, Vahaboglu H. 2003. Expression stability of six housekeeping genes: A proposal for resistance gene quantification studies of Pseudomonas aeruginosa by real-time quantitative RT-PCR. J Med Microbiol 52:403-408.

86. Dagerhamn J, Blomberg C, Browall S, Sjostrom K, Morfeldt E, Henriques-Normark B. 2008. Determination of accessory gene patterns predicts the same relatedness among strains of Streptococcus pneumoniae as sequencing of housekeeping genes does and represents a novel approach in molecular epidemiology. J Clin Microbiol 46:863-868.

87. Adekambi T, Drancourt M, Raoult D. 2009. The rpoB gene as a tool for clinical microbiologists. Trends Microbiol 17:37-45.

88. Tayeb LA, Lefevre M, Passet V, Diancourt L, Brisse S, Grimont PA. 2008. Comparative phylogenies of Burkholderia, Ralstonia, Comamonas, Brevundimonas and related organisms derived from rpoB, gyrB and rrs gene sequences. Res Microbiol 159:169-177.

89. van Passel MW, Kuramae EE, Luyf AC, Bart A, Boekhout T. 2006. The reach of the genome signature in prokaryotes. BMC Evol Biol 6:84.

90. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol 57:81-91.

91. Salmond GP, Reeves PJ. 1993. Membrane traffic wardens and protein secretion in gram-negative bacteria. Trends Biochem Sci 18:7-12.

92. Cornelis P. 2010. Iron uptake and metabolism in pseudomonads. Appl Microbiol Biotechnol 86:1637-1645.

93. Cornelis P, Matthijs S, Van Oeffelen L. 2009. Iron uptake regulation in Pseudomonas aeruginosa. Biometals 22:15-22.

94. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772-780.

95. Katoh K, Frith MC. 2012. Adding unaligned sequences into an existing alignment using MAFFT and LAST. Bioinformatics 28:3144-3146.

96. Katoh K, Kuma K, Toh H, Miyata T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res 33:511-518.

97. Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059-3066.

98. Botha P, Archer L, Anderson RL, Lordan J, Dark JH, Corris PA, Gould K, Fisher AJ. 2008. Pseudomonas aeruginosa colonization of the allograft after lung transplantation and the risk of bronchiolitis obliterans syndrome. Transplantation 85:771-774.

99. Vos R, Vanaudenaerde BM, Geudens N, Dupont LJ, Van Raemdonck DE, Verleden GM. 2008. Pseudomonal airway colonisation: risk factor for bronchiolitis obliterans syndrome after lung transplantation? European Respiratory Journal 31:1037-1045.

100. Gottlieb J, Mattner F, Weissbrodt H, Dierich M, Fuehner T, Strueber M, Simon A, Welte T. 2009. Impact of graft colonization with gram-negative bacteria after lung transplantation on the development of bronchiolitis obliterans syndrome in recipients with cystic fibrosis. Respiratory Medicine 103:743-749.

101. Willner DL, Hugenholtz P, Yerkovich ST, Tan ME, Daly JN, Lachner N, Hopkins PM, Chambers DC. 2013. Re-Establishment of Recipient-Associated Microbiota in the Lung Allograft is Linked to Reduced Risk of Bronchiolitis Obliterans Syndrome. American Journal of Respiratory and Critical Care Medicine.

102. Enderby B, Smith D, Carroll W, Lenney W. 2009. Hydrogen cyanide as a biomarker for Pseudomonas aeruginosa in the breath of children with cystic fibrosis. Pediatr Pulmonol 44:142-147.

103. Laville J, Blumer C, Von Schroetter C, Gaia V, Defago G, Keel C, Haas D. 1998. Characterization of the hcnABC gene cluster encoding hydrogen cyanide synthase and anaerobic regulation by ANR in the strictly aerobic biocontrol agent Pseudomonas fluorescens CHA0. J Bacteriol 180:3187-3196.

104. Partida-Martinez LP, Hertweck C. 2005. Pathogenic fungus harbours endosymbiotic bacteria for toxin production. Nature 437:884-888.

105. Tekeda R. 1958. *Pseudomonas* pigments: I. Pyoluteorin, a new chlorine-containing pigment produced by *Pseudomonas aeruginosa.* Hakko Kgaku Zasshi 36:281-290.

106. Ligon JM, Hill DS, Hammer PE, Torkewitz NR, Hofmann D, Kempf H, Pee H. 2000. Natural products with antifungal activty from *Pseudomonas* biocontrol bacteria. . Pest Manag Sci 56:688-695.

107. Hothersall J, Wu J, Rahman AS, Shields JA, Haddock J, Johnson N, Cooper SM, Stephens ER, Cox RJ, Crosby J, Willis CL, Simpson TJ, Thomas CM. 2007. Mutational analysis reveals that all tailoring region genes are required for production of polyketide antibiotic mupirocin by pseudomonas fluorescens: pseudomonic acid B biosynthesis precedes pseudomonic acid A. J Biol Chem 282:15451-15461.

108. Sutherland R, Boon RJ, Griffin KE, Masters PJ, Slocombe B, White AR. 1985. Antibacterial activity of mupirocin (pseudomonic acid), a new antibiotic for topical use. Antimicrob Agents Chemother 27:495-498.

109. Villiger JW, Robertson WD, Kanji K, Ah Chan M, Fetherston J, Hague IK, Haycock D, Hunter P. 1986. A comparison of the new topical antibiotic mupirocin ('Bactroban') with oral antibiotics in the treatment of skin infections in general practice. Curr Med Res Opin 10:339-345.

110. Odom RB. 1989. Mupirocin (2 percent) ointment in the treatment of primary and secondary skin infections. Cutis 43:599-601.

111. Mertz PM, Marshall DA, Eaglstein WH, Piovanetti Y, Montalvo J. 1989. Topical mupirocin treatment of impetigo is equal to oral erythromycin therapy. Arch Dermatol 125:1069-1073.

112. Yanagisawa T, Lee JT, Wu HC, Kawakami M. 1994. Relationship of protein structure of isoleucyl-tRNA synthetase with pseudomonic acid resistance of Escherichia coli. A proposed mode of action of pseudomonic acid as an inhibitor of isoleucyl-tRNA synthetase. J Biol Chem 269:24304-24309.

113. Hughes J, Mellows G. 1978. On the mode of action of pseudomonic acid: inhibition of protein synthesis in Staphylococcus aureus. J Antibiot (Tokyo) 31:330-335.

114. Hughes J, Mellows G. 1980. Interaction of pseudomonic acid A with Escherichia coli B isoleucyl-tRNA synthetase. Biochem J 191:209-219.

115. Hughes J, Mellows G, Soughton S. 1980. How does Pseudomonas fluorescens, the producing organism of the antibiotic pseudomonic acid A, avoid suicide? FEBS Lett 122:322-324.

116. Hohnadel D, Meyer JM. 1988. Specificity of pyoverdine-mediated iron uptake among fluorescent Pseudomonas strains. J Bacteriol 170:4865-4873.

117. Mossialos D, Meyer JM, Budzikiewicz H, Wolff U, Koedam N, Baysse C, Anjaiah V, Cornelis P. 2000. Quinolobactin, a new siderophore of Pseudomonas fluorescens ATCC 17400, the production of which is repressed by the cognate pyoverdine. Appl Environ Microbiol 66:487-492.

118. Mercado-Blanco J, van der Drift KM, Olsson PE, Thomas-Oates JE, van Loon LC, Bakker PA. 2001. Analysis of the pmsCEAB gene cluster involved in biosynthesis of salicylic acid and the siderophore pseudomonine in the biocontrol strain Pseudomonas fluorescens WCS374. J Bacteriol 183:1909-1920.

119. Gennari M, Dragotto F. 1992. A study of the incidence of different fluorescent Pseudomonas species and biovars in the microflora of fresh and spoiled meat and fish, raw milk, cheese, soil and water. J Appl Bacteriol 72:281-288.

120. Andreani NA, Martino ME, Fasolato L, Carraro L, Montemurro F, Mioni R, Bordin P, Cardazzo B. 2014. Tracking the blue: a MLST approach to characterise the Pseudomonas fluorescens group. Food Microbiol 39:116-126.

121. Lau GW, Hassett DJ, Ran H, Kong F. 2004. The role of pyocyanin in Pseudomonas aeruginosa infection. Trends Mol Med 10:599-606.

122. Laville J, Voisard C, Keel C, Maurhofer M, Defago G, Haas D. 1992. Global control in Pseudomonas fluorescens mediating antibiotic synthesis and suppression of black root rot of tobacco. Proc Natl Acad Sci U S A 89:1562-1566.

123. Maurhofer M, Reimmann C, Schmidli-Sacherer P, Heeb S, Haas D, Defago G. 1998. Salicylic Acid Biosynthetic Genes Expressed in Pseudomonas fluorescens Strain P3 Improve the Induction of Systemic Resistance in Tobacco Against Tobacco Necrosis Virus. Phytopathology 88:678-684.

124. Sacherer P, Defago G, Haas D. 1994. Extracellular protease and phospholipase C are controlled by the global regulatory gene gacA in the biocontrol strain Pseudomonas fluorescens CHA0. FEMS Microbiol Lett 116:155-160.

125. Schnider-Keel U, Seematter A, Maurhofer M, Blumer C, Duffy B, Gigot-Bonnefoy C, Reimmann C, Notz R, Defago G, Haas D, Keel C. 2000. Autoinduction of 2,4-diacetylphloroglucinol biosynthesis in the biocontrol agent Pseudomonas fluorescens CHA0 and repression by the bacterial metabolites salicylate and pyoluteorin. J Bacteriol 182:1215-1225.

126. Pessi G, Haas D. 2001. Dual control of hydrogen cyanide biosynthesis by the global activator GacA in Pseudomonas aeruginosa PAO1. FEMS Microbiol Lett 200:73-78.

127. Reimmann C, Beyeler M, Latifi A, Winteler H, Foglino M, Lazdunski A, Haas D. 1997. The global activator GacA of Pseudomonas aeruginosa PAO positively controls the production of the autoinducer N-butyryl-homoserine lactone and the formation of the virulence factors pyocyanin, cyanide, and lipase. Mol Microbiol 24:309-319.

128. Heeb S, Blumer C, Haas D. 2002. Regulatory RNA as mediator in GacA/RsmA-dependent global control of exoproduct formation in Pseudomonas fluorescens CHA0. J Bacteriol 184:1046-1056.

129. Argaman L, Hershberg R, Vogel J, Bejerano G, Wagner EG, Margalit H, Altuvia S. 2001. Novel small RNA-encoding genes in the intergenic regions of Escherichia coli. Curr Biol 11:941-950.

130. Wassarman KM. 2002. Small RNAs in bacteria: diverse regulators of gene expression in response to environmental changes. Cell 109:141-144.

131. Storz G. 2002. An expanding universe of noncoding RNAs. Science 296:1260-1263.

132. Wagner EG, Altuvia S, Romby P. 2002. Antisense RNAs in bacteria and their genetic elements. Adv Genet 46:361-398.

133. Bassler BL. 2002. Small talk. Cell-to-cell communication in bacteria. Cell 109:421-424.

134. Fuqua WC, Winans SC, Greenberg EP. 1994. Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. J Bacteriol 176:269-275.

135. Daneshvar A, Truelstrup H. 2013. Kinetics of biofilm formation and desiccation survival of Listeria monocytogenes in single and dual species biofilms with Pseudomonas fluorescens, Serratia proteamaculans or Shewanella baltica on food-grade stainless steel surfaces. Biofouling 29:1253-1268.

136. Ude S, Arnold DL, Moon CD, Timms-Wilson T, Spiers AJ. 2006. Biofilm formation and cellulose expression among diverse environmental Pseudomonas isolates. Environ Microbiol 8:1997-2011.

137. Bloemberg GV, Lugtenberg BJ. 2001. Molecular basis of plant growth promotion and biocontrol by rhizobacteria. Curr Opin Plant Biol 4:343-350.

138. Whitehead NA, Barnard AM, Slater H, Simpson NJ, Salmond GP. 2001. Quorum-sensing in Gram-negative bacteria. FEMS Microbiol Rev 25:365-404.

139. El-Sayed AK, Hothersall J, Thomas CM. 2001. Quorum-sensing-dependent regulation of biosynthesis of the polyketide antibiotic mupirocin in Pseudomonas fluorescens NCIMB 10586. Microbiology 147:2127-2139.

140. Laue BE, Jiang Y, Chhabra SR, Jacob S, Stewart GS, Hardman A, Downie JA, O'Gara F, Williams P. 2000. The biocontrol strain Pseudomonas fluorescens F113 produces the Rhizobium small bacteriocin, N-(3-hydroxy-7-cis-tetradecenoyl)homoserine lactone, via HdtS, a putative novel N-acylhomoserine lactone synthase. Microbiology 146 (Pt 10):2469-2480.

141. Navarro MV, Newell PD, Krasteva PV, Chatterjee D, Madden DR, O'Toole GA, Sondermann H. 2011. Structural basis for c-di-GMP-mediated inside-out signaling controlling periplasmic proteolysis. PLoS Biol 9:e1000588.

142. Newell PD, Yoshioka S, Hvorecny KL, Monds RD, O'Toole GA. 2011. Systematic analysis of diguanylate cyclases that promote biofilm formation by Pseudomonas fluorescens Pf0-1. J Bacteriol 193:4685-4698.

143. Newell PD, Boyd CD, Sondermann H, O'Toole GA. 2011. A c-di-GMP effector system controls cell adhesion by inside-out signaling and surface protein cleavage. PLoS Biol 9:e1000587.

144. Rosqvist R, Magnusson KE, Wolf-Watz H. 1994. Target cell contact triggers expression and polarized transfer of Yersinia YopE cytotoxin into mammalian cells. EMBO J 13:964-972.

145. Sory MP, Cornelis GR. 1994. Translocation of a hybrid YopE-adenylate cyclase from Yersinia enterocolitica into HeLa cells. Mol Microbiol 14:583-594.

146. He SY, Nomura K, Whittam TS. 2004. Type III protein secretion mechanism in mammalian and plant pathogens. Biochim Biophys Acta 1694:181-206.

147. Lindgren PB, Peet RC, Panopoulos NJ. 1986. Gene cluster of Pseudomonas syringae pv. "phaseolicola" controls pathogenicity of bean plants and hypersensitivity of nonhost plants. J Bacteriol 168:512-522.

148. Gopalan S, Wei W, He SY. 1996. hrp gene-dependent induction of hin1: a plant gene activated rapidly by both harpins and the avrPto gene-mediated signal. Plant J 10:591-600.

149. Tang X, Frederick RD, Zhou J, Halterman DA, Jia Y, Martin GB. 1996. Initiation of Plant Disease Resistance by Physical Interaction of AvrPto and Pto Kinase. Science 274:2060-2063.

150. Scofield SR, Tobias CM, Rathjen JP, Chang JH, Lavelle DT, Michelmore RW, Staskawicz BJ. 1996. Molecular Basis of Gene-for-Gene Specificity in Bacterial Speck Disease of Tomato. Science 274:2063-2065.

151. Leister RT, Ausubel FM, Katagiri F. 1996. Molecular recognition of pathogen attack occurs inside of plant cells in plant disease resistance specified by the Arabidopsis genes RPS2 and RPM1. Proc Natl Acad Sci U S A 93:15497-15502.

152. Ma Q, Zhai Y, Schneider JC, Ramseier TM, Saier MH, Jr. 2003. Protein secretion systems of Pseudomonas aeruginosa and P fluorescens. Biochim Biophys Acta 1611:223-233.

153. Rainey PB. 1999. Adaptation of Pseudomonas fluorescens to the plant rhizosphere. Environ Microbiol 1:243-257.

154. Jackson RW, Preston GM, Rainey PB. 2005. Genetic characterization of Pseudomonas fluorescens SBW25 rsp gene expression in the phytosphere and in vitro. J Bacteriol 187:8477-8488.

155. Vinatzer BA, Jelenska J, Greenberg JT. 2005. Bioinformatics correctly identifies many type III secretion substrates in the plant pathogen Pseudomonas syringae and the biocontrol isolate P. fluorescens SBW25. Mol Plant Microbe Interact 18:877-888.

156. Engel J, Balachandran P. 2009. Role of Pseudomonas aeruginosa type III effectors in disease. Curr Opin Microbiol 12:61-66.

157. Shanahan P, O'Sullivan D J, Simpson P, Glennon JD, O'Gara F. 1992. Isolation of 2,4-diacetylphloroglucinol from a fluorescent pseudomonad and investigation of physiological parameters influencing its production. Appl Environ Microbiol 58:353-358.

158. Cronin D, Moenne-Loccoz Y, Fenton A, Dunne C, Dowling DN, O'Gara F. 1997. Role of 2,4-Diacetylphloroglucinol in the Interactions of the Biocontrol Pseudomonad Strain F113 with the Potato Cyst Nematode Globodera rostochiensis. Appl Environ Microbiol 63:1357-1361.

159. Barahona E, Navazo A, Martinez-Granero F, Zea-Bonilla T, Perez-Jimenez RM, Martin M, Rivilla R. 2011. Pseudomonas fluorescens F113 mutant with enhanced competitive colonization ability and improved biocontrol activity against fungal root pathogens. Appl Environ Microbiol 77:5412-5419.

160. Fenton AM, Stephens PM, Crowley J, O'Callaghan M, O'Gara F. 1992. Exploitation of gene(s) involved in 2,4-diacetylphloroglucinol biosynthesis to confer a new biocontrol capability to a Pseudomonas strain. Appl Environ Microbiol 58:3873-3878.

161. Pernthaler J. 2005. Predation on prokaryotes in the water column and its ecological implications. Nat Rev Microbiol 3:537-546.

162. Neidig N, Paul RJ, Scheu S, Jousset A. 2011. Secondary metabolites of Pseudomonas fluorescens CHA0 drive complex non-trophic interactions with bacterivorous nematodes. Microb Ecol 61:853-859.

163. Chauhan A, Layton AC, Williams DE, Smartt AE, Ripp S, Karpinets TV, Brown SD, Sayler GS. 2011. Draft genome sequence of the polycyclic aromatic hydrocarbon-

degrading, genetically engineered bioluminescent bioreporter Pseudomonas fluorescens HK44. J Bacteriol 193:5009-5010.

164. Madi A, Lakhdari O, Blottiere HM, Guyard-Nicodeme M, Le Roux K, Groboillot A, Svinareff P, Dore J, Orange N, Feuilloley MG, Connil N. 2010. The clinical Pseudomonas fluorescens MFN1032 strain exerts a cytotoxic effect on epithelial intestinal cells and induces Interleukin-8 via the AP-1 signaling pathway. BMC Microbiol 10:215.

165. Choi HJ, Seo CH, Park SH, Yang H, Do KH, Kim J, Kim HK, Chung DH, Ahn JH, Moon Y. 2011. Involvement of epidermal growth factor receptor-linked signaling responses in Pseudomonas fluorescens-infected alveolar epithelial cells. Infect Immun 79:1998-2005.

166. Hirvonen MR, Huttunen K, Roponen M. 2005. Bacterial strains from moldy buildings are highly potent inducers of inflammatory and cytotoxic effects. Indoor Air 15 Suppl 9:65-70.

167. Richard A, Rossignol G, Comet JP, Bernot G, Guespin-Michel J, Merieau A. 2012. Boolean models of biosurfactants production in Pseudomonas fluorescens. PLoS One 7:e24651.

168. Rossignol G, Merieau A, Guerillon J, Veron W, Lesouhaitier O, Feuilloley MG, Orange N. 2008. Involvement of a phospholipase C in the hemolytic activity of a clinical strain of Pseudomonas fluorescens. BMC Microbiol 8:189.

169. Sperandio D, Rossignol G, Guerillon J, Connil N, Orange N, Feuilloley MG, Merieau A. 2010. Cell-associated hemolysis activity in the clinical strain of Pseudomonas fluorescens MFN1032. BMC Microbiol 10:124.

170. Dacheux D, Goure J, Chabert J, Usson Y, Attree I. 2001. Pore-forming activity of type III system-secreted proteins leads to oncosis of Pseudomonas aeruginosa-infected macrophages. Mol Microbiol 40:76-85.

171. Raaijmakers JM, De Bruijn I, Nybroe O, Ongena M. 2010. Natural functions of lipopeptides from Bacillus and Pseudomonas: more than surfactants and antibiotics. FEMS Microbiol Rev 34:1037-1062.

172. Hallet B. 2001. Playing Dr Jekyll and Mr Hyde: combined mechanisms of phase variation in bacteria. Curr Opin Microbiol 4:570-581.

173. Hogardt M, Heesemann J. 2013. Microevolution of Pseudomonas aeruginosa to a chronic pathogen of the cystic fibrosis lung. Curr Top Microbiol Immunol 358:91-118.

174. Anderson M, Davey R. 1994. Pseudobacteraemia with Pseudomonas fluorescens. Med J Aust 160:233-234.

175. Collignon P, Dreimanis D, Beckingham W. 1999. Pseudobacteraemia due to Pseudomonas fluorescens. J Hosp Infect 43:321-322.

176. Gottlieb T, Funnell G, Gosbell I. 1991. Pseudomonas fluorescens pseudobacteraemia. Med J Aust 155:854-855.

177. Namnyak S, Hussain S, Davalle J, Roker K, Strickland M. 1999. Contaminated lithium heparin bottles as a source of pseudobacteraemia due to Pseudomonas fluorescens. J Hosp Infect 41:23-28.

178. Siebor E, Llanes C, Lafon I, Ogier-Desserrey A, Duez JM, Pechinot A, Caillot D, Grandjean M, Sixt N, Neuwirth C. 2007. Presumed pseudobacteremia outbreak resulting from contamination of proportional disinfectant dispenser. Eur J Clin Microbiol Infect Dis 26:195-198.

179. Simor AE, Ricci J, Lau A, Bannatyne RM, Ford-Jones L. 1985. Pseudobacteremia due to Pseudomonas fluorescens. Pediatr Infect Dis 4:508-512.

180. Smith J, Ashhurst-Smith C, Norton R. 2002. Pseudomonas fluorescens pseudobacteraemia: a cautionary lesson. J Paediatr Child Health 38:63-65.

181. Whyte A, Lafong C, Malone J, Golda BP. 1999. Contaminated lithium heparin bottles as a source of pseudobacteraemia. J Hosp Infect 42:342-343.

182. Adeyemi AI, Sulaiman AA, Solomon BB, Chinedu OA, Victor IA. 2010. Bacterial bloodstream infections in HIV-infected adults attending a Lagos teaching hospital. J Health Popul Nutr 28:318-326.

183. Bompard Y, Lambert T, Gantzer A, Chastel A, Voinnesson A, Aufrant C. 1988. [Use of imipenem-cilastatin in neonatal septicemias caused by gram-negative bacilli multiresistant to beta-lactam antibiotics]. Pathol Biol (Paris) 36:521-524.

184. Popov DA, Vostrikova T. 2011. [The first experience of application of PCR techniques in real-time mode to diagnose bacteriemia during postoperational period in cardiosurgery patients]. Klin Lab Diagn:49-52.

185. Rais-Bahrami K, Platt P, Naqvi M. 1990. Neonatal pseudomonas sepsis: even early diagnosis is too late. Clin Pediatr (Phila) 29:444.

186. Rodriguez Ramirez J, Pena Quintana P, Cabrera Galvan JJ, Gomez Diaz J. 1989. [Severe intrahepatic cholestasis in sepsis caused by Pseudomonas fluorescens]. Rev Clin Esp 185:106-107.

187. Shah SS, Kagen J, Lautenbach E, Bilker WB, Matro J, Dominguez TE, Tabbutt S, Gaynor JW, Bell LM. 2007. Bloodstream infections after median sternotomy at a children's hospital. J Thorac Cardiovasc Surg 133:435-440.

188. Sutter VL. 1968. Identification of Pseudomonas species isolated from hospital environment and human sources. Appl Microbiol 16:1532-1538.

189. Brook I. 1987. Microbiology of human and animal bite wounds in children. Pediatr Infect Dis J 6:29-32.

190. Dalamaga M, Karmaniolas K, Chavelas C, Liatis S, Matekovits H, Migdalis I. 2005. Pseudomonas fluorescens cutaneous abscess and recurrent bacteremia following a dog bite. Int J Dermatol 44:347-349.

191. Carpenter EM, Dicks D. 1982. Isolation of Pseudomonas fluorescens after suprapubic catheterisation. J Clin Pathol 35:581.

192. Pappas G, Karavasilis V, Christou L, Tsianos EV. 2006. Pseudomonas fluorescens infections in clinical practice. Scand J Infect Dis 38:68-70.

193. Rutenburg AM, Koota GM, Schweinburg FB. 1958. The efficacy of kanamycin in the treatment of surgical infections. Ann N Y Acad Sci 76:348-362.

194. Redding PJ, McWalter PW. 1980. Pseudomonas fluorescens cross-infection due to contaminated humidifier water. Br Med J 281:275.

195. Thangkhiew I. 1986. Successful treatment with ceftazidime of a Pseudomonas fluorescens chest infection in a myasthenic patient. J Antimicrob Chemother 18:428-429.

196. Zervos M, Nelson M. 1998. Cefepime versus ceftriaxone for empiric treatment of hospitalized patients with community-acquired pneumonia. The Cefepime Study Group. Antimicrob Agents Chemother 42:729-733.

197. Bahrani-Mougeot FK, Paster BJ, Coleman S, Barbuto S, Brennan MT, Noll J, Kennedy T, Fox PC, Lockhart PB. 2007. Molecular analysis of oral and respiratory bacterial species associated with ventilator-associated pneumonia. J Clin Microbiol 45:1588-1593.

198. Klinger JD, Thomassen MJ. 1985. Occurrence and antimicrobial susceptibility of gram-negative nonfermentative bacilli in cystic fibrosis patients. Diagn Microbiol Infect Dis 3:149-158.

199. Sutton CL, Kim J, Yamane A, Dalwadi H, Wei B, Landers C, Targan SR, Braun J. 2000. Identification of a novel bacterial sequence associated with Crohn's disease. Gastroenterology 119:23-31.

200. Dalwadi H, Wei B, Kronenberg M, Sutton CL, Braun J. 2001. The Crohn's disease-associated bacterial protein I2 is a novel enteric t cell superantigen. Immunity 15:149-158.

201. Fukata M, Michelsen KS, Eri R, Thomas LS, Hu B, Lukasek K, Nast CC, Lechago J, Xu R, Naiki Y, Soliman A, Arditi M, Abreu MT. 2005. Toll-like receptor-4 is required for

intestinal response to epithelial injury and limiting bacterial translocation in a murine model of acute colitis. Am J Physiol Gastrointest Liver Physiol 288:G1055-1065.

202. Wei B, Huang T, Dalwadi H, Sutton CL, Bruckner D, Braun J. 2002. Pseudomonas fluorescens encodes the Crohn's disease-associated I2 sequence and T-cell superantigen. Infect Immun 70:6567-6575.

203. Ashorn S, Raukola H, Valineva T, Ashorn M, Wei B, Braun J, Rantala I, Kaukinen K, Luukkaala T, Collin P, Maki M, Iltanen S. 2008. Elevated serum anti-Saccharomyces cerevisiae, anti-I2 and anti-OmpW antibody levels in patients with suspicion of celiac disease. J Clin Immunol 28:486-494.

204. Ashorn S, Valineva T, Kaukinen K, Ashorn M, Braun J, Raukola H, Rantala I, Collin P, Maki M, Luukkaala T, Iltanen S. 2009. Serological responses to microbial antigens in celiac disease patients during a gluten-free diet. J Clin Immunol 29:190-195.

205. Mundwiler ML, Mei L, Landers CJ, Reveille JD, Targan S, Weisman MH. 2009. Inflammatory bowel disease serologies in ankylosing spondylitis patients: a pilot study. Arthritis Res Ther 11:R177.

206. Yu JE, De Ravin SS, Uzel G, Landers C, Targan S, Malech HL, Holland SM, Cao W, Harpaz N, Mayer L, Cunningham-Rundles C. 2011. High levels of Crohn's disease-associated anti-microbial antibodies are present and independent of colitis in chronic granulomatous disease. Clin Immunol 138:14-22.

207. Erb-Downward JR, Thompson DL, Han MK, Freeman CM, McCloskey L, Schmidt LA, Young VB, Toews GB, Curtis JL, Sundaram B, Martinez FJ, Huffnagle GB. 2011. Analysis of the lung microbiome in the "healthy" smoker and in COPD. PLoS One 6:e16384.

Chapter II

Materials and Methods

Isolation, storage and 16S typing of clinical *P. fluorescens* strains

Isolation of *Pseudomonas* strains from clinical samples was performed between 1999 and 2009 from thirty-four treatment centers across the United States. The majority were isolated from sputum from Cystic Fibrosis patients; others from undetermined respiratory illnesses. All isolates were banked and stored at -80°C. The universal primer set 8F and 1492F was used to amplify a portion of the 16S rRNA gene through PCR, which was then Sanger-sequenced using an ABI 3730XL sequencer (1). Partial 16S sequences were screened through NCBI's nonredundant nucleotide database and the isolates were given a primarily species identification. Individual strains were either streaked out on LB agar for phenotypic assays or grown up aerobically overnight in Luria Broth at 34°C DNA isolation.

Phenotypic assays

To determine the temperature range for growth, a single colony was selected from an agar plate and grown in Luria Broth aerobically for 18 hours at room temperature (~22°C), 32°C and 37°C. The optical density at 600 nm was measured and compared to non-inoculated broth.

Plates were incubated at 34°C for twenty-four hours and colony form-units were counted and used to calculate doubling time.

All phenotypic assays were performed from a single colony grown up from an -80°C glycerol stock maintained in the lab. Fluorescence was assayed using an UV light. Standard gram-stain was used to determine whether each strain was gram-negative or positive. Motility was assayed by looking for movement on a wet mount under 100x magnification.

DNA Isolation

Genomic DNA was extracted from 2 mL of overnight culture growth using the DNeasy Blood and Tissue Kit (Qiagen Cat# 69506) following manufacturer's recommendations.

Genome sequencing, assembly and annotation

Whole genome sequencing was performed at the University of Michigan Sequencing Core on the Illumina HiSeq 2000 platform with a 100-bp paired-end library. Illumina reads were assembled *de novo* in contigs with the DNAstar SeqMan NGen Version 12 software. For the genomes we selected for deposal onto NCBI's Whole Genome online database, the contigs were then ordered with the Mauve aligner using a previously sequenced *Pseudomonas* genome as a reference. Annotation was performed using the automated online software RAST (2-4).

Nucleotide sequence accession numbers.

Accession numbers of the strains whose draft genomes that were deposited at DDBJ/EMBL/GenBank. The accession numbers for the draft genomes of subclade I isolates are LCZB00000000, LCZC00000000, LDET00000000 AU11706, AU13852 and AU20219, respectively; for subclade II isolates are LCZD00000000, LCZE00000000 for subclade II isolates AU5633 and AU11114, respectively; for subclade III isolates are JRXT00000000,

JRXU00000000, JRXV00000000, JRXW00000000, JRXX00000000, JRYA00000000, JRXY00000000, for isolates AU2989, AU6026, AU10973, AU11518, AU14440, AU14705, AU14917, respectively and for subclade IV isolates are LCYR00000000, LCYS00000000, LCYT00000000, LCYW00000000, LCYX00000000, LCYY00000000, LCYZ0000000, LCZA00000000 for isolates AU1044, AU2390, AU7350, AU10414, AU11122, AU11164, AU11235, AU11136, AU12597 and AU12644, respectively. The subclade IV isolates are currently in the process of being deposited and are not yet available for public viewing.

Taxonomic classification

Each newly sequenced clinical strain was screened for the presence of the eight housekeeping genes, DnaE, PpsA, RecA, RpoB, GuaA, MutL, PyrC and AcsA. For each housekeeping gene, 10 and 20 homologous gene sequences from other gammaproteobacteria were downloaded from the public archive at Ensembl Bacteria (http://bacteria.ensembl.org/index.html) (5). For example, the GyrB coding domain from the following 16 bacterial strains was collected: *Bacteroides vulgatus* PC510; *Bacteroides ovatus* ATCC 8483; *Bacteroides coprocola* DSM 17136; *Escherichia coli* UMNK88; *Escherichia coli* DEC5E; *Escherichia coil* IAI1; *Streptococcus pyogenes* MGAS15252; *Streptococcus pyogenes* MGAS1882; *Streptococcus pneumonia* 2090008; *Pseudomonas fluorescens* F113; *Pseudomonas putida* BIRD-1; *Pseudomonas fluorescens* SS101; *Pseudomonas aeruginosa* LESB58; *Pseudomonas aeruginosa* PAO579; *Pseudomonas putida* and *Pseudomonas putida* HB3267. A large fasta file was created for each housekeeping gene that contained the homologous gene sequences downloaded from Ensembl Bacteria and this large fasta file was used to screen a locally created database that contained all 60 newly sequenced clinical *Pseudomonas spp.* genomes using the local megablast function (6). For genomes with a significant hit to a query housekeeping gene, the nucleotide sequence of the housekeeping was located and saved for later phylogenetic tree creation.

Multi-locus sequence analysis (MLSA) was performed using between six and eight concatenated sequences of the following housekeeping genes: DnaE, PpsA, RecA, RpoB, GuaA, MutL, PyrC and AcsA; an approach modified from the MLSA approach used in Loper *et al* (7). Nucleotide differences along the concatenated sequences of these housekeeping genes was used to infer a phylogenetic tree provide species and subclade level taxonomic classification of the clinical *Pseudomonas* isolates. Using this MLSA approach, the following *Pseudomonas* strains were determined to belong to the *P. fluorescens*-species complex and its subsequence subclades.

For OTUs of interest in the mouse and human datasets, we obtained the consensus sequence that represented the 97% shared nucleotide sequence within that OTU for further taxonomic classification. These OTU-representative (OTU-rep) sequences were used for species- level classification utilizing three approaches; a V3-V5 16S rRNA phylogenetic tree, a BLAST screen against NCBI's microbial nucleotide database and alignment against select *Pseudomonas* genomes.

The V3-V5 16S rRNA phylogenetic tree was used to evaluate a likely species-level classification for the *Pseudomonas* OTUs from the mouse and human lung microbiota datasets and the clinical *Pseudomonas* isolates. Reference, previously sequenced *Pseudomonas* strains were also included in this analysis. The primer set 357F and 926R was used to locate the V3-V5 hypervariable region of the 16S rRNA gene in newly sequenced and reference *Pseudomonas spp*. genomes (8). MAFFT was used to align the concatenated housekeeping genes for the MLSA tree and the sequences for the V3-V5 16S rRNA 16S gene tree (9, 10).

The BLAST screen against NCBI's microbial nucleotide database was performed using the OTU-representative sequence of each *Pseudomonas* OTU. For the alignment against select *Pseudomonas* genomes, NCBI's 'bl2seq' function was used, with the OTU-rep sequence of each *Pseudomonas* OTU as the query and the full genome of the representative *Pseudomonas* strain as the subject.

Analysis of individual genes and gene clusters

The presence or absence of secondary metabolite genes was assayed through the blastn function provided by NCBI using a chosen query sequence listed in **Appendix Figure IV.4** and a cutoff of evalue <1E15 and sequence identity >70% (11, 12). The nucleotide sequences of the *czc*A homologues were discovered in RAST and used to infer a phylogenetic tree using MAFFT (3, 4, 9, 10). To find the amino acid sequence of p*Fit* from newly sequenced subclade I, II and III *P. fluorescens* strains we utilized NBCI's blastp function and screen each genome with the amino acid sequence from *P. fluorescens* A506 as a query. The amino acid sequence of *pfi*T from the subclade I, II and III strains was then aligned using MUSCLE (13). The nucleotide sequences of FimA and VasH were discovered in RAST (3, 4). Alignment of the FimA nucleotide acid homologues was performed using the MAFFT algorithm (9, 10). The nucleotide sequences of VasH homologues were aligned in Mauve (14). The *pfi*T, FimA and VasH alignments were visualized using DNAstar's MegaAlignPro software.

Whole Genome Analysis

Pan, accessory, soft core and core genomes were calculated using the COGtriangle clustering in GET_HOMOLOGUES (15). Shared average nucleotide identity (ANI) was determined using the online software at http://enve-omics.ce.gatech.edu/ani (16). A similarity matrix of the ANI between the subclade III strains was created and clustering performed using the Unweighted Pair Group Method with Arithmetic mean (UPGMA) algorithm.

Animal Models

Male Bl/6 mice between the ages of 8 and 16 weeks were either sourced from an on-site mouse colony at the University of Michigan or were purchased from The Jackson Laboratory company. Mice were either untreated or received intranasal inoculations of 20 ul of 1 x 10^8

Aspergilllus fumigatus spores/ml or PBS for four or eight weeks. From the on-site sourced mice the following number of mice was in each treatment group: Untreated Control Lungs (n-13); PBS Control 4 Week Lungs (n=8); Inflamed 4 Week Lungs (n-11); PBS Control 8 Week Lungs (n=6); Inflamed 8 Week Lungs (n=9). For the Jackson sourced mice the following number of mice was in each treatment group: Untreated Control Lungs (n=7); PBS Control 4 Weeks Lungs (n=6); Inflamed 4 Week Lungs (n=5); PBS Control 8 Week Lungs (n=6); Inflamed 8 Week Lungs (n=3). For each mouse group, lung and tongue snips were collected for whole CFU plating, DNA isolation and 16S rRNA amplification.

Tissue Harvest and Sample Preparation

Lungs and tongues were sterilely removed and placed in 15 mL polystyrene tubes on ice. 2 mL of sterile PBS was added to each sample tube and homogenized. The homogenizer was washed and cleaned with ethanol in between samples. Samples for CFU analysis were immediately plated while samples for DNA analysis were transferred to 1.5mL centrifuge tubes, flash frozen in liquid nitrogen and stored at -80°C.

CFU plating

Lungs homogenates were serially diluted between 10^{-1} and 10^{-10}. Straight homogenate and all dilutions were plated in triplicate on LB agar and incubated aerobically at 37°C for 1-2 days.

Quantitative PCR

Bacterial 16S rDNA copies were quantified using real-time PCR utilizing TaqMan hydrolysis probes. Primers were targeted to the V1-V2 regions of the 16S rDNA gene using the following sequences: 5'-AGAGTTTGATCCTGGCTCAG-3' and 5'-CTGCTGCCTYCCGTA-3'. The probe was developed for the primer BSR65/17 landing site using the following sequence: 5´-TCGACTTGCATGTRTTA-3´. 16S clones derived from a *Haemophilus* species were used for generation of a standard curve. After an initial denaturation of 5 minutes at 95°C, 40 cycles of amplification were performed: 30 seconds at 94°C, 30 seconds at 50°C and 30 seconds at 72°C. A final elongation step was performed at 72°C.

454 Pyrosequencing

The bar-coded primer sets corresponding to 357F and 929R, designed by the Broad Institute were used to amplify the V3-V5 hypervariable regions of the bacterial 16S rRNA gene (17). Primary PCR cycling conditions were 95°C for 2 min., followed by 30 cycles of (95°C 20 sec, 50°C 30 sec, 72°C 5 min), holding at 4°C. Quality control and sequencing was carried out at the University of Michigan Pyrosequencing Core, using the Roche 454 GS Junior according to established protocols (18). Water control and mock community specimens were analyzed with each sequencing run as quality controls.

Data analysis

Sequence data were processed and analyzed with the software mothur v.1.27.0 according to the Schloss 454 SOP (http://www.mothur.org). with OTUs binned at 97% identity. All bacterial signal detected in water is subtracted from all other samples before analysis. Only OTUs that were greater than 1% of the sample population is kept for rank abundance and ordination techniques. All OTUs were included in diversity and microbial regression analysis. For ordination and rank abundance analysis, samples were normalized to the percent of total reads. All files were imported and analyzed in R using the R-package vegan 2.0-4 for diversity

analyses and ordinations, and a custom R script for sorting classification results into tables (ClassifierSorter v.2). Classification of OTUs was carried out using the mothur implementation of the RDP Classifier and the RDP taxonomy training set (http:// rdp.cme.msu.edu). *Pseudomonas aeruginosa* and *Pseudomonas fluorescens* were discriminated from each other using the NCBI Basic Local Alignment Search Tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Statistical analysis

Statistical analyses were performed using Prism 5 (GraphPad Software) for ANOVA, t-test and regression analysis, and vegan and R for all diversity, rank abundance and ordination analyses.

Chapter II

References

1. Eden PA, Schmidt TM, Blakemore RP, Pace NR. 1991. Phylogenetic analysis of Aquaspirillum magnetotacticum using polymerase chain reaction-amplified 16S rRNA-specific DNA. International journal of systematic bacteriology 41:324-325.
2. Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT. 2009. Reordering contigs of draft genomes using the Mauve aligner. Bioinformatics 25:2071-2073.
3. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, 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.
4. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res 42:D206-214.
5. Kersey PJ, Allen JE, Christensen M, Davis P, Falin LJ, Grabmueller C, Hughes DS, Humphrey J, Kerhornou A, Khobova J, Langridge N, McDowall MD, Maheswari U, Maslen G, Nuhn M, Ong CK, Paulini M, Pedro H, Toneva I, Tuli MA, Walts B, Williams G, Wilson D, Youens-Clark K, Monaco MK, Stein J, Wei X, Ware D, Bolser DM, Howe KL, Kulesha E, Lawson D, Staines DM. 2014. Ensembl Genomes 2013: scaling up access to genome-wide data. Nucleic Acids Res 42:D546-552.
6. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421.

7. Loper JE, Hassan KA, Mavrodi DV, Davis EW, 2nd, Lim CK, Shaffer BT, Elbourne LD, Stockwell VO, Hartney SL, Breakwell K, Henkels MD, Tetu SG, Rangel LI, Kidarsa TA, Wilson NL, van de Mortel JE, Song C, Blumhagen R, Radune D, Hostetler JB, Brinkac LM, Durkin AS, Kluepfel DA, Wechter WP, Anderson AJ, Kim YC, Pierson LS, 3rd, Pierson EA, Lindow SE, Kobayashi DY, Raaijmakers JM, Weller DM, Thomashow LS, Allen AE, Paulsen IT. 2012. Comparative genomics of plant-associated Pseudomonas spp.: insights into diversity and inheritance of traits involved in multitrophic interactions. PLoS genetics 8:e1002784.

8. Dickson RP, Erb-Downward JR, Freeman CM, Walker N, Scales BS, Beck JM, Martinez FJ, Curtis JL, Lama VN, Huffnagle GB. 2014. Changes in the lung microbiome following lung transplantation include the emergence of two distinct Pseudomonas species with distinct clinical associations. PloS one 9:e97214.

9. Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059-3066.

10. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular biology and evolution 30:772-780.

11. Pearson WR. 2014. BLAST and FASTA Similarity Searching for Multiple Sequence Alignment. 1079:75-101.

12. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403-410.

13. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792-1797.

14. Darling AC, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome research 14:1394-1403.

15. Contreras-Moreira B, Vinuesa P. 2013. GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. Applied and environmental microbiology 79:7696-7701.

16. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. International journal of systematic and evolutionary microbiology 57:81-91.

17. Group JCHMPDGW. 2010. Human Microbiome Consortium 16S 454 Sequencing Protocol.

18. Daigle D, Simen BB, Pochart P. 2011. High-throughput sequencing of PCR products tagged with universal primers using 454 life sciences systems. Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.] Chapter 7:Unit7 5.

Chapter III

Multi-locus Sequence Analysis (MLSA) of newly sequenced clinical *Pseudomonas spp.* isolates

BACKGROUND

Multi-locus sequence typing (MLST) and multi-locus sequence analysis (MLSA) are molecular methods for taxonomic classification based on nucleotide differences at loci in bacteria strains (1-5). Maiden et al. first developed MLST in 1998 as a portable method for the typing of pathogenic bacteria clones (2). MLST and MLSA are adapted from multilocus enzyme electrophoresis (MLEE), which has long been used to type new bacteria strains (5, 6). In MLEE, metabolic enzymes are separated via gel electrophoresis based on differences in electrostatic charge. Distinct stained profiles are created, called 'electromorphs or electrophoretic variants', and bacteria are typed based on divergent or shared electromorphs (1, 5, 6). MLEE is based on the concept that nucleotide changes in genes that encode for metabolic enzymes lead to changes in protein structure that are measurable by differences in migration rates (6). However, the MLEE approach to bacterial typing has disadvantages; it is difficult to compare between labs,

the technique is time consuming and hard to learn, and it indirectly measures genetic evolution through changes in protein structure (1, 7). MLST and MLSA were developed as attempt to get around the disadvantages of MLEE.

Measuring changes in protein structure indirectly through MLEE made sense in a time when measuring genetic diversity directly through sequencing was unaffordable and not widely available. Advancements in technology, beginning with the advent of the polymerase chain reaction, increased the accuracy and availability of genetic sequencing, making it possible to measure genetic diversity at the level single nucleotide polymorphism (SNPs) at any loci between bacteria strains (1, 8, 9). The almost universal reach of the Internet, the development of new software and bioinformatics tools, and the wide use of public nucleotide repositories (such those provided by NCBI and EMBL) allows for the newly instantaneous exchange of genetic information (10, 11). These advancements in technology have allowed nucleotide-based methods, such as MLST and MLSA, to become the most widely used methods for bacterial typing (12).

Both MLST and MLSA utilize nucleotide diversity in housekeeping genes to measure phylogenetic diversity between bacteria strains. Housekeeping genes are those that code for proteins that perform basic metabolic functions and, therefore, are under purifying selection (ie. mutations that alter the protein function are more readily purged during evolution). In a typical MLST scheme, six or more housekeeping genes are selected and differences in allele patterns, called ‘sequence types’ (STs) are compared between bacteria strains in question (1, 2, 13). In the first MLST schemes, 400-600 bp gene fragments of at least six housekeeping genes were collected from strains of interest and arbitrary numbers were assigned to the sequenced alleles for each gene (2). The nucleotide length of 400-600 bp was chosen because that was the longest reliable length produced by a single run of the sequencing machines in the mid-1990s

(1). Strains with the same pattern of alleles were then assigned the same sequence type (ST) and phylogenetic trees were created based on similarity and differences in STs (1, 2). MLST measures differences in alleles at chosen sites in a bacterial genome and not the frequency of individual nucleotide changes at those loci; MLSA differs from MLST by building phylogenies based solely on the differences in nucleotides at sequenced loci (4, 13).

While both MLSA and MLST schemes begin with the first initial step- selecting six or more housekeeping genes and collecting the sequences of those housekeeping genes from strains of interest- they differ in what information goes into the building of the subsequent phylogenetic tree. In MLST the phylogenetic tree is built based on differences in sequence type; in MLSA the phylogenetic tree is built based on differences in nucleotides along the concatenated sequence of the housekeeping genes (3, 13). While MLST is often used to delineate between individual strains within a known species, the higher discriminatory power of MLSA is typically used for bacteria strains of completely unknown bacterial species (13).

Taxonomic classification via MLSA is also more discriminatory than that by 16S rRNA gene sequence comparison. There is a higher level of genetic variability between non-16S rRNA housekeeping genes, and the longer sequence length of the concatenated genes provides more opportunities for nucleotide differences between strains (13). In addition, the use of multiple loci decreases the chance that horizontal gene transfer obscures the phylogenetic relationship between strains (13). Of the multiple different sequence-based typing methods (16S rRNA gene; MLST and MLSA), MLSA provides the highest level of discriminatory power and was thus chosen to provide taxonomic classification to the unknown *Pseudomonas spp.* strains sequenced in this study.

The taxonomy of the *P. fluorescens* species complex is currently limited and lacks any representative strains isolated from clinical samples (14). To expand the taxonomy of the *P. fluorescens* species complex, as well as the *Pseudomonas* genus, we sequenced 60 *Pseudomonas* isolates from individuals with chronic respiratory disease. We modified a previously published MLSA scheme and inferred a phylogenetic tree that included the 60 newly sequenced clinical strains as well as 21 representative previously sequenced *Pseudomonas* strains. As an internal control, we re-sequenced *P. fluorescens* SBW25 along with the clinical strains and included the in-house sequenced SBW25 in our taxonomic analysis (labeled 'UM SBW25'). The 60 newly sequenced strains were placed into various species groups within the phylogenetic tree, with 22 strains placed within the *P. fluorescens* species complex. Two additional MLSA schemes were used to more preciously delineate the individual subclades within the *P. fluorescens* species complex. The results here greatly expand the taxonomy within the *P. fluorescens* complex and provide the first taxonomic classification of *P. fluorescens* species bacteria isolated from human clinical samples.

RESULTS AND DISCUSSION

To taxonomically classify the newly sequenced *Pseudomonas spp.* clinical strains we modified a MLSA scheme previously reported to delineate the different subclades in the *P. fluorescens* species complex. In Loper *et al.,* ten, fully sequenced *P. fluorescens* species complex strain underwent taxonomic classification through a ten-gene MLSA approach (14). The main finding from this taxonomic analysis was that the ten strains formed three distinct subclades within the *P. fluorescens* species complex. Further comparative genomics revealed that strains from different subclades varied in their genomic and physical attributes. From the ten gene MLSA scheme presented in Loper *et al.,* we selected eight and screened the 60 clinical strains for a homologous sequence for each housekeeping gene, as further detailed in

the materials and methods (**Chapter II**). **Appendix Figure III.1** lists the eight genes selected and their expressed function. The results from the homologous gene search are presented in **Appendix Figure III.2.** Significant nucleotide matches to all eight of the housekeeping genes were found in a majority of the newly sequenced strains. However, in a portion of strains (14 / 60), the screen was unable to find a homologue to *pyrC* or *acsA.* The lack of a homologue to these genes is either due to the draft nature of the genomes, imprecision of the gene screen or actual gene loss; these possibilities are further discussed. From the results of the gene screen, we created a MLSA scheme that consisted of the concatenated sequences of the six housekeeping genes *dnaE, ppsA, recA, rpoB, guA* and *mutL.* The six-gene MLSA scheme was then used to infer a phylogenetic tree and taxonomically classify the 60 newly sequenced clinical strains.

Using the six-gene MLSA scheme, we inferred a phylogenetic tree that places the 60 newly sequenced clinical strains within species-level groupings in the *P. fluorescens* species complex (**Figure III.1**). Included in the tree are 21 previously sequenced and publically available *Pseudomonas* genomes that assist in discerning the different species clades within the genus. Our MLSA approach successfully placed the previously sequenced reference strains in their known taxonomic groups (14, 15). The 60 newly sequenced strains were placed into different species clades and subclades, as presented in **Appendix Figure III.1**. Twenty-two of the strains were placed within the *P. fluorescens* species complex; three in subclade I, two in subclade II, seven in subclade III and ten formed a fourth, novel subclade IV. While subclades I, II and III also contain publically available *P. fluorescens* strains isolated from environmental samples, the new fourth subclade contains only newly sequenced, clinical *P. fluorescens* strains. The six-gene MLSA scheme classified 22 of the 60 clinical strains in the *P. fluorescens* species complex and lead to the discovery a fourth, novel subclade that did not contain any representative previously sequenced environmental isolates.

The branching pattern of the *P. fluorescens* species complex points to a high level of diversity between subclade IV and the other subclades. In the phylogenetic tree (**Figure III.1**) the branch length relates to the expected average number of nucleotide substitutions per site between strains (scale bar, 0.01 nucleotide substitutions). The branch that separates subclade IV from its closest other subclade in the *P. fluorescens* species complex is 0.07 average nucleotide substitutions (**Figure III.1**). In contrast, the branches that separate subclades I, II and III are no more than 0.02 average nucleotide substitutions. The individual strains within subclade IV show fewer nucleotide differences between each other than members of subclade I, II or III. On average, subclade III strains differ from each other 0.04 nucleotide differences, with a range between 0.02 and 0.08. Subclade II strains, on average, differ from each 0.06 nucleotide differences, with a range between 0.04 and 0.12 and subclade I strains differ, on average, 0.04 nucleotide differences, with a range between 0.04 and 0.05. In contrast, subclade IV strains differ on average 0.02 nucleotides, with a range of 0.01 to 0.03. The long branch length separating subclade IV from the other subclades points to a high level of genetic dissimilarity between subclade IV and other members of the *P. fluorescens* species complex and close clustering within subclade IV suggests a high level of genetic similarity between subclade IV strains.

While a six-gene MLSA scheme was used to classify all 60 clinical isolates, modified MLSA schemes were use to specifically analyze the different *P. fluorescens* species subclades. The taxonomic tree presented in **Figure III.1** is very large and hard to read. For the additional comparative genomic analysis of the clinical strains in the *P. fluorescens* species complex we created simplified taxonomic trees that that focused on specific subclades of interest. For comparative genomic of subclade III clinical strains and environmental strains, we created an 8-

gene MLSA scheme (**Figure III.2**) and for detailed analysis of the distinct genomic attributes of subclade IV strains, we created a 7-gene MLSA scheme (**Figure III.3**).

For the comparative genomic analysis of the subclade III strains presented in **Chapter IV**, we utilized an eight-gene MLSA scheme that included the coding regions of the following housekeeping proteins: *dnaE, ppsA, recA, rpoB, guaA, mutL, pyrC* and *acsA* (**Figure III.2**). As shown in **Appendix Figure III.2**, all seven subclade III strains contained a homologue to the eight genes included in the initial screen. The eight-gene based MLSA scheme was able to successfully place previously sequenced *P. fluorescens* species complex strains into their known subclades. The trees inferred from the six-gene (**Figure III.1**) and eight-gene (**Figure III.2**) MLSA schemes agree with their placement of the seven clinically isolated strains AU2989, AU6026, AU10973, AU11518, AU14440, AU14705 and AU14917 in subclade III. The eight-gene MLSA scheme successful delineated the phylogenetic relationship of the clinical strains in subclade III and was utilized for additional comparative genomic analysis presented in **Chapter IV**.

For the focused analysis on subclade IV, we inferred a phylogenetic tree similar to the one created for subclade III analysis (**Figure III.2**), using a seven-gene MLSA scheme (**Figure III.3**). As shown in **Appendix Figure III.1,** nine out of the ten subclade IV strains lacked a homolog to the AcsA gene. While this could be due to the draft nature of their genomes, as discussed later in **Chapter V,** subclade IV strains have undergone significant gene loss and, therefore, might truly lack gene homologues to *acsA.* Extensive analysis of the genomic differences between subclade IV strains and other members of the *P. fluorescens* species complex is found in **Chapter V**. The phylogenetic tree inferred from the six-gene MLSA (**Figure III.1**) and the seven-gene MLSA (**Figure III.3**) both agree in the tight clustering of the subclade

IV strains and the long branch length that separates this subclade from the other members of the *P. fluorescens* species complex.

While the placement of the clinical strains in subclade III and subclade IV remains consistent between taxonomic schemes, the placement of clinical strains in subclades I and II varies between the six-gene scheme in **Figure III.1** and seven-gene MLSA scheme in **Figure III.3.** In the large, six-gene MLSA scheme that includes all 60 clinical strains, AU5633 and AU11114 are placed within subclade II with reference strains *P. fluorescens* Pf0-1 and R124 and strains AU11706, AU13852 and AU20219 are placed within subclade I with reference strains *P. protegens* Pf-5 and CHA0 (**Figure III.1**). However, in the seven-gene MLSA in **Figure III.3**, strains AU11114 and AU11706 are placed together on a branch within subclade I (**Figure III.3**). The seven-gene MLSA scheme includes PyrC, which is missing in the six-gene scheme (**Appendix Figure III.1**). In addition, the number of strains included in the taxonomic analysis differs between the six- and seven-gene MLSA phylogenetic trees. Both the length of the concatenated gene sequence and the number of strains analyzed can affect the final inferred phylogenetic tree. To test which is affecting the delineation of strains in subclades I and II, the seven-gene phylogenetic scheme was used to infer a phylogenetic tree that included the 56 clinical strains that contained all seven of these genes (**Appendix Figure III.1**). As shown in **Figure III.4,** when these additional strains are evaluated with the seven-gene MLSA scheme, subclades I and II change their structure again. AU11114 is once again clustered with the clinical strain AU5633 and the reference strains Pf0-1 and R124, as seen in the large, six-gene MLSA (**Figure III.1**). However, in the seven-gene MLSA scheme with 56 clinical strains (**Figure III.4**), subclade II becomes two separate, smaller subclades, with AU5633, AU11114, Pf0-1 and R124 in one and strains Q2-87 and Q8r-96 in another. The changes in membership in subclade II point to the fact that, even with the 22 *P. fluorescens* clinical strains added in this study, the taxonomy of the *P. fluorescens* species complex is still in flux. As more strains are added,

membership in subclades II and I (which currently only include 6 and 5 strains, respectively) will likely change further. However, while subclades I and II changed in membership depending on the MLSA scheme utilized and the strains included in the analysis, membership in subclades III and IV remained consistent. With the newly sequenced clinical strains added to the *Pseudomonas* genus phylogenetic tree, we were able to identify 22 newly sequenced clinical *P. fluorescens* species strains and make initial, subclade destinations that are further analyzed in the comparative genomics presented in **Chapters IV** and **V**.

CONCLUSION

Through MLSA we were able to provide an initial species-level classification to 60 newly sequenced clinical strains of *Pseudomonas.* Twenty-two of the clinical strains clustered into the *P. fluorescens* species complex, representing the first report of fully sequenced *P. fluorescens* strains isolated from clinical samples. Ten of the clinical *P. fluorescens* strains form a new fourth subclade within the *P. fluorescens* species complex. The strains within the new subclade IV cluster tightly together, suggesting they are genetically very similar to each other, while the long branch length that separates subclade IV suggests subclade IV strains are genetically dissimilar from other members of the species complex. Through the analysis provided here we demonstrate that MLSA can successfully classify the newly sequenced clinical *Pseudomonas* isolates into species and subclade groupings and that the strains within the newly formed fourth subclade are phylogenetically and genetically distinct from all other members of the *P. fluorescens* species complex.

Figure III.1: Phylogenetic tree of 60 sequenced clinical *Pseudomonas spp.* strains with six-gene MLSA scheme.

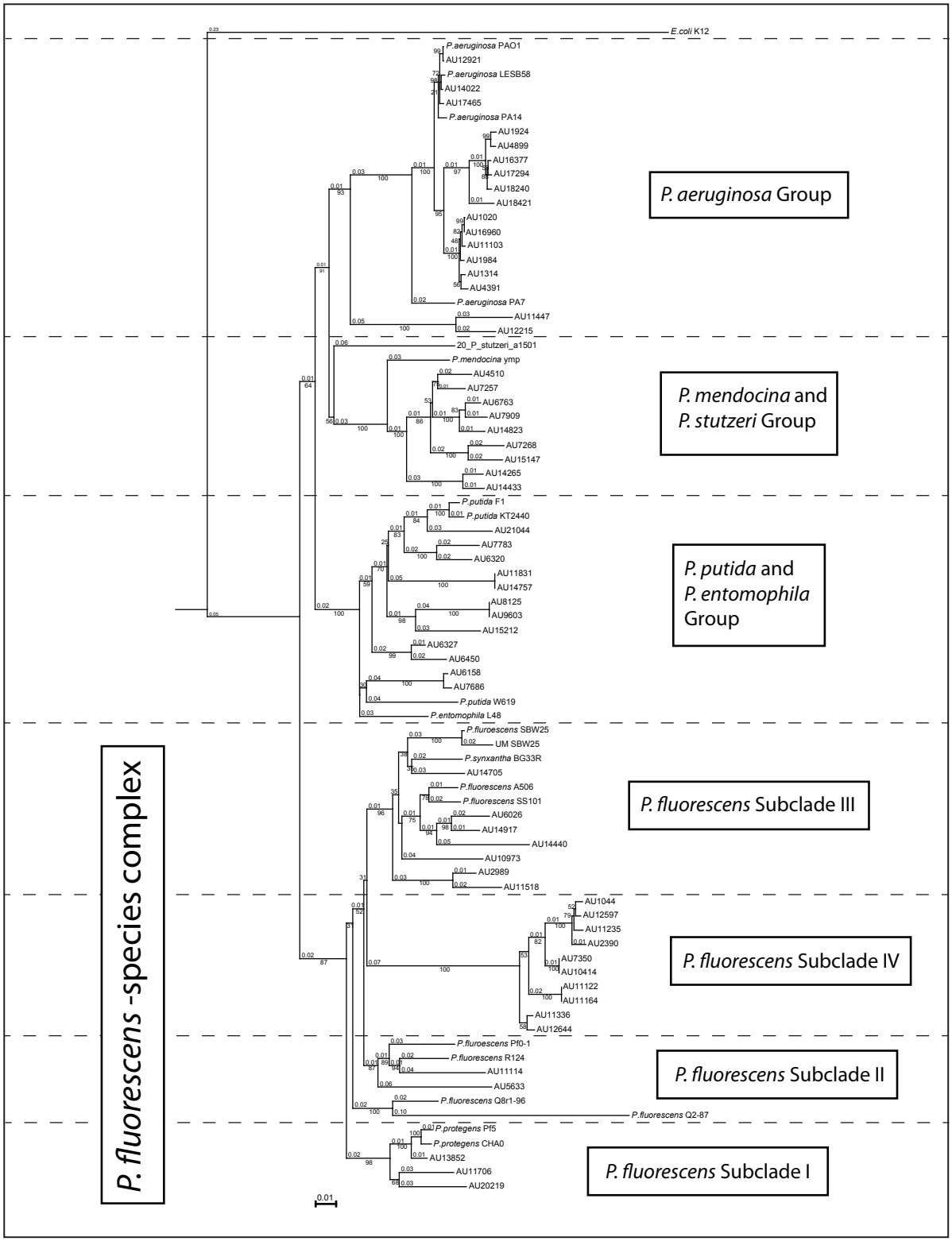


Figure III.1: Phylogenetic tree of 60 sequenced clinical *Pseudomonas spp.* strains with six-gene MLSA scheme. Phylogenetic tree was created based on the concatenated sequences of the following six housekeeping genes: DnaE (DNA polymerase III alpha subuit); PpsA (phosphoenolpyruvate synthase); RecA (Recombinase A); RpoB (RNA polymerase subunit beta); GuaA (GMP synthease) and MutL (DNA mismatch repair protein).

Figure III.2: Phylogenetic tree with eight-gene MLSA scheme to characterize *P. fluorescens* subclade III.

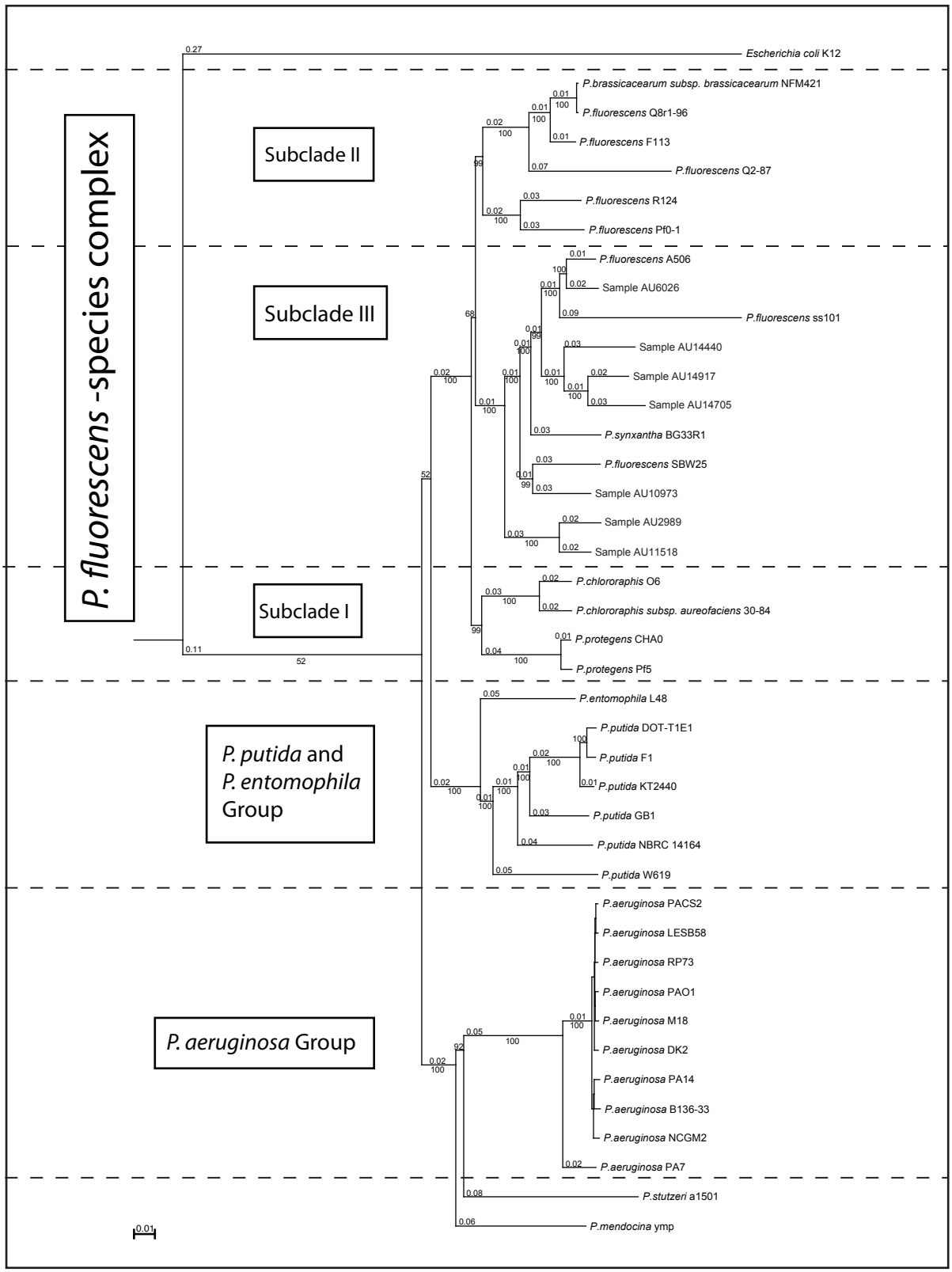


Figure III.2: Phylogenetic tree with eight-gene MLSA scheme to characterize *P. fluorescens* subclade III.
Phylogenetic tree inferred from the concatenated sequences of the following eight housekeeping genes: DnaE (DNA polymerase III alpha subuit); PpsA (phosphoenolpyruvate synthase); RecA (Recombinase A); RpoB (RNA polymerase subunit beta); GuaA (GMP synthease); MutL (DNA mismatch repair protein); PyrC (pyrimidine biosynthetic enzyme dihydroorotase) and AcsA (acetyl-CoA synthetase).

Figure III.3: Phylogenetic tree with seven-gene MLSA scheme to characterize *P. fluorescens* subclade IV.

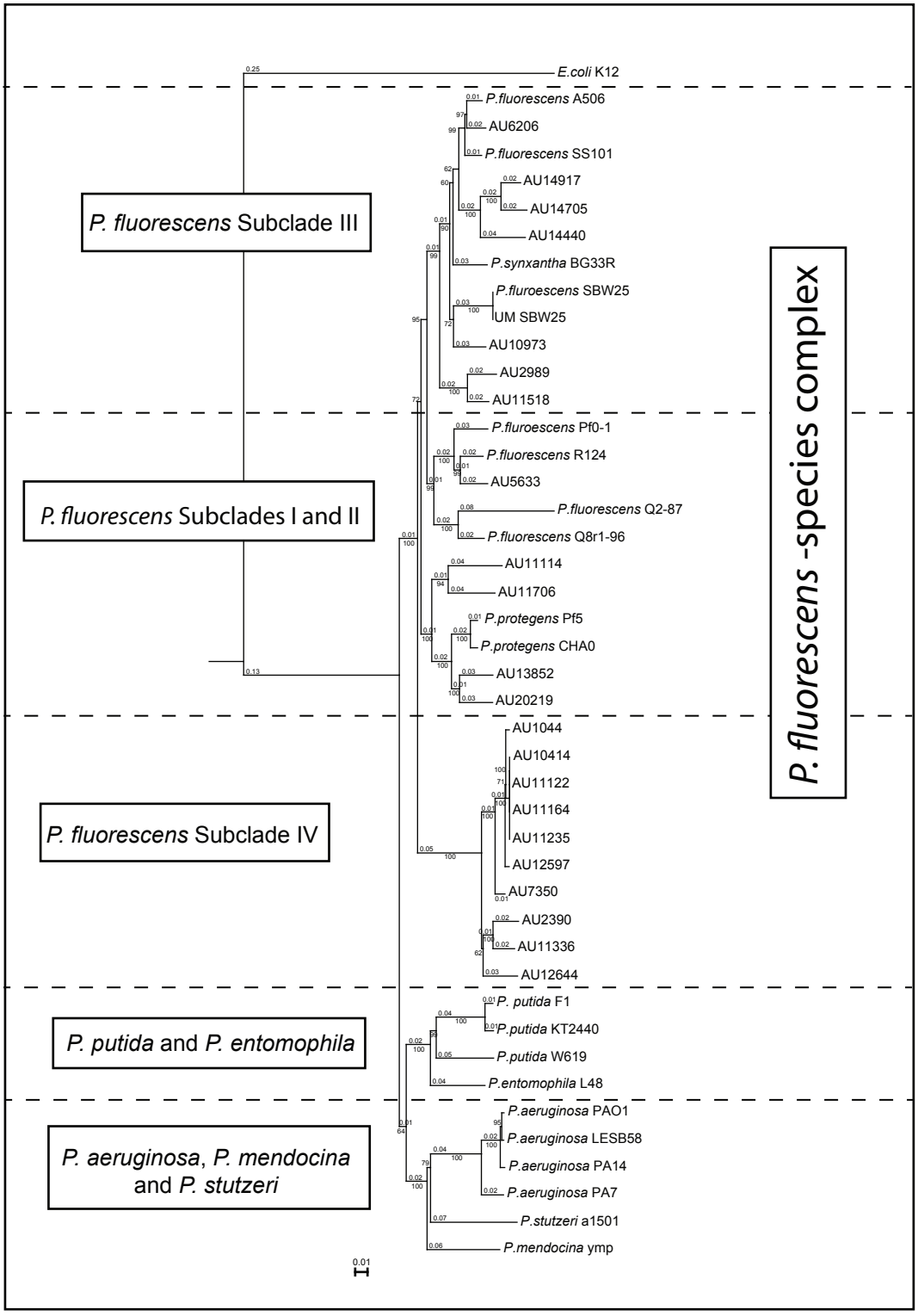


Figure III.3: Phylogenetic tree with seven-gene MLSA scheme to characterize *P. fluorescens* subclade IV.
Phylogenetic tree was created based on the concatenated sequences of the following seven housekeeping genes: DnaE (DNA polymerase III alpha subuit); PpsA (phosphoenolpyruvate synthase); RecA (Recombinase A); RpoB (RNA polymerase subunit beta); GuaA (GMP synthease), MutL (DNA mismatch repair protein) and PyrC (pyrimidine biosynthetic beta).

Figure III.4: Phylogenetic tree of 56 sequenced clinical *Pseudomonas spp.* strains with seven-gene MLSA scheme.

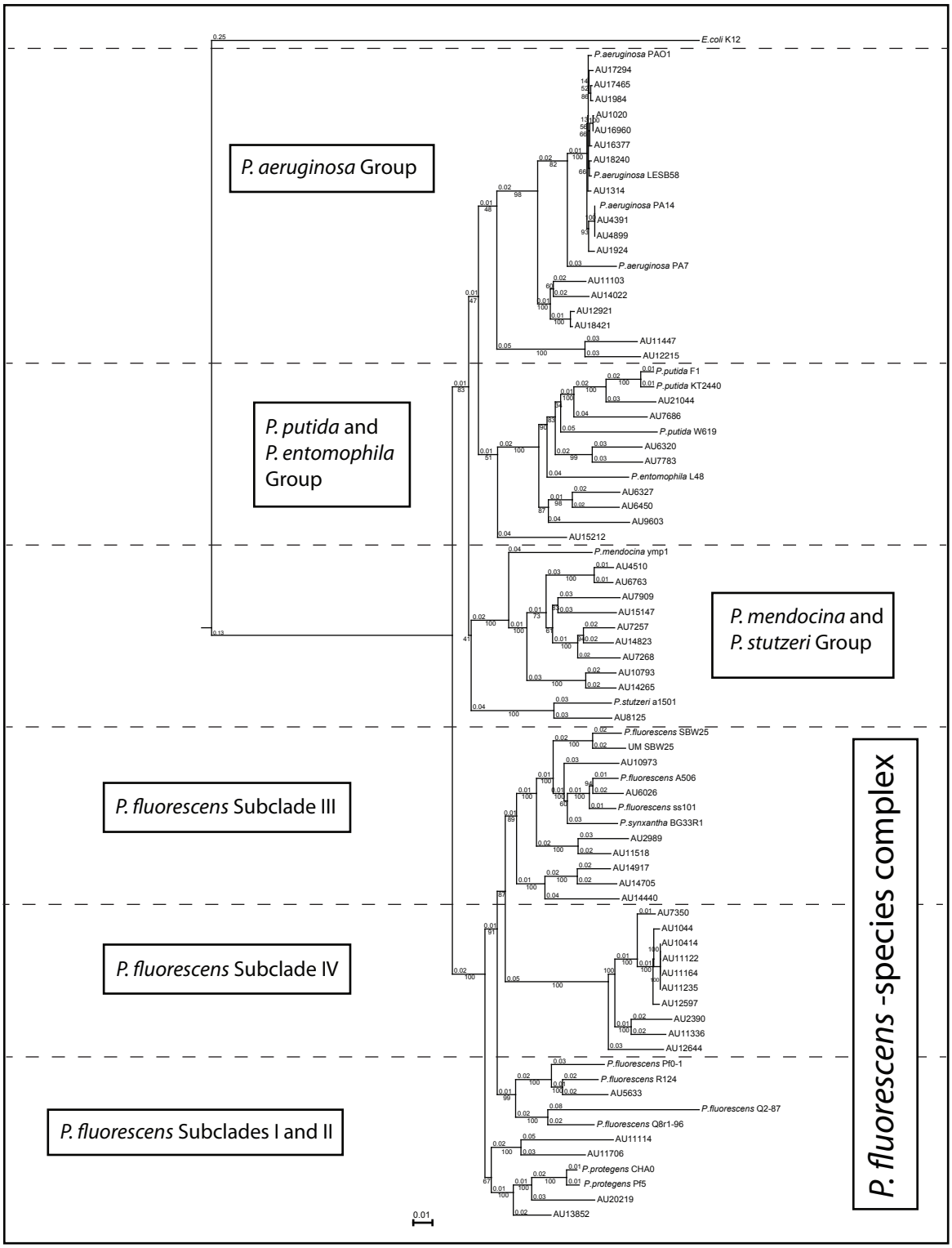


Figure III.4: Phylogenetic tree of 56 sequenced clinical *Pseudomonas spp.* strains with seven-gene MLSA scheme. Phylogenetic tree was created based on the concatenated sequences of the following seven housekeeping genes: DnaE (DNA polymerase III alpha subuit); PpsA (phosphoenolpyruvate synthase); RecA (Recombinase A); RpoB (RNA polymerase subunit beta); GuaA (GMP synthease); MutL (DNA mismatch repair protein) and AscA (acetyl-coenzyme A synthetase).

Chapter III

References

1. Maiden MC. 2006. Multilocus sequence typing of bacteria. Annu Rev Microbiol 60:561-588.
2. Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG. 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci U S A 95:3140-3145.
3. Bennasar A, Mulet M, Lalucat J, Garcia-Valdes E. 2010. PseudoMLSA: a database for multigenic sequence analysis of Pseudomonas species. BMC Microbiol 10:118.
4. Gevers D, Cohan FM, Lawrence JG, Spratt BG, Coenye T, Feil EJ, Stackebrandt E, Van de Peer Y, Vandamme P, Thompson FL, Swings J. 2005. Opinion: Re-evaluating prokaryotic species. Nat Rev Microbiol 3:733-739.
5. Urwin R, Maiden MC. 2003. Multi-locus sequence typing: a tool for global epidemiology. Trends Microbiol 11:479-487.
6. Selander RK, Caugant DA, Ochman H, Musser JM, Gilmour MN, Whittam TS. 1986. Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. Appl Environ Microbiol 51:873-884.
7. Enright MC, Spratt BG. 1999. Multilocus sequence typing. Trends Microbiol 7:482-487.
8. Maiden MC. 2000. High-throughput sequencing in the population analysis of bacterial pathogens of humans. Int J Med Microbiol 290:183-190.
9. Cole ST, Saint Girons I. 1994. Bacterial genomics. FEMS Microbiol Rev 14:139-160.
10. Chan MS, Maiden MC, Spratt BG. 2001. Database-driven multi locus sequence typing (MLST) of bacterial pathogens. Bioinformatics 17:1077-1083.

11. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL. 2008. GenBank. Nucleic Acids Res 36:D25-30.

12. Li W, Raoult D, Fournier PE. 2009. Bacterial strain typing in the genomic era. FEMS Microbiol Rev 33:892-916.

13. Vinatzer BA, Bull CT. 2009. The Impact of Genomic Approaches on our Understanding of Diversity and Taxonomy of Plant Pathogenic Bacteria *In* Jackson RW (ed.), Plant Pathogenic Bacteria: Genomics and Molecular Biology. Caister Academic Press, Reading, UK.

14. Loper JE, Hassan KA, Mavrodi DV, Davis EW, 2nd, Lim CK, Shaffer BT, Elbourne LD, Stockwell VO, Hartney SL, Breakwell K, Henkels MD, Tetu SG, Rangel LI, Kidarsa TA, Wilson NL, van de Mortel JE, Song C, Blumhagen R, Radune D, Hostetler JB, Brinkac LM, Durkin AS, Kluepfel DA, Wechter WP, Anderson AJ, Kim YC, Pierson LS, 3rd, Pierson EA, Lindow SE, Kobayashi DY, Raaijmakers JM, Weller DM, Thomashow LS, Allen AE, Paulsen IT. 2012. Comparative genomics of plant-associated Pseudomonas spp.: insights into diversity and inheritance of traits involved in multitrophic interactions. PLoS Genet 8:e1002784.

15. Mulet M, Lalucat J, Garcia-Valdes E. 2010. DNA sequence-based analysis of the Pseudomonas species. Environ Microbiol 12:1513-1530.

Appendix Figure III.1: Housekeeping genes used for screening and classifying the newly sequenced *Pseudomonas spp.* strains.

Gene	Expressed protein	Included in original gene screen ?	6 gene MLSA scheme?	7 gene MLSA scheme?	8 gene MLSA scheme?
AscA	acetyl-coenzyme A synthetase	X			X
DnaE	DNA polymerase III alpha subunit	X	X	X	X
GuaA	GMP synthease	X	X	X	X
MutL	DNA mismatch repair protein	X	X	X	X
PpsA	phosphoenolpyruvate synthase	X	X	X	X
PyrC	pyrimidine	X		X	X
RecA	recombinase A	X	X	X	X
RpoB	RNA polymerase subunit beta	X	X	X	X

Appendix Figure III.1: Housekeeping genes used for screening and classifying the newly sequenced *Pseudomonas spp.* strains.
'X' indicates that the gene was included in the original screen and in the corresponding MLSA scheme.

Appendix Figure III.2: Housekeeping genes in the newly sequenced *Pseudomonas spp*. strains.

DnaE	PpsA	RecA	RpoB	GuaA	MutL	PyrC	AcsA
1020	1020	1020	1020	1020	1020	1020	1020
1314	1314	1314	1314	1314	1314	1314	1314
1924	1924	1924	1924	1924	1924	1924	1924
1984	1984	1984	1984	1984	1984	1984	1984
2989	2989	2989	2989	2989	2989	2989	2989
4391	4391	4391	4391	4391	4391	4391	4391
4510	4510	4510	4510	4510	4510	4510	4510
4899	4899	4899	4899	4899	4899	4899	4899
5633	5633	5633	5633	5633	5633	5633	5633
6320	6320	6320	6320	6320	6320	6320	6320
6327	6327	6327	6327	6327	6327	6327	6327
6450	6450	6450	6450	6450	6450	6450	6450
6763	6763	6763	6763	6763	6763	6763	6763
7257	7257	7257	7257	7257	7257	7257	7257
7268	7268	7268	7268	7268	7268	7268	7268
7350	7350	7350	7350	7350	7350	7350	7350
7686	7686	7686	7686	7686	7686	7686	7686
7783	7783	7783	7783	7783	7783	7783	7783
7909	7909	7909	7909	7909	7909	7909	7909
9603	9603	9603	9603	9603	9603	9603	9603
10973	10973	10973	10973	10973	10973	10973	10973
11103	11103	11103	11103	11103	11103	11103	11103
11114	11114	11114	11114	11114	11114	11114	11114
11447	11447	11447	11447	11447	11447	11447	11447
11518	11518	11518	11518	11518	11518	11518	11518
11706	11706	11706	11706	11706	11706	11706	11706
12215	12215	12215	12215	12215	12215	12215	12215
12921	12921	12921	12921	12921	12921	12921	12921
13852	13852	13852	13852	13852	13852	13852	13852
14022	14022	14022	14022	14022	14022	14022	14022
14265	14265	14265	14265	14265	14265	14265	14265
14440	14440	14440	14440	14440	14440	14440	14440
14705	14705	14705	14705	14705	14705	14705	14705
14823	14823	14823	14823	14823	14823	14823	14823
14917	14917	14917	14917	14917	14917	14917	14917
15147	15147	15147	15147	15147	15147	15147	15147
15212	15212	15212	15212	15212	15212	15212	15212
16377	16377	16377	16377	16377	16377	16377	16377

16960	16960	16960	16960	16960	16960	16960	16960
17294	17294	17294	17294	17294	17294	17294	17294
17465	17465	17465	17465	17465	17465	17465	17465
18240	18240	18240	18240	18240	18240	18240	18240
18421	18421	18421	18421	18421	18421	18421	18421
20219	20219	20219	20219	20219	20219	20219	20219
21044	21044	21044	21044	21044	21044	21044	21044
DnaE	**PpsA**	**RecA**	**RpoB**	**GuaA**	**MutL**	**PyrC**	**AcsA**

DnaE	**PpsA**	**RecA**	**RpoB**	**GuaA**	**MutL**	**PyrC**	**AcsA**
6026	6026	6026	6026	6026	6026	6026	6026
1044	1044	1044	1044	1044	1044	1044	
2390	2390	2390	2390	2390	2390	2390	
6158	6158	6158	6158	6158	6158		6158
10414	10414	10414	10414	10414	10414	10414	
11122	11122	11122	11122	11122	11122	11122	
11164	11164	11164	11164	11164	11164	11164	
11235	11235	11235	11235	11235	11235	11235	
11336	11336	11336	11336	11336	11336	11336	
11831	11831	11831	11831	11831	11831		11831
12597	12597	12597	12597	12597	12597	12597	
12644	12644	12644	12644	12644	12644	12644	
14433	14433	14433	14433	14433	14433		14433
14757	14757	14757	14757	14757	14757		14757
8125	8125	8125	8125	8125	8125	8125	
60	**60**	**60**	**60**	**60**	**60**	**56**	**50**
DnaE	**PpsA**	**RecA**	**RpoB**	**GuaA**	**MutL**	**PyrC**	**AcsA**

Appendix Figure III.2: Housekeeping genes in the newly sequenced *Pseudomonas spp*. strains.
The housekeeping genes used in the intial screen are listed along the top and bottom of the table. If a strain contains a homolog to a gene, it is listed within that gene's column. Gray blocks indicate that no homolog was found in that particular strain. The number of strains that contain that particular gene are listed along the bottom. Subclade I strains are highlighted in orange. Subclade II strains are highlighted in purple. Subclade III strains are highlighted in blue. Sublcade IV strains are highlighted in green.

Chapter IV

Comparative Genomics of *Pseudomonas fluorescens* Subclade III Strains from Human Lungs

Modified from

Comparative genomics of *Pseudomonas fluorescens* subclade III strains from human lungs. Scales BS., *et al.* BMC Genomics 2015 *in press.*

BACKGROUND

Pseudomonas spp. are gram-negative, aerobic, nonsporulating, motile coccobacillus bacteria within the *Gammaproteobacteria* class. *Pseudomonas fluorescens* bacteria produce a fluorescent pigment (pyocyanin) from which the species gets its name (1). With ongoing advancements in taxonomy and comparative genomics, many *Pseudomonas* isolates originally identified as the "species" *fluorescens* are now being re-classified as novel *Pseudomonas* species within the *P. fluorescens* "species-complex" (1-4). Reclassification has identified up to 52 independent ‘species’ within the *P. fluorescens* species-complex (5). Members of the *P. fluorescens* species-complex have versatile metabolic capabilities, allowing them to thrive in a large range of environments (2). *P. fluorescens* is most commonly identified as a microbe of the soil, plant and rhizosphere environments and is widely studied for plant-promoting capabilities

(6-16). Strains of *P. fluorescens* produce a wide-range of secondary metabolites that are important for its association with plants. These include antibiotics, such as phenazine (9, 17), polyketides (8, 13), cyclic lipopeptides, biosurfactants (18); phytohormones (19) and metabolites that alter plant hormone levels (20, 21). However, many traits of *P. fluorescens* also benefit its survival in a mammalian host, such as the production of siderophores (22-24) and bioactive metabolites (7, 13, 25-29) and the ability to form biofilms (30-34). While environmental strains of *P. fluorescens* have optimal growth temperatures between 4°C and 27°C, there are documented cases *P. fluorescens* strains with an expanded growth temperature range that allow them to colonize a human host (31, 35-37) and interact with human cells *in vitro* (38, 39).

Acute infections caused by *P. fluorescens* do occur in susceptible individuals but these are often 'opportunistic' in nature. The most common causes for acute infections by *P. fluorescens* are through the contamination of blood products (40-45) or the use of contaminated hospital equipment (46-49). However, an increasing amount of evidence suggests that some strains of *P. fluorescens* are members of the human microbiota and interact with the human host in ways that contribute to certain chronic diseases (1). For example, there is a strong link between the presence of circulating antibodies to the *P. fluorescens*-specific protein *pfi*T and such autoimmune diseases as Crohn's disease and irritable bowel disease (IBD) (50-54); celiac disease (55-57); chronic granulomatous disease (58) and ankylosing spondylitis (59). We have previously reported that in the absence of acute disease, *P. fluorescens* is routinely cultured from clinical respiratory specimens at a low rate in a hospital setting (60). *P. fluorescens* strains are often isolated from respiratory specimens from individuals with cystic fibrosis (CF), though this often goes unreported. Additional culture-independent analyses have identified *P. fluorescens* as a low-abundance member of the human microbiome (60-64).

There is a large degree of genetic diversity among bacterial strains classified as *P. fluorescens*, which likely reflects the wide range of growth capabilities in these bacteria. Using multi-locus sequence analysis (MLSA) and 16S rRNA gene classification approaches, the *P. fluorescens* species complex can be divided into three smaller taxonomic subclades (1, 2, 65-68). Loper *et al.,* used ten housekeeping genes (*acsA, aroE, dnaE, guaA, gyrB, mutL, ppsA, pyrC, recA,* and *rpoB*) to divide seven fully sequenced strains within the *P. fluorescens*-species complex into these three distinct subclades (65). Subclade I consists of strains *P. protegens* Pf-5, *P. chlororaphis* 30-84 and *P. chlororaphis* O6. Subclade II consists of strains *P. fluorescens* Pf0-1, *P. brassicacearum* Q8r1-96 and *P. fluorescens* Q2-87. Subclade III consists of strains *Pseudomonas sp.* (now known to be *P. synxantha*) BG33R, *P. fluorescens* A506 and *P. fluorescens* SBW25. Strains within a subclade share a higher proportion of conserved domains than the species-complex as a whole. While 2,789 genes were found to comprise the *P. fluorescens* species-complex 'core genome', the core genomes within a subclade were much higher: 4,188, 3,729 and 3,893 genes for subclades I, II and III respectively (65). Comparative genomic analysis revealed that these subclade divisions also correlate with potentially important functional differences. For example, the genes necessary to produce a type-III secretion system (T3SS) were found in subclades II and III, but not subclade I (65).

Though much work has been done on the taxonomy and genomics of the *P. fluorescens*-species complex, as of yet, no strain isolated from a mammalian (or human source) has been determined. In this report, we provide the first comparative genomics analysis of *P. fluorescens* strains isolated from human clinical sources. The seven *P. fluorescens* clinical strains were isolated from respiratory specimens from individuals with CF and represent the first publically available draft genomes of human-associated *P. fluorescens* strains. Phylogenetic analysis reveals that these seven clinical strains fall within subclade III of the *P. fluorescens* species-complex. Our analysis reveals an extremely high degree of similarity between the

clinical and environmental isolates; however, there are a number of genomic differences between the clinical and environmental subclade III strains, which likely reflect selective pressures unique in the CF lung compared to the environment.

RESULTS AND DISCUSSION

Collection of *P. fluorescens* isolates

The newly sequenced clinical strains of *P. fluorescens* were isolated over a seven-year period from respiratory samples collected from various hospitals across the United States. The seven *P. fluorescens* strains were collected between March 2001 and January 2008 from five different hospitals. Six were isolated from adult patients with CF (five from sputum samples and one from a throat swab). One was isolated from an infant patient with CF (**Figure IV.1**). Initial sequence-based analysis of the 16S rRNA-encoding gene at the time of isolation was consistent with *P. fluorescens* and indicated that they were not *P. aeruginosa* (data not shown). Since the isolates were sourced from different CF patients from five separate hospitals across a seven year period, each isolation likely represents a distinct bacteria strain.

Phenotypic and growth properties

The phenotypic and growth properties of the newly sequenced clinical strains were compared to those of a previously sequenced representative *P. fluorescens* strain, SBW25.. The clinical strains all displayed phenotypic characteristics of bacteria within the *P. fluorescens*-species complex, e.g. motile, gram-negative, coccobacillus, lactose-fermentation negative, citrate growth positive, and fluorescence under ultra violet (UV) light (data not shown). While the clinical strains were phenotypically similar to the environmental strains in the above listed traits, they differed in their ability to grow at temperatures above 30°C. Environmental strains of *P. fluorescens* typically have an optimal growth range between 22 and 30°C (1). After 24 hours, all

of the clinical strains were able to grow at 32°C or higher in the laboratory, while the environmental subclade III strain SBW25 was unable to grow at temperatures higher than 30°C (data not shown). As we have discussed in a recent review, the temperature of the respiratory tract in a healthy individual is not uniform; it is a gradient from near that of inspired air (at the nares) to core body temperature (in the alveoli) (60). This temperature gradient is most likely altered in a CF lung but a temperature gradient will still exist due to the movement of ambient temperature air in and out of the airways. Thus, all of the isolates in our study can easily grow at temperatures found in the lungs, some potentially further down than others. Other investigators have reported that human isolates of *P. fluorescens* have an increased temperature range of growth persmissivity (1, 31, 35-37), consistent with our observations. The newly sequenced clinical strains are able to grow at temperatures at or above 32°C but otherwise show the same phenotypic and growth characteristics of previously studied strains of *P. fluorescens*.

Genomic assembly

The seven clinical *P. fluorescens* strains were *de novo* sequenced on the Illumina HiSeq platform. Genomic DNA was prepared for each strain and the output paired-end read sequences were partially assembled using DNAstar's SeqMan NGen software. Assembly statistics of each of the newly sequenced strains are shown in **Figure IV.2**. Contigs were assembled into numerous scaffolds. Average coverage across genome ranged from 44X (AU10973) to 68X (AU14705). An example of the type of coverage seen across each scaffold is seen in **Appendix Figure IV.1**.

Genomic features

The genome size, percent G+C content and # of coding sequences were calculated for each of the newly sequenced clinical strains and compared to that of previously sequenced

subclade III strains. A summary of the genomic features of the *P. fluorescens* strains is shown in **Figure IV.3**. The clinical strains have genome sizes that range between 6.1 megabase pairs (Mbps) and 6.88 Mbps. This is similar to the range seen in environmental strains (5.95 - 6.72 Mbps). The GC content per genome is also similar between clinical and environmental strains (59.5 - 60.8% versus 59.6 - 60.5%, respectively). The only plasmid found in any strain was in A506 (65). Due to the draft nature of the genomes and the difficulty of sequencing and assembling across multiple copies of nearly identical genes, we were only able to identity all of the 16S rRNA genes in the newly sequenced clinical strains. Published accounts of the previously sequenced environmental strains inform us that *P. fluorescens* strains typically contain between 5-6 copies of the 16S rRNA gene, while we were only able to identity between 2-5 copies in each of the clinical strains. However, all rRNA gene sequences in our strains were >97% identical to that in *P. fluorescens* strain A506, with some >99% identical (**Appendix Figure IV.2**). Overall, the clinical strains had the same general genomic characteristics as those found in the environmental *P. fluorescens* strains.

Phylogenetic analysis

To taxonomically classify the newly sequenced clinical strains, we utilized a multi-locus sequence approach (MLSA) modified from the scheme described by Loper *et al.* (65) (**Figure 1**). As a methodologic control (sequencing and assembly), we sequenced and de novo assembled SBW25 to determine its placement in the *P. fluorescens* phylogeny. All publically available, sequenced *P. fluorescens* isolates and some additional strains within the *Pseudomonas* genus were included in the phylogenetic analysis. *P. fluorescens* subclade I contains two *P. protegens* strains- Pf5 and CHA0- and two *P. chlororaphis* strains- O6 and *aureofaciens* 30-84. Subclade II contains *P. fluorescens* strains Pf01, R124, F113, Q2-87, Q8r1-96 and *P. brassicacearum subsp. brassicacearum* NFM421. Within these two subclades, we were also able to recapitulate smaller, intra-clade branching of the individual strains. Within

subclade II there are two smaller branches, one of F113, Q8r1-96, Q2-87 and one of Pf01 and R124 (65, 69). Within subclade I there are also two smaller branches, one *P. protegens* branch and one *P. chlororaphis* branch. The final four strains studied in Loper *et al.,* A506, SBW25, SS101 and BG33R, group together tightly into a third subclade in both their original MLSA and our larger analysis (**Figure 1**). In our analysis, the previously sequenced strains of *P. fluorescens* formed three distinct phylogenetic subclades within the phylogenetic tree, consistent with the initial publication of these strains (65, 69).

Phylogenetic analysis placed the clinical *P. fluorescens* strains within subclade III. Strain AU6026, isolated from CF sputum in Seattle, WA (**Figure IV.1**), segregates close to environmental strain A506, which was taken from a leaf surface isolate in California (65) (**Figure 1**). Clinical strain AU10973, isolated from a patient in Salt Lake, UT in 2006, segregates closest with strain SBW25, isolated from a sugar beet in England. Overall, there was no clustering pattern of clinical vs. environmental isolates in our analysis; however, all seven of the clinical isolates clearly fell with subclade III of the *P. fluorescens* species-complex.

Additional phylogenetic analysis of the clinical isolates was performed using average nucleotide identity (ANI) and unweighted pair group method with arithmetic mean (UPGMA) clustering (**Figure IV.4** and **Appendix Figure IV.3**). ANI uses the shared nucleotide identity between bacteria genomes to infer taxonomic relationships, where 95% is considered the cut-off for species delineation (68, 70). We calculated the ANI between each subclade III strain (**Appendix Figure IV.3**)**.** The percent ANI shared between members of subclade III ranged from 86 - 98.9%. The ANI between *P. aeruginosa* PAO1 and every *P. fluorescens* subclade III strain was below 79%. Both of these numbers correspond to previously determined intra and inter-species shared ANI (66). Many of the phylogenetic relationships seen between strains in the MLSA are recapitulated in the tree inferred through ANI (**Figure 1 and Figure IV.4**). AU6026

branches most closely to the environmental strain A506; AU10973 branches closely to SBW25 and the three strains AU14440, AU14917 and AU14705 branch closely to each and closely to the environmental strain BG33R1. One major difference between the two trees is that AU2989 and AU11518 no longer form a separate branch away from the other strains in the phnenogram as seen in the MLSA. The ANI between AU11518 and every subclade III strain is below 88% (**Appendix Figure IV.3**). One explanation for the discrepancy between the trees is that AU2989 and AU11518 share more genetic similarity across housekeeping genes than across the genome as whole. In summary, both MLSA and ANI analysis clearly show that each of the clinical strains can be classified within subclade III; however, many of the ANI values are below 95% (**Appendix Figure IV.3**), raising the possibility that subclade III contains multiple species with the *P. fluorescens* species-complex.

Secondary Metabolites

Members of the *P. fluorescens* species-complex are well-documented to produce numerous secondary metabolites necessary for living on plants and in the soil and rhizosphere (65). The genetic sequence of every known gene involved in the production of the secondary metabolites listed in **Figure IV.5** was used to screen the newly sequenced clinical *P. fluorescens* strains using the standalone blastn function provided by NCBI (71) (**Appendix Figure IV.4**). The panel of genes involved in the production of secondary metabolites found in the clinical strains agrees with what has been reported from previously sequenced *P. fluorescens* subclade III strains, in terms of presence, absence and variable presence (65). None of the strains screened contain the genes necessary to make any of the seven *P. fluorescens*-associated antibiotics (**Figure IV.5**). In contrast, every strain contained the full gene cluster to produce the siderophore pyoverdine (72, 73) (**Figure IV.5**). The fluorescent nature of *P. fluorescens* is due to the fluorescent pigment of the pyoverdine class of siderophores (74).

Siderophores are required by bacteria to acquire iron from the environment (22). In addition to the major siderophore, pyoverdine, *P. fluorescens* strains produce multiple secondary siderophores. The genes necessary to produce the secondary siderophore, pseudomonine are present in the clinical strain AU14705 and the environmental strains A506 and BG33R. Another secondary siderophore, hemophore, is involved in the chelation of heme from eukaryotic hosts (75). The gene cluster required for the biosynthesis and efflux of this siderophore was found in clinical strains AU14705 and AU14917 and environmental strains SS101 and BG33R1.

Subclade III *P. fluorescens* strains secrete enzymes involved in acquiring nutrients from the environment and interacting with eukaryotic hosts and surrounding microbiota. Chitinases are enzymes produced by environmental strains that can break down fungal cell walls. The function of chitinases is an important contributor to the biological control nature of certain *Pseudomonas spp*. strains (76). The environmental strains A506, SS101 and BG33R1 are all known to produce chitinase (65) (**Figure IV.5**). Three clinical strains, AU6026, AU14705 and AU14917, also contain the gene cluster necessary to produce this enzyme. The gene required to produce the enzyme pectate lyase (77) was found only in the environmental strain SBW25 (65) and our in-house sequenced genome of this strain. The exoprotease AprA is another important contributor to the biocontrol nature of multiple environmental strains. In *P. fluorescens* A506, the production of AprA is associated with its biological control of fire blight disease of pear and apple plants (78, 79). and in *P. fluorescens* CHA0, the production of AprA helps protect the strain from protozoan grazing (80). Secretion of AprA requires the presence of three adjacent genes, AprD, AprE and AprF (81). Every member of subclade III contained all four genes of the AprA gene cluster (**Figure IV.5**).

Plant-associated members of the *P. fluorescens* species complex produce compounds that are able to influence the growth and development of plants. The plant hormone indole-3-

acetic acid (IAA) is produced by plants to help regulate essential biological processes, such as phytostimulation (19) (**Figure IV.5**). The gene cluster involved in the breakdown of IAA (iacHABICDEFG) was only present in the genome of environmental strain BG33R1 (65). Acetoin is a volatile compound produced by bacteria during fermentation and also functions a growth-promoting metabolite for plants (82). The gene combination of acoRABC and budC involved in acetonin catabolism was found in the clinical strains AU2989, AU6026 and AU11518 and the environmental strains A506 and SS101. Overall, analysis of secondary metabolite gene clusters revealed no significant difference between clinical and environmental strains of *P. fluorescens*.

Pan, Accessory and Core Genomes

To determine the pan, accessory, and core genomes of the subclade III *P. fluorescens* strains, we calculated the number of COGS (clusters of orthologous groups of proteins) shared among the strains. The COG algorithm detects candidate sets of orthologous proteins, with the given requirement that included proteins are also found in at least three evolutionarily distant species (83, 84). Each protein represented by a COG is therefore thought to have evolved from a single ancestral gene (83). On average, the genomes of the strains in our study contained 5592 COGs per genome, with a range of 5332 - 6123. The pan genome, which includes all COGs, shared and unshared, among all the genomes, consists of 11,795 individual COGs (**Figure IV.6**). The core genome, which refers to only those COGs that are shared between all genomes, contains 3612 COGs (**Figure IV.6)**. The accessory genome, which is all the COGs not part of the core, consists of 8183 individual COGs. The number of shared predicted proteins in the core is higher in this study than previously reported in Loper *et al*, where all the *P. fluorescens* strains shared 2789 only predicted proteins (65). However, in Loper *et al.,* the ten strains analyzed came from all three phylogenetic subclades and represented a more genetically diverse group of strains than the ones in this study. The closer evolutionary

relationship of the strains analyzed in this study also contributes to a smaller pan genome than found in Loper *et al* (11795 vs. 13782).

Comparison between the genomes of the clinical strains and the environmental isolates reveals that they share the same general assortment of COGs. However, while the identity of the conserved protein families is the same across the subclade III genomes, certain COGs are found in significantly higher numbers in the clinical strains. Clinical strains contain additional copies of COGs involved in metal-resistance, in particular, those of the *czc*-gene family (**Figure IV.7a**). Proteins produced through the *czc* gene cluster are important in regulating resistance to potentially toxic metals, such as zinc, copper, cobalt and cadmium, in gram-negative bacteria (85-89). Though members of the *czc* gene family are also found in environmental strains, clinical strains have up to twice as many members of this gene family per genome (**Figure IV.7a**). This suggests that either duplication or genetic transfer of the metal-resistance gene cluster has occurred recently in the clinical strains.

GC Islands

We performed total and segmented GC content scans across the genomes of the newly sequenced clinical subclade III strains to look for genomic regions of recent horizontal gene transfer. The web-based software GC-Profile (http://tubic.tju.edu.cn/GC-Profile/) was used to create the GC scans (90). The GC-Profile tool segments an inputted genome based on parameters provided by the user and provides the GC content across the genome segments, as well as the cumulative GC content of the entire genome.

The range in total GC content of the newly sequenced clinical strains (59.5 to 60.8%) is similar to the range seen in the previously sequenced environmental strains (59.6 to 60.5%) (**Figure IV.3**). However, the segmented GC scans revealed genomic regions where the GC

content drops drastically from the overall average (**Figure IV.8**). We refer to these regions of lower GC content as GC Islands (GCI). GC islands with near or more than a 15% difference from the overall genome GC average were then selected for further investigation. We used the coordinates of the GCIs provided by the GC-Profile tool to select the nucleotide sequence of the GCI from each isolate's genome. This nucleotide region was then used to screen the non-redundant nucleotide database on NCBI with the online megablast function. The results of the GCI nucleotide screens are presented in **Appendix Figure IV.5** and suggest that some islands are more genetically similar to *Pseudomonas* species outside of the *P. fluorescens* species-complex.

The nucleotide screens of the GCIs from the clinical strains resulted in closest hits from both within and outside the *Pseudomonas* genus. The closest hits of the GCIs of strains AU2989 and AU6026 were to members of the *P. fluorescens* species complex, as were GCIs 1 and 3 of AU11518, GCI 1 of AU14440 and AU14705 and GCIs 2 and 4 of AU14917 (**Appendix Figure IV.5**). Strain AU10973 contained no GCIs. The top hits of GCI 2 of AU11518 were outside of the *P. fluorescens* species-complex and were a mix of unknown *Pseudomonas spp*., *P. stutzeri*, *P. aeruginosa*, and plasmids from *Klebsiella pneumonia* and *Enterobacter cloacae* (**Appendix Figure IV.5**). The GCI 2 of AU14440 showed mixed results, with two hits in the *P. fluorescens* subclade III (A506 and SBW25) and the other hits to other *Pseudomonas spp.* (*P. aeruginosa* and *P. plecogloassicida*). GCIs 1 and 3 of AU14705 showed the most varied top hits. AU14705 GCI 1 had only one significant result, that of the *Burkholderia ambifaria* MC40-6 chromosome 2. GCI 3 of AU14705 matched to non-*P. fluorescens spp.* in the *Pseudomonas* genus (*P. stutzeri, P. resionvorans*, *P. aeruginosa*, and one unknown *Pseudomonas spp*.)

Areas of GC difference within a bacterial genome are often evidence of horizontal gene transfer (91-93). Horizontal gene transfer between bacteria species is one way in which new

traits, such as secretion systems (94-96), antibiotic resistance (97-99) and other potential pathogenic factors (100, 101) evolve. Analysis of the GC islands within the clinical strains suggests that some of these islands were acquired through horizontal transfer events from other species outside the *P. fluorescens* species-complex and, perhaps, even non-pseudomonad bacteria.

The *pfi*T / I2 Region

The *pfi*T / I2 region of *P. fluorescens* encodes for a novel superantigen that is highly associated with various enteric and autoimmune diseases (102). The difference between *pfi*T and I2 is subtle and was based on a selective screen for a peptide present in active Crohn's Disease (CD) lesional tissue but absent in healthy tissue; the microbial-based sequence was termed 'I2' (103). The microbial source of the I2 sequence was later identified as the *P. fluorescens* genome (104). Cloning and genome walking revealed a 208-amino acid open reading frame termed *pfi*T, of which residues 49 and 148 contain the I2-specific portion of the protein. At the nucleotide and amino acid level, *pfi*T and I2 are two distinct sequences (I2 is smaller and resides within the larger *pfi*T sequence), the expressed protein is often called both I2 and *pfi*T interchangeably in the literature. Though *pfi*T/I2 has been identified as *P. fluorescens* specific immunostimulatory protein, no one has yet characterized the *pfi*T /I2 sequence in the genomes of fully sequenced *P. fluorescens* strains.

To investigate differences at the amino acid level, we screened each fully sequenced subclade III strain using a local blastp search with the published *pfi*T amino acid sequence (103). The top hit was selected from each genome and used to create a large multiple protein alignment (**Figure IV.9**). The published *pfi*T protein sequence was used as a reference and only the residues that differ from the reference were colored. Only 29 of the 208 (13.9%) showed any variation between the strains. Seven of these amino acid changes are seen in just one of the

subclade III strains. Two residues are altered in every strain, a threonine to an isoleucine at position 80 and a lysine to a glutamic acid at position 108. At position 16 every strain but AU14440 and A506 had a serine to a glycine substitution. Strain AU14440 has a serine to an alanine substitution at this position and A506 has no substitution at this position. The *pfi*T protein was previously determined to belong to the TetR-family transcriptional regulators (102). None of substitutions that occur in the *pfi*T sequence from subclade III strains occur in residues known to be important for DNA binding in TetR-family transcriptional regulators (102). Amino acid analysis of the *pfi*T region in the subclade III strains reveals that no substitutions have occurred in residues known to be important to function of expressed protein.

The *pfi*T/I2 protein has immunostimulatory properties indicative of a T-cell superantigen (102, 104, 105). These superantigen properties include the ability to stimulate naïve and memory T-cells; dependence on MHC-II recognition and a lack of antigen processing prior to T-cell stimulation (104). Within the disease literature, elevated antibodies to the I2 expressed protein has been linked to Chron's disease and IBD(50-54); celiac disease (55-57); chronic granulomatous disease (58) and ankylosing spondylitis (59). The presence of anti-I2 antibodies has been studied for its predictive power in CD-related complications (106-108) and CD-related surgery (51, 53, 108, 109). It has been proposed that *P. fluorescens* is a commensal member of the gastrointestinal tract's microbiota and that during disease there is epithelial cell 'leakage' that leads to presence of serum antibodies to microbial-specific proteins (110). Contributing to this theory, *P. fluorescens* has been shown to alter the permeability of intestinal epithelial cells *in vitro* (111, 112) and therapies that target the microbiota also effect serum anti-I2 Abs (53, 54). An aberrant antibody response to the *pfi*T region of indigenous strains of *P. fluorescens* could therefore contribute to the pathogenesis of such diseases as Crohn's and IBD.

Secretion Systems

Type II Secretion System

The Type II secretion system (T2SS), also known as the general secretion (gsp) pathway, is a highly conserved secretion system within Gram-negative bacteria (113). The T2SS is prevalent among gamma-proteobacteria (114, 115) and the *Pseudomonas* genus (113). The T2SS contains twelve core genetic components: an outer membrane secretin (gspD); a cytoplasmic ATPase (gspE), an inner membrane protein (gspF); major (gspG) and minor (gspH, I, J, K) pseudopilins, proteins that assist with ATPase attachment to the inner membrane and form an inner membrane platform (gspL, M), a pre-pseudopilin peptidase/methyltransferase (gspO) and a protein possibly involved in substrate recognition or secretion interactions (gspC) (116) (114). Gamma-proteobacteria can contain either one (114), or multiple (117-119) set(s) of T2SS genes (gspCDEFGHIJKLMO). Not all T2SSs contain the same assortment of these genes, *Azoarcus*, *Chromobacterium*, *Mesorhizobium*, *Y. pestis* and *Y. pseudotuberculosis* all lack gspM. *Acinetobacter* and *Bdellovibrio* lack gspM and gspL. *Bradyrhizobium* strains do not contain gspM, gspC or gspI. In *Acinetobacter,* secretion is still functional even without the gspM and gspL genes (120). Type II secretion is important for bacteria survival in environmental niches and human pathogenesis (113). Tissue destruction, cytotoxicity, adherence, and spread of transmission are all examples of pathogenic processes enabled by T2SS (113).

Every subclade III strain but AU14440 contained between one and three T2SS gene clusters (**Figure IV.10** and **Appendix Figure IV.6**). The clinical strains AU2989 and AU14705 contain two T2SS gene clusters. The previously sequenced environmental strain SBW25 and the SBW25 strain sequenced for this study both contain three T2SS gene clusters. These results agree with what has been previously reported for T2SS gene clusters in subclade III strains (65). Clinical strain, AU14440, is missing five general secretion genes (GspJKLMN). These missing genes are necessary for the standard pilus structure of the T2SS (116) (114) and,

therefore, AU14440 likely does not contain the standard type-II secretion system found in other *P. fluorescens* strains.

Type III Secretion System

Type-III secretion systems (T3SS) are needle-like structures that bacteria use to deliver effector proteins directly into nearby host cells (121). T3SSs fall into multiple families based on the genetic structure and the macromolecular structure of the needle complex. Differences in the three-dimensional structure of the T3SS are often reflected in the lifestyle of the bacteria in which the secretion is found. The Hrp1 T3SS (so called because it can trigger a hypersensitive response in plants) (122) is the most common T3SS found in the *P. fluorescens* species-complex (65, 123-126). The macromolecular structure that forms the 'needle' in the Hrp1 type-III secretion system is longer than the needle complexes in other T3SS families and this is thought to be due to the thick plant cellular wall that must be traversed by the Hrp1 T3SS (78). The environmental subclade III strains SBW25, BG33R, A506, SS101 and the subclade II strains Q8r1-96 and Q2-87 all contain at least one copy of a Hrp1 family T3SS (65). However, this system is not found in all *P. fluorescens* strains, as both Pf0-1 and Pf-5 do not have an Hrp1 T3SS gene cluster (65, 127). The function of the Hrp1 T3SS system in *P. fluorescens* strains is not completely understood. In some strains the system appears to protect against eukaryotic cell predation (128, 129). In *P. fluorescens* SBW25, the Hrp1 system is induced during colonization of the sugar beet rhizosphere (130) and can also induce a hypersensitive response in tobacco (124, 131). The Hrp1 T3SS of *P. fluorescens* Pf29Arp also appears to be involved in colonization of the wheat rhizosphere (65). However, strains such as Pf-5 and Pf0-1 are still able to colonize plant roots without harboring an Hrp1 T3SS (65, 100). Though many *P. fluorescens* strains harbor an Hrp1 T3SS, it is not ubiquitous throughout the species-complex and the exact function of the Hrp1 T3SS in the *P. fluorescens* is still not known.

Here we report the number of genes annotated as belonging to the Hrp1 T3SS in each of the subclade III *P. fluorescens* strains (**Figure IV.10** and **Appendix Figure IV.6**). Though the presence or absence of a T3SS has been reported in multiple *P. fluorescens* strains, the exact collection of T3SS-related genes has not. HrpD, HrpG, HrpQ and HrpT all vary in whether or not they are included in the Hrp1 T3SS gene cluster. The exact function of each gene in the Hrp1 T3SS has not been experimentally demonstrated, but the importance of each gene to the function of T3SS can be inferred by its presence in an isolate with functional T3SS. Environmental strain SBW25 has a functional T3SS (130) and lacks HrpD and HrpQ, suggesting these two genes are not required for T3SS function. Whether or not HrpG and HrpT are required cannot be asserted in a similar fashion until the functionality of a T3SS in additional subclade III strains is tested. Genes related to the Hrp1 T3SS were found in each of the subclade III strains analyzed, but it can not be determine through genetic analysis whether or not the T3SS apparatus would be functional.

Type IV Secretion System

The Type-IV secretion system (T4SS), also known as the Type-IV pili or fimbria, is a three-dimensional filament on the surface of bacteria that is important in adhesion, protease secretion and motility (132). One of the most studied motility phenotypes requiring the T4SS is that of twitching motility (133, 134). Twitching motility is a social activity by which bacteria cells move together to colonize new surfaces in response to low nutrient conditions. Host colonization and the formation of fruiting bodies and biofilms all utilize the twitching motility phenotype (91, 135-137). Twitching motility is most often exhibited in gram-negative bacteria of the *Proteobacteria* phyla (134) and *P. aeruginosa* serves as the primary model for the genetic and functional analysis of twitching motility. In *P. aeruginosa*, genes associated with Type-IV pili and twitching motility are important for infection (138-140) and biofilm formation (135, 141).

Despite the widespread occurrence of Type-IV pili genes in *P. fluorescens*, twitching motility is not always demonstrated. One recent study was able to induce a twitching motility phenotype in the subclade II strain, *P. fluorescens* R124, under low iron conditions (69). However, another subclade II strain, *P. fluorescens* Pf0-1, and the subclade III strain *P. fluorescens* SBW25 did not show a twitching phenotype under the same iron-low media conditions. All three *P. fluorescens* strains- R124, SBW25 and Pf0-1 - contain nearly the same composition of genes associated with T4SS (69). A possible explanation for this discrepancy is that strains of *P. fluorescens* from different environments require different environmental stressors to induce twitching motility. *P. fluorescens* Pf0-1 was isolated from loamy agricultural soil (2, 142) and *P. fluorescens* SBW25 was isolated from the leaf surface of sugar beet plant (2). In contrast, *P. fluorescens* R124 was isolated from a silica cave, an environment characterized by being high in mineral content but low in other nutrients (69). In addition, the *P. fluorescens* strain R124 showed active twitching motility despite lacking some genetic elements required in *P. aeruginosa* twitching motility. In *P. aeruginosa*, the T4SS genes *pil*ABCDEFGHIJKLMNOPQRSTUVWXY1Y2Z are required for twitching motility (143). The subclade III strains analyzed here are missing between one and seven of the T4SS genes required for twitching motility in *P. aeruginosa* (**Figure IV.10** and **Appendix Figure IV.6**). As was seen in the T2SS gene cluster (**Figure IV.10** and **Appendix Figure IV.6**), AU14440 is missing the largest number of T4SS-associated genes, lacking homologues of *pil*BDGHLRT and the genes that encode for Adhesin, FimA and FimD. Due to the large number of T4SS gene homologues missing in AU14440, it is likely that this isolate does not contain a functional T4SS. Every clinical subclade III strain contains a majority of the genes known to be required for twitching motility in *P. aeruginosa*; with the strain AU14440 missing the highest number of T4SS related genes.

The number of gene copies that encode for theT4SS proteins Adhesin, FimA and FimD varies considerably between the subclade III strains. FimA, a type IV fimbria major subunit protein, shows the largest variation in gene copy per genome, ranging from no gene copies per genome (AU2989, AU11518 and AU14440) to four gene copies per genome (AU10973). To better understand what these duplicate gene copies represent, we further analyzed the *fim*A genes from a sample group of subclade III strains (**Appendix Figure IV.7**). We aligned the genetic sequences of *fim*A from A506, AU6026 and AU10973 using the MAFFT algorithm (**Appendix Figure IV.7**). The alignment reveals that the genes annotated as *fim*A fall into two major groups; *fim*A 'a' group and *fim*A 'b' group. Subclade III strains A506 and AU10973 contain *fim*A homologues in both these two genetic groups. In addition, A506 and AU10973 both contain *fim*A homologues that fall within the same group but vary by single point mutations. This analysis reveals that *P. fluorescens* subclade III strains contain multiple homologues of the *fim*A gene and also duplicate copies of the same homologue that vary by point mutations along the sequence. This redundancy could be a function of the annotation software miscalling two separate genes as expressing same final product or that the *fim*A gene has been duplicated in certain subclade III strains. As the *fim*A gene has only been annotated in a select number of *Pseudomonas spp.* (http://www.pseudomonas.com/) it is not possible to know if this level of redundancy is seen across the *P. fluorescens* species complex or the rest of the *Pseudomonas* genus. Further analysis is required to understand the function of multiple *fim*A genes in the subclade III strains.

Widespread Colonizing Island (WCI) / Tad locus

The Tad locus is a new subtype of type II secretion system (94). Proteins expressed by the Tad locus are important for biofilm formation and colonization of hosts in multiple bacteria species. Due to its importance in colonizing new environmental niches the locus containing the *tad* genes has been named the 'widespread colonization island' (WCI) (144). In *P. aeruginosa,*

the genes within the WCI are regulated by the same quorum sensing circuit that regulates multiple virulence factors (145, 146). The WCI has not previously been studied in *P. fluorescens*. In our analysis, both the previously sequenced and the newly sequenced subclade III *P. fluorescens* genomes were screened for genes previously determined to be included in the WCI (94). This was done by putting each genome through the RAST annotation pipeline (147, 148) and searching for genes labeled as 'Tad', 'Cpa', 'Rcp' or as belonging to the WCI subsystem. All subclade III strains contained the ten WCI-associated genes; *cpa*ABCE and *tad*ABCDG. These are the same composition of WCI genes found in *P. aeruginosa* strains, where WCI function has been successfully documented. Based on the shared genes in the Tad locus, the *P. fluorescens* subclade III strains analyzed in this study contains the genes necessary to make a functional WCI.

Type VI Secretion System

The Type VI secretion system (T6SS) is a macromolecular complex that either transports bacterial effectors directly into target cells or releases them into the extracellular medium (149). First described in *Vibrio cholera* (150) and *P. aeruginosa* (151), the T6SS have been shown to be involved in both virulence against host cells (152, 153) and competition against nearby bacteria species (154, 155). In *P. aeruginosa*, T6SS activity is upregulated in response to the T6SS activity of sister cells (156). This is thought to be a mechanism by which *P. aeruginosa* can specifically target potentially aggressive cells within environments containing multiple bacterial species (157). Strains of *P. fluorescens* have also been shown to express an active T6SS (57). However, the presence of T6SS genes has only been demonstrated in environmental strains of *P. fluorescens.* Here provide the first analysis of a T6SS gene cluster in clinical *P. fluorescens* strains.

Every *P. fluorescens* strain within this analysis contains at least one copy of *imp*ABCDGHIJK, *vas*DH, *ppA*, *ppK*A, *dot*U, *icm*F, *clp*B, *clp*V1, *vgr*G1 and *vgr*G2 genes. Many of the T6SS genes found across all *P. fluorescens* strains are highly conserved across all bacteria strains that contain a T6SS (158) (**Figure IV.10** and **Appendix Figure IV.6**). It has been reported that at least 11 structural proteins and two effector proteins (mainly Hcp and VgrG) are required to assemble a functional T6SS (129, 158). Here we surveyed the T6SS gene clusters across the subclade III strains and found that each strain carries between 1-3 T6SS gene clusters (129). The majority of subclade III strains, except AU10973, appear to have at least a partial second T6SS gene cluster (**Figure IV.10**). Each subclade III strain also contains multiple copies of the effector proteins Hcp and VgrG. Multiple copies of effector proteins are thought to be tied to the multiple utilizations of the T6SS within a bacteria cell, allowing a bacterium to differently utilize the T6SS depending on environmental conditions (129). The newly sequenced, human-associated subclade III strains contain multiple copies of the structural T6SS genes and a large number of T6SS effectors, similar to what has been reported in other *P. fluorescens* strains.

VasH is a sigma-54 dependent transcriptional regulator that has been shown to regulate transcription of the T6SS (159-161). The sigma factor 54 is an alternate sigma factor that functions as a global regulator in multiple plant and animal associated bacteria (162, 163). Sigma factor 54 interacts with RNA-polymerase and a sigma-54 dependent transcriptional regulator (such as VasH) to initiate transcription at sigma-54 dependent promoters (163, 164). The number of *vas*H gene copies annotated in the subclade III strains ranges from two (AU10973 and AU14705) to five (AU14440). To investigate the differences between these *vas*H gene copies, we aligned the nucleotide sequences of the *vas*H genes from A506, AU6026, AU10973 and AU14440 in Mauve, along with the *vas*H gene from *Vibrio cholera* (**Appendix Figure IV.8**). The separate *vas*H homologues fall into separate groups based on nucleotide

sequence similarity. Major differences are seen between *vas*H homologues in both length of gene and shared nucleotide domains. We also visualized the nucleotide differences between the *vas*H homologues within a phylogenetic tree inferred from the Mauve alignment (**Appendix Figure IV.9**). The *vas*H 'a' homologues show the highest similarity to the *vas*H from *Vibrio cholera* (**Appendix Figures IV.8 and IV.9**). There are three additional *vas*H homologue branches based on nucleotide similarity. A506, AU6026, AU14440 contain *vas*H homologues that fall in each of the nucleotide similarity groups. AU14440 also has duplicate copy of the *vas*H homologue that branches out with the *Vibrio cholera vas*H. The two *vas*H homologues from AU10973 branch out into separate nucleotide similarity groups. The different gene copies of *vas*H fall within separate evolutionary groups based on nucleotide similarity comparison, suggesting that these could represent distinct cases of horizontal gene transfer.

Metal Resistance Genes

While there was no major difference between the consortium of genes that made up the core genomes of the environmental and clinical strains, the clinical strains did contain more homologues of genes involved in metal resistance. In particular, the clinical strains contained elevated numbers of genes in the *czc*-family gene cluster (**Figure IV.7a**). The *czc* gene cluster is involved in zinc, cadmium and cobalt resistance in various gram-negative bacteria, including *P. aeruginosa* (88, 165, 166). The *czcCBA* gene cluster encodes for a resistance nodulation cell division (RND)-type efflux pump that actively shuttles the metal cations out of the bacterial cell (165, 167). CzcA constitutes the inner part of the of the RND efflux pump and is a chemiosmotic cation/proton antiporter driven by a H+ gradient (165, 167). Environmental strains contained on average 18 genes belonging to the *czc* gene cluster (range 17-19) while clinical strains contained on average 33 *czc*-related genes (range 23-47). Focusing on just the gene that encodes for the *czc*A protein, environmental strains each contained three *czcA* homologues while clinical strains contain between three and eight *czcA* homologues (**Figure IV.7b**).

To determine whether the presence of additional copies of the *czc*A gene was due to recent gene acquisition, the GC content across each *czc*A homologue was compared to the GC content across the entire genome of each strain. Every *czc*A homologue in the environmental strains had similar GC content as that found across the entire genome (**Figure IV.7b**). In contrast, multiple *czc*A homologues from the clinical strains showed significantly lower GC content content when compared to GC content of the entire genome. Three *czc*A homologues (from AU6026, AU11518 and AU14440) had more than a 5.5% difference in GC content when compared to the entire genome, while two *czc*A homologues (from AU2989 and AU14440) had more than a 12% difference in GC content (**Figure IV.7b** and **Figure IV.11**). This level of difference in GC content suggests that these *czc*A gene clusters could represent recent events of horizontal gene transfer.

To further investigate whether the clinical strains gained additional *czc*A homologues through horizontal gene transfer, we built a phylogenetic tree containing the predicted amino acid sequences of *czc*A homologues from clinical and environmental subclade III strains, as well as the amino acid sequences of representative *Pseudomonas* and non-*Pseudomonas* strains (**Figure IV.11**). Each clinical strain had three *czc*A homologues that grouped with the three *czc*A in the environmental strains (**Figure IV.11**). The clinical strains also contained a *czc*A homologue whose most similar neighbor is the CzcA protein from *P. putida*. The *czc*A homologues with significantly lower GC content are labeled on the tree. The three *czc*A homologues with a GC content difference of more than 5.5% all group together in the tree and are most closely related to a *czc*A homologue from the *P. putida* strains F1, KT2440 and GB-1 as well as the a *czc*A homologue from *P. fluorescens* A11518 and AU14440. The *czc*A homologues from subclade III clinical strains with more than 12% difference in GC content from genome content did not branch off with any *czc*A homologue from a *Pseudomonas* strain. The

nearest strain to the two *czc*A homologues with more than 12% difference was *Bradyrhizobium japonicum* USDA 110 (**Figure IV.11**), further suggesting that these gene copies were acquired through recent horizontal gene transfer. Clinical subclade III strains contain homologous to *czc*A that show evidence of recent horizontal transfer from non-*Pseudomonas* bacteria strains.

Efflux pumps are required in bacteria to regulate metal ion concentrations in the cell and have been linked to antibiotic resistance. Metals, such as zinc, are necessary for bacterial growth in trace amounts. However, when in abundance, metal cations become toxic to the bacterial cell. Zinc will bind to free thiol groups, which can disrupt protein function (168). The RND-type efflux pump produced by the CzcCBA proteins is regulated by the CzcR-S two-component system (165). In *P. aeruginosa*, treatment with zinc leads to the expression of the *czcRS* operon, which then leads to transcriptional activation of *czcCBA* (166). There is a link between the metal resistance conferred through the *czcCBA* efflux pump and resistance to antibiotics, such as imipenem. The same two-component regulator that turns on transcription of *czcCBA* also negatively regulates the gene *opr*D. This gene encodes for the specific porin, OprD, which is the primary route by which carbapenems, such as impienem, enter into a bacterial cell (169). Interestingly, though the *czc*A gene is typically located in a gene cluster consisting of *czc*B and *czc*C (*czcCBA*), the *czc*A homologue in AU2989 and AU14440 with >12% GC content occurs in a differently ordered gene cluster. This *czc*A homologue is preceded upstream by a ‘probable Co/Zn/Cd efflux system membrane fusion protein and followed downstream by the genes *czc*C and *Opr*D. This gene cluster arrangement is not seen with any of the other *czc*A homologues in the subclade III strains. The presence of *opr*D just downstream of these genes suggests that this porin was acquired during the same genetic acquisition event as the *czc*-related genes.

Resistance to metals and antibiotics correlate in bacteria isolated from zinc-contaminated environments (170, 171). In laboratory conditions, exposure to zinc can lead to spontaneous mutations in the *czcS* sensor gene in *P. aeruginosa,* making the bacteria resistant to imipenem exposure (166). Increased resistance to antibiotics has been seen in multiple different metal-contaminated environments, such as freshwater streams (172), costal areas (173), metal-contaminated ash settling basins (174). The clinical strains in this study were isolated from patients with CF (**Figure IV.1**). To control the chronic bacterial infections associated with CF, individuals with this disease are put on numerous antibiotics, often multiple antibiotic at the same time (175). As a result, chronically infected bacteria in the CF lung accumulate numerous mutations that lead to resistance to antibiotics (176, 177). The increased number of homologues involved in the *czc*-family protein encoding genes in the clinical strains suggests that these bacteria are adapting to the antibiotic-heavy environment of the CF lung by acquiring additional copies of efflux pumps. Strains AU2989 and AU14440 both have a *czc*-related efflux pump whose nucleotide sequence is >12% different in GC content then rest of the genome, is most closely related to non-*Pseudomonas* bacteria via amino acid similarity and occurs in a novel ordered gene cluster, suggesting that these two strains have gained a novel *czc*-related efflux pump through a recent horizontal gene transfer event from a non-*Pseudomonas* related bacteria.

CONCLUSIONS

Here we present the first comparative genomic analysis of *Pseudomonas fluorescens* isolates from clinical samples. Previously, all sequenced *P. fluorescens* strains had been isolated from environmental samples, such as the soil, plant leaves and loam (1, 65). Comparing the genome of seven strains of *P. fluorescens* isolated from the lungs of individuals with CF (**Figure IV.1**) to four previously sequenced environmental strains within the same subclade (III), we found that all eleven strains were very similar when global and individual

genomic features were compared (**Figures IV.3, IV.5** and **IV.10**). However, the clinical strains analyzed in this study differed from the environmental strains in two important ways. Firstly, the strains isolated from clinical strains had an increased temperature growth range that allowed them to grow between 32°C and 37°C, while the representative environmental strain SBW25 could not grow above 27°C. Environmental strains of *P. fluorescens* are considered psychotropic, which is one reason why *P. fluorescens* was long believed unable to colonize a mammalian or human host. The finding that the clinical isolates studied in this paper can grow at 32°C or higher agrees with other accounts of strains of *P. fluorescens* isolated (31, 35-37) Increased temperature permissively in *P. fluorescens* is often linked to increased pathogenicity against human cells (38, 39). The ambient temperature of the respiratory tract is on average 32°C - 35°C (178), which is the ideal growth temperature for the strains analyzed in this study but is above the permissive range for environmental *P. fluorescens* strains. The clinical strains also differed significantly from environmental strains in the number of protein-coding genes involved in the resistance nodulation cell division (RND)-type of efflux pump that regulate zinc, cadmium and cobalt ions in many gram-negative bacteria (88, 165, 166). Efflux pumps control the concentration of metal ions that pass through a bacteria cell, and there is strong evidence for a link between resistance to metal ions and antibiotics (166, 169). GC content and phylogenetic analysis of the *czcA* gene involved in zinc, cadium and cobalt resistance suggests that clinical strains obtained additional copies through horizontal gene transfer from bacteria outside of the *Pseudomonas* genus. The increased temperature permissivity and additional genes involved in metal-resistance efflux pumps found in the clinical *P. fluorescens* strains are both potential mechanisms of adaption by which these traditionally environmentally-derived bacteria may have adapted to the very specific environment of the human CF lung.

Humans are exposed to environmental microbes, such as *P. fluorescens*, on a continual basis, whether through breathing in the microbes that exist in showerhead biofilms (179), from building walls (180) or the air surrounding us at all times (181). *Pseudomonas spp.* are known to be highly metabolically versatile bacteria, allowing them to adapt to numerously diverse habitats (182). *Pseudomonas aeruginosa* is perhaps the most famous of the *Pseudomonas spp.* to colonize and infect human hosts. Environmental and clinical strains of *P. aeruginosa* are often very similar genomically (170, 171), with some of the major differences including increased temperature growth range (183) and the acquisition of additional antibiotic resistance traits (184). (185-188). The very definition of an “environmental”, “clinical” or “human” strain is in fact an artificial construct; bacteria will colonize and persist in any niche that provides the requirements necessary for growth. Our results suggest that the *P. fluorescens* strains isolated from clinical samples have been selected for by the human host to have different phenotypic and genetic traits than those typically found in *P. fluorescens* strains isolated outside of the human host.

Table IV.1: Date, location and source of *P. fluorescens* strains.

Isolate ID	Isolation Date	Isolation Location	Isolation Source
AU2989	**4/4/01**	**Hartford, CT**	**CF Throat Swab**
AU6026	**7/28/03**	**Seattle, WA**	**CF Sputum**
AU10973	**4/6/06**	**Salt Lake City, UT**	**CF Sputum**
AU11518	**7/12/06**	**Hartford, CT**	**CF Infant**
AU14440	**10/19/07**	**Little Rock, AR**	**CF Sputum**
AU14705	**11/13/07**	**Augusta, GA**	**CF Sputum**
AU14917	**1/11/08**	**Little Rock, AR**	**CF Sputum**

Figure IV.2: Assembly statistics of newly sequenced *P. fluorescens* strains.

		AU2989	AU6026	AU10973	AU11518	AU14440	AU14705	AU14917
Assembly Totals	# of Contigs	119	101	65	136	323	179	136
	Contigs >2K	64	70	42	66	238	42	82
	Assembled Sequences	3470296	2721507	3857173	3986627	3689916	4236424	4073286
	Unassembled Sequences	113880	64231	98101	177547	154426	309290	179064
	Sequences not assembled due to complete trimming	6672	5466	8507	7330	6450	9690	9423
	Sequences removed due to small contig size	44657	24398	28587	99733	74646	227149	83582
	All Sequences	3584176	2785738	3955274	416417	3844342	454714	4252350
	Contig N50	196 kbases	165 kbases	303 kbases	230 kbases	47 kbases	323 kbases	117 kbases
	Average Coverage	56	44	62	64	60	68	66
Average Totals	Sequences per Contig	29162	26945	59341	29313	11423	23667	29950
Average Lengths	Contigs	52308	60585	95565	46636	21338	34118	46707
	Assembled Sequences	97	97	97	97	97	97	97
	Unassembled Sequences	72	62	66	79	78	85	76
	All Sequences	96	96	96	96	97	96	96
Average Quality	Assembled Sequences	33	33	33	33	33	33	33
	Unassembled Sequences	27	27	26	29	29	30	29
	All Sequences	33	22	33	33	33	33	33
Assembled Pair Statistics	Read Pairs	1792088	1392869	1977637	2082087	1922171	2272857	2126175
	Assembled Pairs	1705204	1342761	1898009	1959906	1807056	2073494	2001733
	Pairs Consistent Within a Contig	1698401	1337220	1893522	1951647	1796379	2065932	1992757
	Pairs Inconsistent Within a Contig	247	256	60	163	56	87	100

Figure IV.2: Assembly statistics of newly sequenced *P. fluorescens* strains.
Assembly statistics of newly sequenced *P. fluorescens* strains. The paired-end reads from Illumina HiSeq were de novo assembled using the DNAstar SeqMan NGen.

Figure IV.3: Genomic features of newly sequenced *P. fluorescens* strains.

	Genome Size (Mbp)	G+C Content (%)	# of Contigs & Scaffolds	# of Coding Sequences	# of RNAs	Plasmid	Reference
AU2989	6.2	60.4	49	5569	58	No	Scales, 2015
AU6026	6.11	60	34	5427	67	No	Scales, 2015
AU10973	6.13	60.8	23	5499	61	No	Scales, 2015
AU11518	6.31	60.3	61	5684	63	No	Scales, 2015
AU14440	6.88	59.5	109	6123	71	No	Scales, 2015
AU14705	6.1	59.9	97	5379	69	No	Scales, 2015
AU14917	6.82	60.1	36	5592	68	No	Scales, 2015
A506	5.96	60	2	5267	69 (tRNAs)	Yes (57.0)	Loper, 2015
SBW25	6.72	60.5	2	5921	66 (tRNAs)	No	Silby, 2009
BG33R	6.29	59.6	4	5511	68 (tRNAs)	No	Loper, 2015
SS101	6.17	60	2	5374	68 (tRNAs)	No	Loper, 2015

Figure IV.4: Phenogram based on average nucleotide identity (ANI) between subclade III strains.

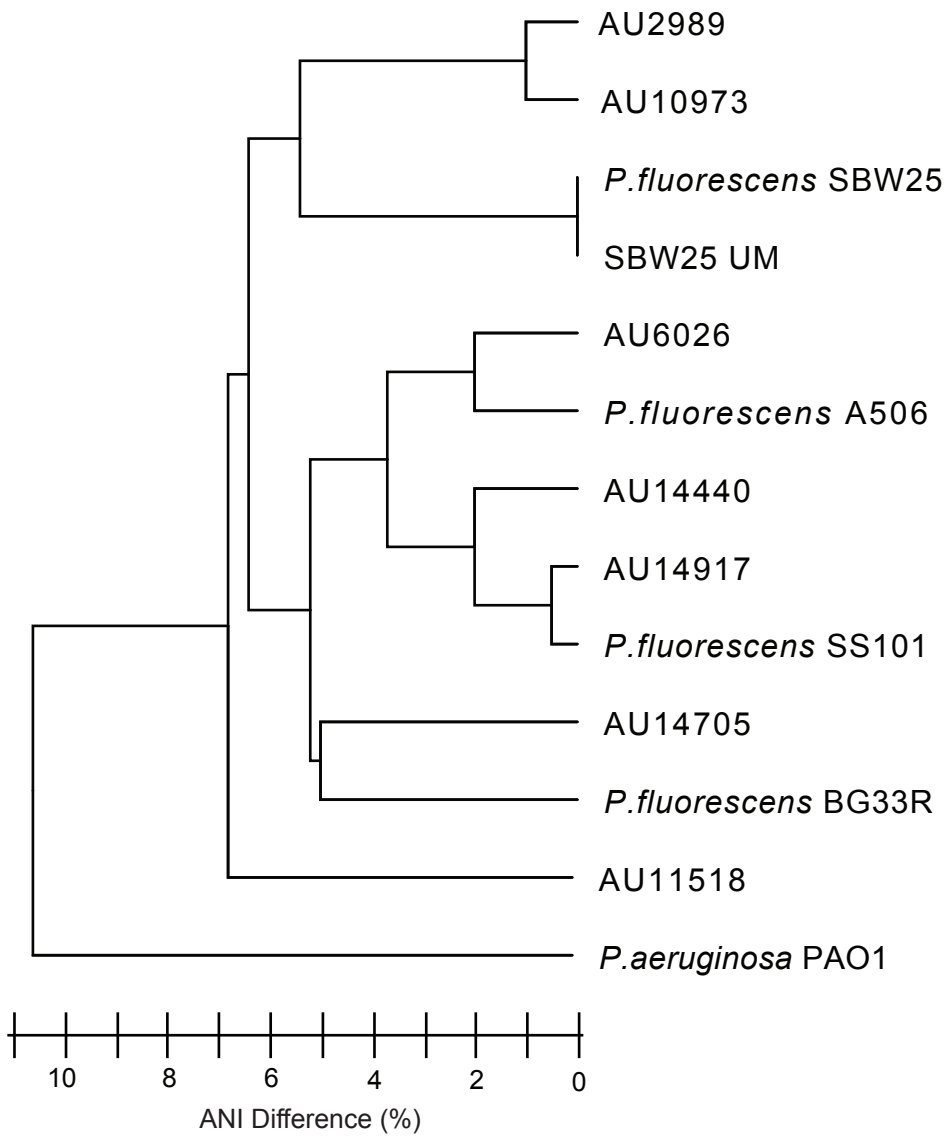


Figure IV.4: Phenogram based on average nucleotide identity (ANI) between subclade III strains.
Phenogram created based on a similarity matrix of the average nucleotide identity (ANI) between subclade III strains. Clustering performed using the Unweighted Pair Group Method with Arithmetic mean (UPGMA) algorithm.

Figure IV.5: Secondary metabolite genes and gene clusters in subclade III strains.

		Antibiotics							Siderophores					Exoenzymes			Plant-Bacterial Communication							
		DAPG	HCN	Phenazine	PyrroInitrin	Rhizoxins	Pyoluteorin	Mupirocin	Pyoverdine	Pseudomonine	Enantio-pyochelin	Achromobactin	Hemophore	Chitanse	AprA	Pectate Lyase	IAA Biosynthesis	IAA Catabolism	PAA catabolism	ACC deaminase	2,3-bd biosynthesis	Acetoin catabolism (LP)	Acetoin catabolism (DP)
Human-associated Isolates	AU2989								x						x								x
	AU6026								x					x	x								x
	AU10973								x						x								
	AU11518								x						x								x
	AU14705								x	x*			x	x	x								
	AU14917								x				x	x	x								
Environment-associated Isolates	SBW25 sample								x						x	x							
	P.fluorescens SBW25								x						x	x							
	P.fluorescens A506								x	x				x	x								x
	P. fluorescens SS101								x				x	x	x								x
	P. synxantha BG33R1								x	x*			x	x	x			x					

Figure IV.5: Secondary metabolite genes and gene clusters in subclade III strains.
The presence or absence of a gene or gene cluster within each sequenced genome. The nucleotide sequence for the gene(s) involved in the biosynthesis of each secondary metabolite was used to query each genome with the blastn function provided by NCBI. The chosen query nucleotide sequence for each gene is detailed in App_IV.5. A 'X' corresponds to a blastn hit with an evalue $<1x10^{15}$ and query/sequence identity >70%, as detailed in Person *et al*, 2014. For both AU14705 and BG33R1, one gene of the pseudomonine cluster fell below the cutoff, but since the gene cluster was perviously verified in Loper *et al,* 2012, it was marked with an 'X' for both strains.

Figure IV.6: Pan, accessory and core analysis of the elevan subclade III strains.

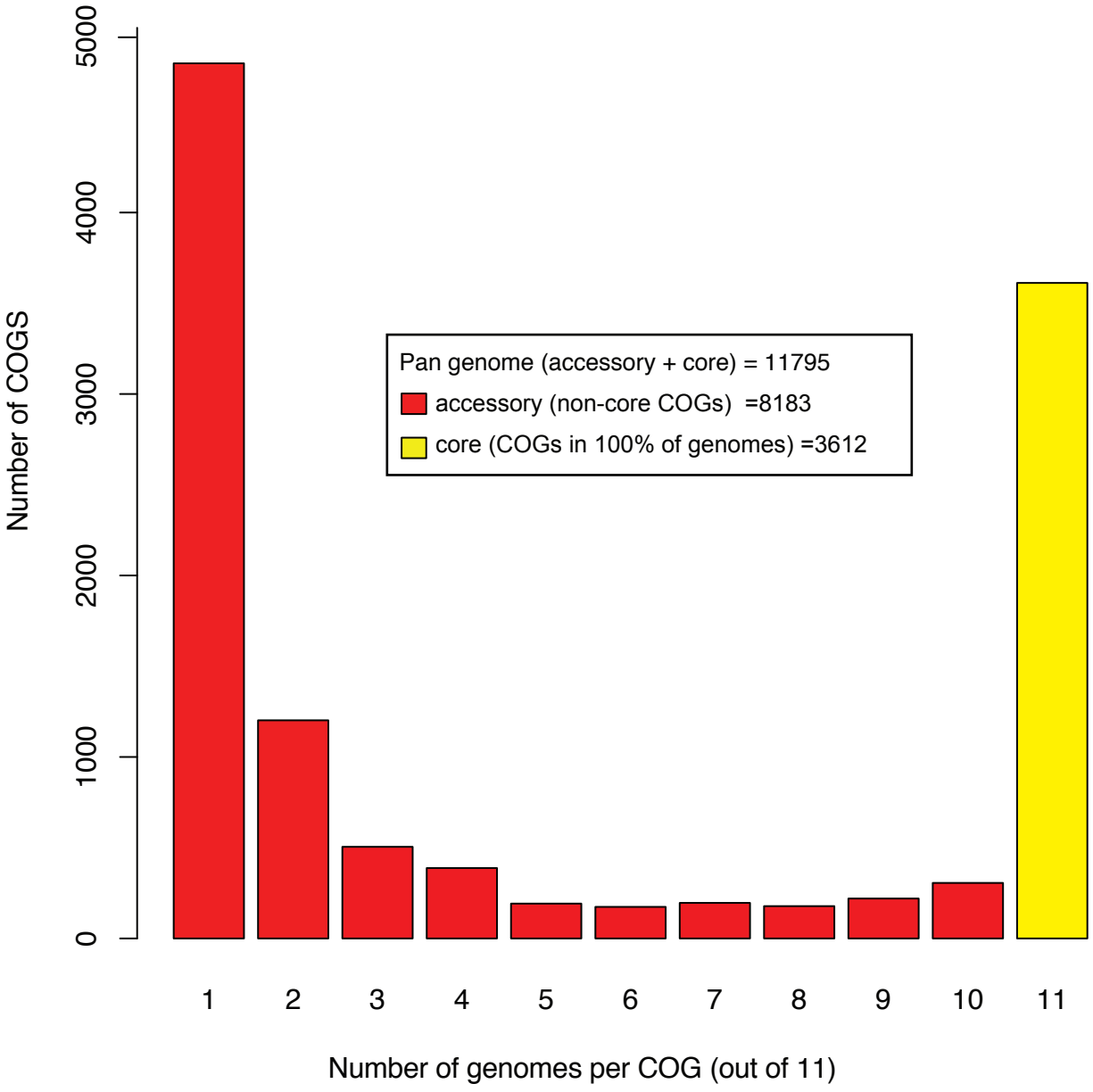


Figure IV.6: The pan, accessory and core genomes in subclade III strains.
Each of the 11795 COGS was analyzed to determine how many genomes encoded that particular COG (1-11, x-axis). The number of COGS encoded by only one genome, two genomes, etc was along determined (y-axis). The pan genome contains the accessory and core . The average number of COGs per genome is 5592 (range: 5332-6123). The pan genome and its compartments was calculated using the COGtriganle clustering algorithm [195] in GET_HOMOLOGUES [196].

Figure IV.7: The *czc* gene cluster in environmental and clinical subclade III strains.

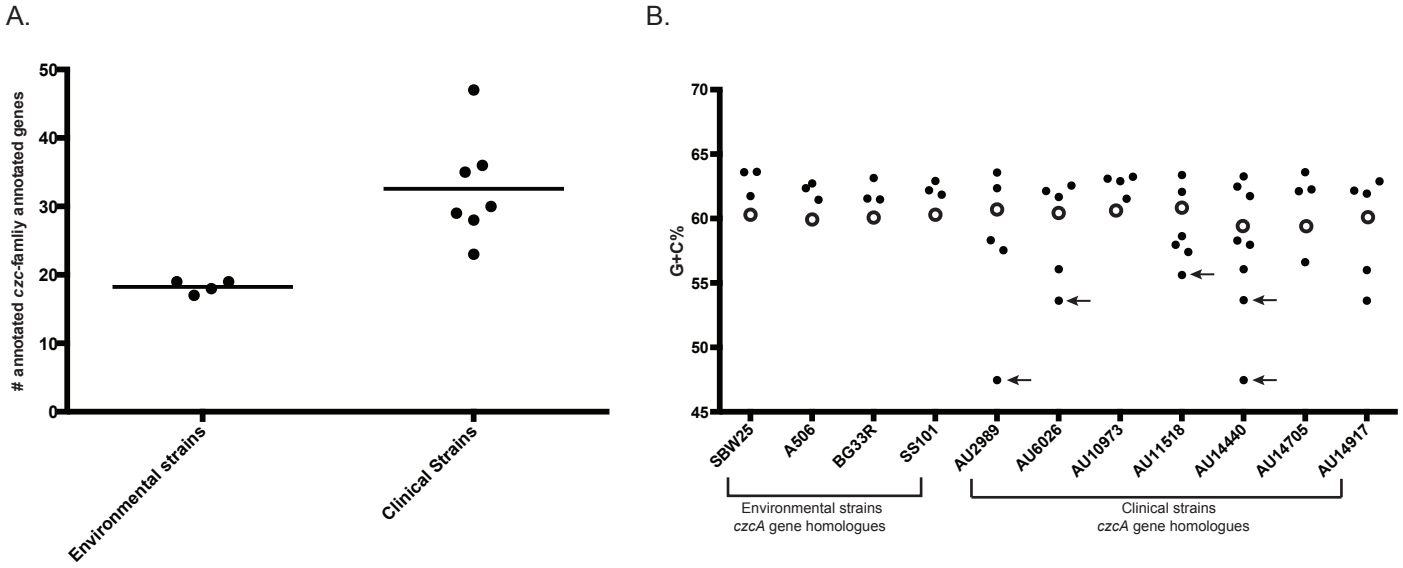


Figure IV.7: The *czc* gene cluster in environmental and clinical subclade III strains.
A. The total number of genes annotated as belonging to the czcA gene cluster in environmental and clinical subclade III strains. **B**. The percent G+C content of the czcA gene homologues. Open circles correspond to the G+C content found across the entire genome. The closed circles correspond to the G+C content found in the czcA gene homologues. Arrows indicate to the czcA homologs enclosed in boxes in **Figure IV.11**.

Figure IV.8: G+C content across the genomes of newly sequenced subclade III strains.

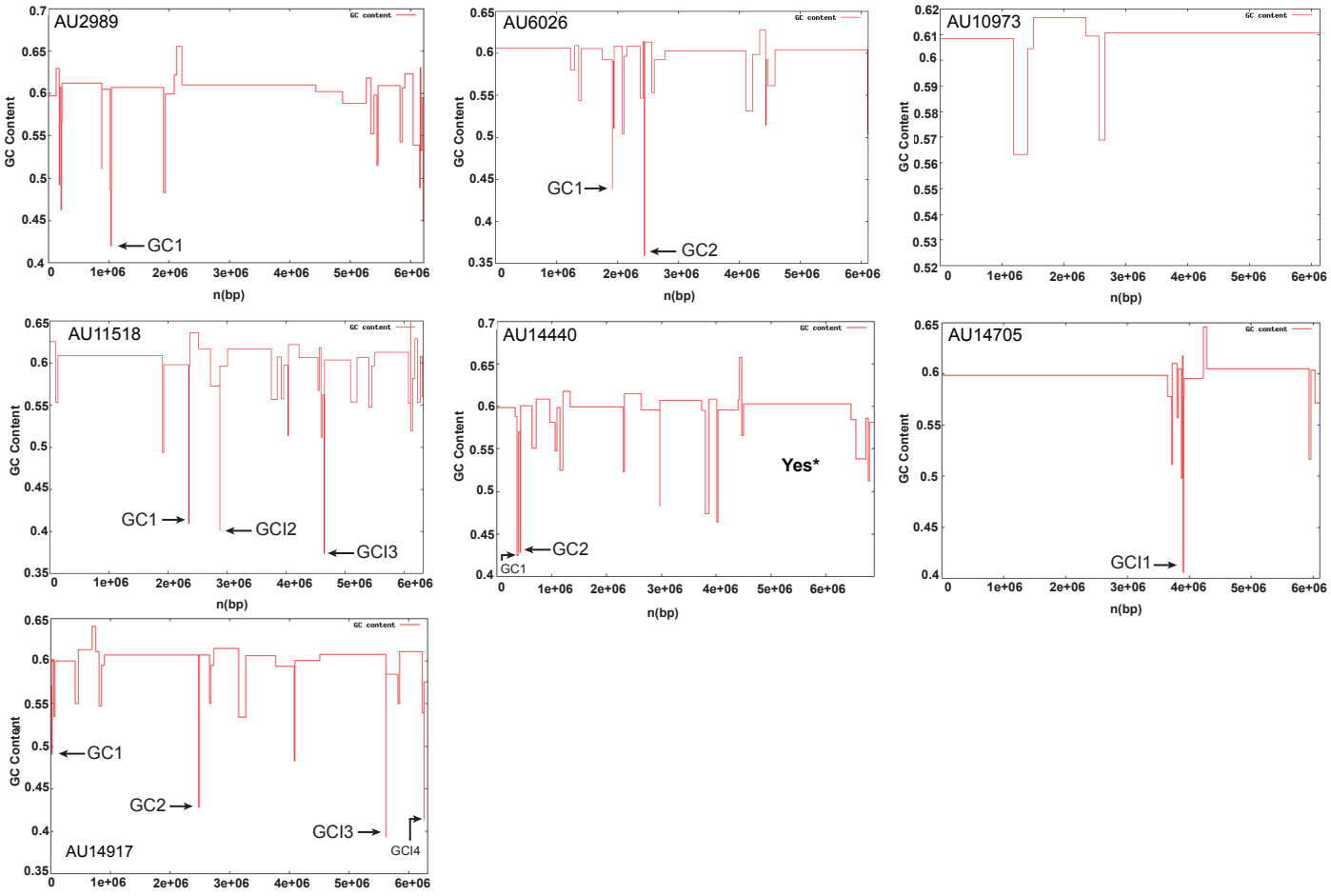


Figure IV.8: G+C content across the genomes of newly sequenced subclade III strains.
GC content calculated using the online GC profile tool [91]. GC content displayed on y-axis and position in draft genome displayed on x-axis. Arrows indicate GC islands that have been further analyzed via NCBI blastn, results in **App_IV.6**.

Figure IV.9: Alignment of the pfiT amino acid sequences from subclade III strains.

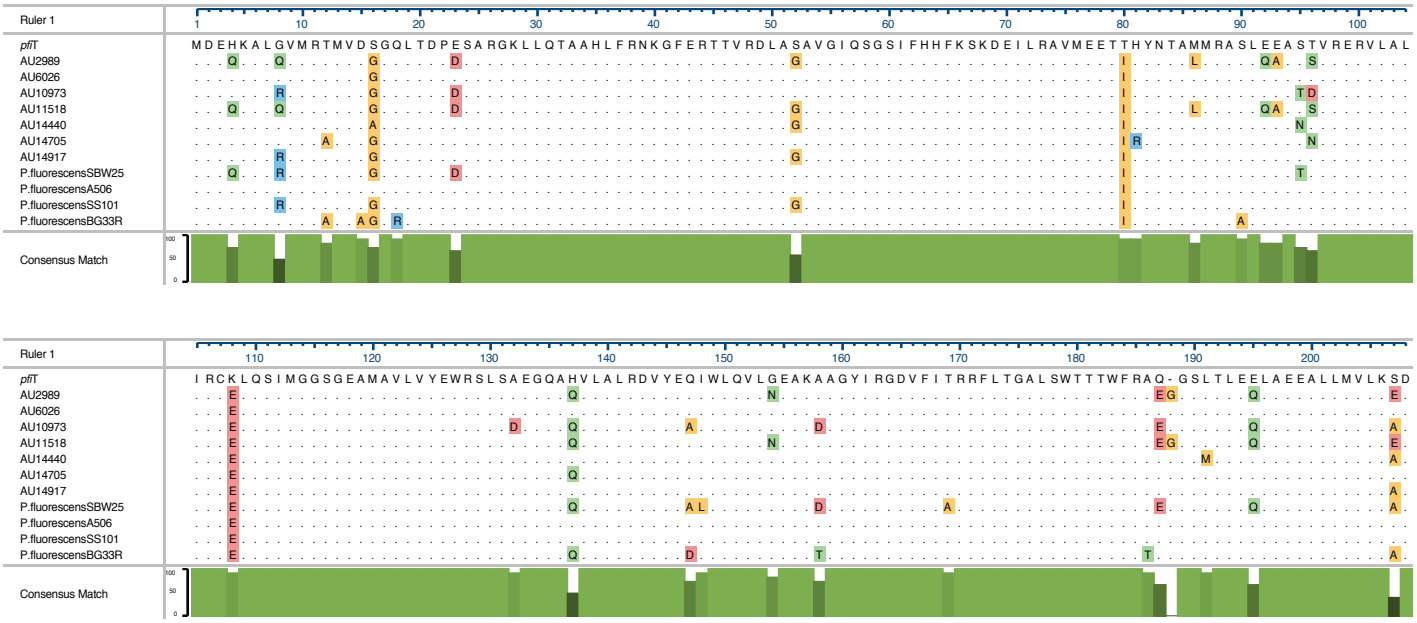


Figure IV.9: Alignment of the pfiT amino acid sequences from subclade III strains.
Amino acid sequence of the pfiT protein aligned using MUSCLE [197]. Published sequence of pfiT [105] used as reference (labeled pfiT above). Only amino acids differing from reference are shown and colored. Percent nucleotide consensus match visualized in green underneath alignment. Alignment visualized with DNAstar MegAligPro.

Figure IV.10: Genes involved in Type-II, III, IV, VI and WCI secretion system in subclade III strains.

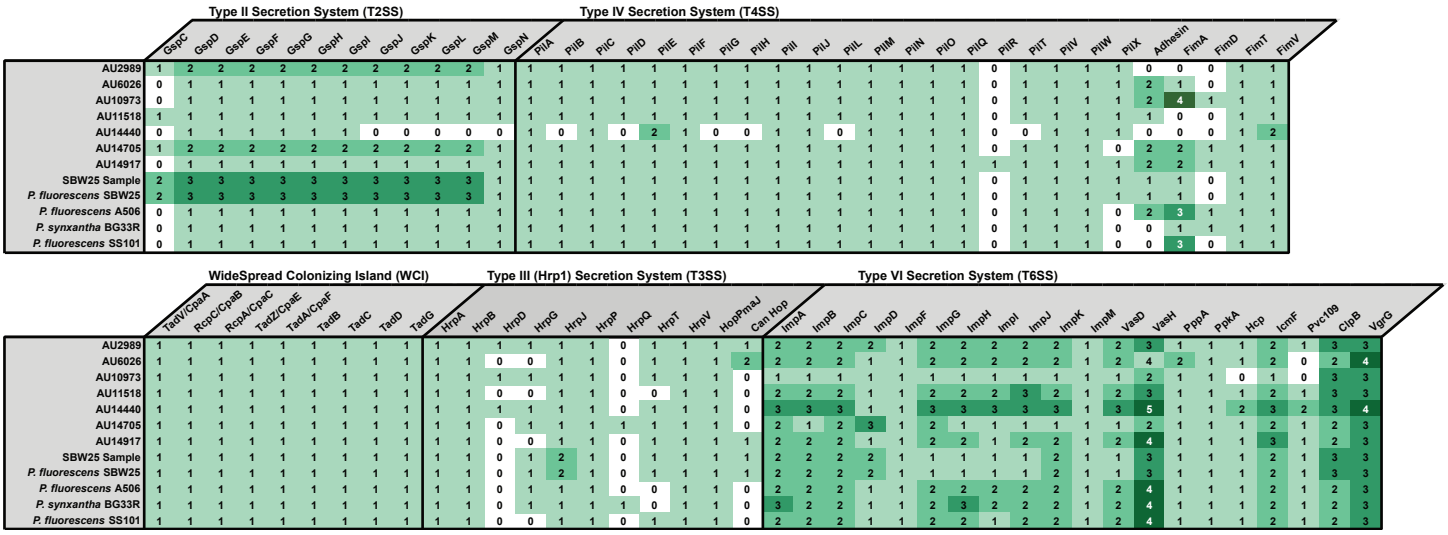


Figure IV.10: Genes involved in Type-II, III, IV, VI and WCI secretion system in subclade III strains.
Genes annotated using the RAST pipeline [148]. Full gene names listed in **Appendix File IV.8.**

Figure IV.11: Neighbor-joining tree based on the amino acid sequence of *czc*A homologues.

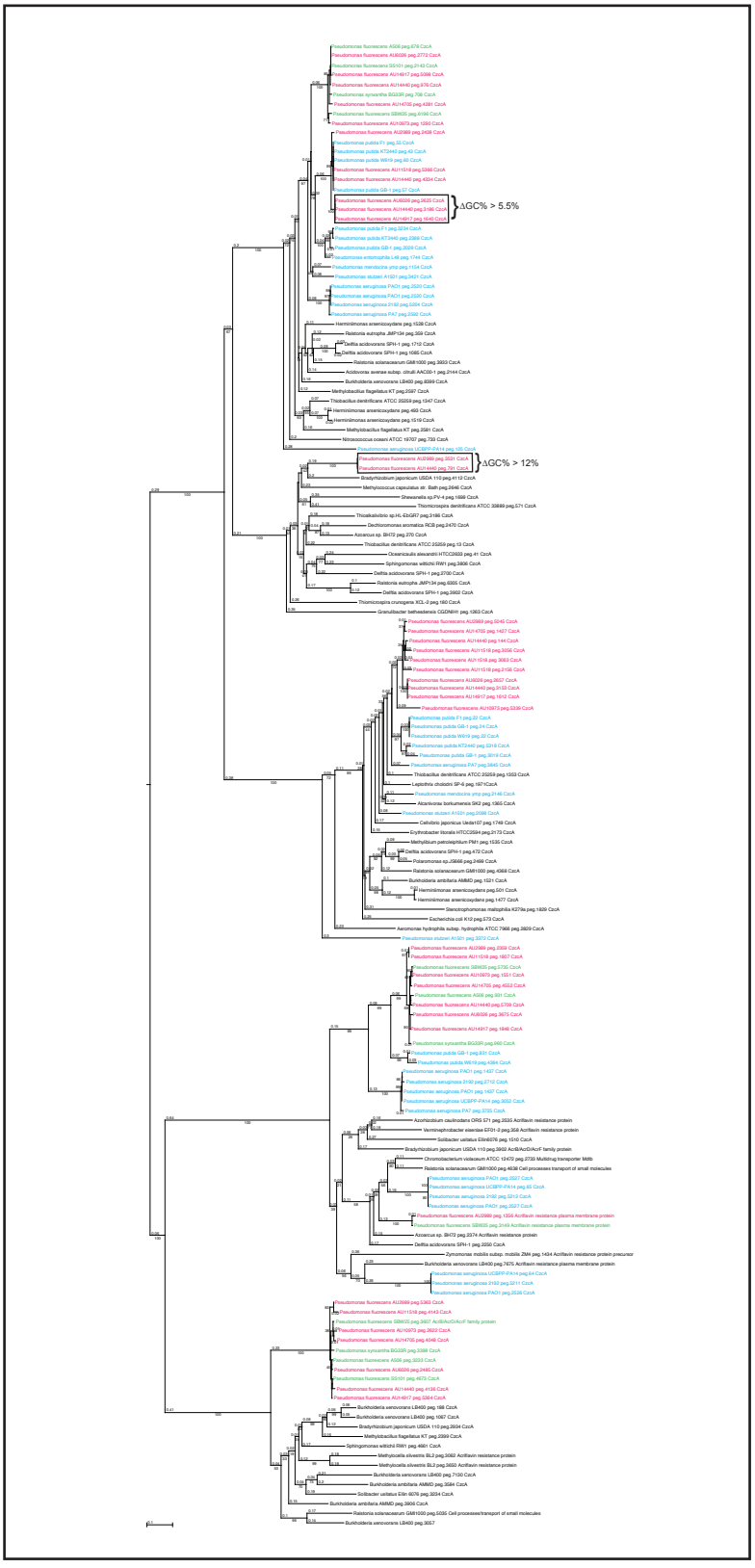


Figure IV.11: Neighborhood-joining phylogenetic tree based on the amino acid sequence of *czc*A homologues.
*czc*A homologues from clinical strains are highlighted in red; homologs from environmental strains are highlighted in green; homologues from other *Pseudomonas spp*. are highlighted in green. Amino acid sequences of *czc*A homologues discovered in RAST [149]. All other bacteria species are publically available on RAST. The amino acid sequences were aligned with MAFFT [194, 193].

Chapter IV

References

1. Scales BS, Dickson RP, LiPuma JJ, Huffnagle GB. 2014. Microbiology, genomics, and clinical significance of the Pseudomonas fluorescens species complex, an unappreciated colonizer of humans. Clin Microbiol Rev 27:927-948.
2. Silby MW, Cerdeno-Tarraga AM, Vernikos GS, Giddens SR, Jackson RW, Preston GM, Zhang XX, Moon CD, Gehrig SM, Godfrey SA, Knight CG, Malone JG, Robinson Z, Spiers AJ, Harris S, Challis GL, Yaxley AM, Harris D, Seeger K, Murphy L, Rutter S, Squares R, Quail MA, Saunders E, Mavromatis K, Brettin TS, Bentley SD, Hothersall J, Stephens E, Thomas CM, Parkhill J, Levy SB, Rainey PB, Thomson NR. 2009. Genomic and genetic analyses of diversity and plant interactions of Pseudomonas fluorescens. Genome Biol 10:R51.
3. Almeida LA, Araujo R. 2013. Highlights on molecular identification of closely related species. Infect Genet Evol 13:67-75.
4. Konstantinidis KT, Ramette A, Tiedje JM. 2006. Toward a more robust assessment of intraspecies diversity, using fewer genetic markers. Appl Environ Microbiol 72:7286-7293.
5. Peix A, Ramirez-Bahena MH, Velazquez E. 2009. Historical evolution and current status of the taxonomy of genus Pseudomonas. Infect Genet Evol 9:1132-1147.
6. Haas D, Defago G. 2005. Biological control of soil-borne pathogens by fluorescent pseudomonads. Nat Rev Microbiol 3:307-319.
7. Laursen JB, Nielsen J. 2004. Phenazine natural products: biosynthesis, synthetic analogues, and biological activity. Chem Rev 104:1663-1686.
8. Weller DM, Landa BB, Mavrodi OV, Schroeder KL, De La Fuente L, Blouin Bankhead S, Allende Molar R, Bonsall RF, Mavrodi DV, Thomashow LS. 2007. Role of 2,4-

diacetylphloroglucinol-producing fluorescent Pseudomonas spp. in the defense of plant roots. Plant Biol (Stuttg) 9:4-20.

9. Mavrodi DV, Blankenfeldt W, Thomashow LS. 2006. Phenazine compounds in fluorescent Pseudomonas spp. biosynthesis and regulation. Annu Rev Phytopathol 44:417-445.

10. Ramette A, Moenne-Loccoz Y, Defago G. 2003. Prevalence of fluorescent pseudomonads producing antifungal phloroglucinols and/or hydrogen cyanide in soils naturally suppressive or conducive to tobacco black root rot. FEMS Microbiol Ecol 44:35-43.

11. Moynihan JA, Morrissey JP, Coppoolse ER, Stiekema WJ, O'Gara F, Boyd EF. 2009. Evolutionary history of the phl gene cluster in the plant-associated bacterium Pseudomonas fluorescens. Appl Environ Microbiol 75:2122-2131.

12. Takahashi M, Matsumoto S, Iwasaki S, Yahara I. 1990. Molecular basis for determining the sensitivity of eucaryotes to the antimitotic drug rhizoxin. Mol Gen Genet 222:169-175.

13. Gross H, Loper JE. 2009. Genomics of secondary metabolite production by Pseudomonas spp. Nat Prod Rep 26:1408-1446.

14. Tsuruo T, Oh-hara T, Iida H, Tsukagoshi S, Sato Z, Matsuda I, Iwasaki S, Okuda S, Shimizu F, Sasagawa K, et al. 1986. Rhizoxin, a macrocyclic lactone antibiotic, as a new antitumor agent against human and murine tumor cells and their vincristine-resistant sublines. Cancer Res 46:381-385.

15. Schnider U, Keel C, Blumer C, Troxler J, Defago G, Haas D. 1995. Amplification of the housekeeping sigma factor in Pseudomonas fluorescens CHA0 enhances antibiotic production and improves biocontrol abilities. J Bacteriol 177:5387-5392.

16. Sarniguet A, Kraus J, Henkels MD, Muehlchen AM, Loper JE. 1995. The sigma factor sigma s affects antibiotic production and biological control activity of Pseudomonas fluorescens Pf-5. Proc Natl Acad Sci U S A 92:12255-12259.

17. Pierson LS, 3rd, Pierson EA. 2010. Metabolism and function of phenazines in bacteria: impacts on the behavior of bacteria in the environment and biotechnological processes. Appl Microbiol Biotechnol 86:1659-1670.

18. Raaijmakers JM, De Bruijn I, Nybroe O, Ongena M. 2010. Natural functions of lipopeptides from Bacillus and Pseudomonas: more than surfactants and antibiotics. FEMS Microbiol Rev 34:1037-1062.

19. Spaepen S, Vanderleyden J, Remans R. 2007. Indole-3-acetic acid in microbial and microorganism-plant signaling. FEMS Microbiol Rev 31:425-448.

20. Xiao Z, Xu P. 2007. Acetoin metabolism in bacteria. Crit Rev Microbiol 33:127-140.

21. Leveau JH, Gerards S. 2008. Discovery of a bacterial gene cluster for catabolism of the plant hormone indole 3-acetic acid. FEMS Microbiol Ecol 65:238-250.

22. Neilands JB. 1981. Iron absorption and transport in microorganisms. Annu Rev Nutr 1:27-46.

23. Youard ZA, Mislin GL, Majcherczyk PA, Schalk IJ, Reimmann C. 2007. Pseudomonas fluorescens CHA0 produces enantio-pyochelin, the optical antipode of the Pseudomonas aeruginosa siderophore pyochelin. J Biol Chem 282:35546-35553.

24. Wuest WM, Sattely ES, Walsh CT. 2009. Three siderophores from one bacterial enzymatic assembly line. J Am Chem Soc 131:5056-5057.

25. Keel C, Weller DM, Natsch A, Defago G, Cook RJ, Thomashow LS. 1996. Conservation of the 2,4-diacetylphloroglucinol biosynthesis locus among fluorescent Pseudomonas strains from diverse geographic locations. Appl Environ Microbiol 62:552-563.

26. Umio S, Kawanishi T, Kamishita T, Mine Y. 1987. Antifungal Composition Employing Pyrrolnitrin in Combination wtih an Imidazole Compound. Japan.

27. Fuller AT, Mellows G, Woolford M, Banks GT, Barrow KD, Chain EB. 1971. Pseudomonic acid: an antibiotic produced by Pseudomonas fluorescens. Nature 234:416-417.

28. Bork K, Brauers J, Kresken M. 1989. Efficacy and safety of 2% mupirocin ointment in the treatment of primary and secondary skin infections--an open multicentre trial. Br J Clin Pract 43:284-288.

29. Strock LL, Lee MM, Rutan RL, Desai MH, Robson MC, Herndon DN, Heggers JP. 1990. Topical Bactroban (mupirocin): efficacy in treating burn wounds infected with methicillin-resistant staphylococci. J Burn Care Rehabil 11:454-459.

30. Tuttlebee CM, O'Donnell MJ, Keane CT, Russell RJ, Sullivan DJ, Falkiner F, Coleman DC. 2002. Effective control of dental chair unit waterline biofilm and marked reduction of bacterial contamination of output water using two peroxide-based disinfectants. J Hosp Infect 52:192-205.

31. Donnarumma G, Buommino E, Fusco A, Paoletti I, Auricchio L, Tufano MA. 2010. Effect of temperature on the shift of Pseudomonas fluorescens from an environmental microorganism to a potential human pathogen. Int J Immunopathol Pharmacol 23:227-234.

32. Bianciotto V, Andreotti S, Balestrini R, Bonfante P, Perotto S. 2001. Mucoid mutants of the biocontrol strain pseudomonas fluorescens CHA0 show increased ability in biofilm formation on mycorrhizal and nonmycorrhizal carrot roots. Mol Plant Microbe Interact 14:255-260.

33. O'Toole GA, Kolter R. 1998. Initiation of biofilm formation in Pseudomonas fluorescens WCS365 proceeds via multiple, convergent signalling pathways: a genetic analysis. Mol Microbiol 28:449-461.

34. Baum MM, Kainovic A, O'Keeffe T, Pandita R, McDonald K, Wu S, Webster P. 2009. Characterization of structures in biofilms formed by a Pseudomonas fluorescens isolated from soil. BMC Microbiol 9:103.

35. Chapalain A, Rossignol G, Lesouhaitier O, Merieau A, Gruffaz C, Guerillon J, Meyer JM, Orange N, Feuilloley MG. 2008. Comparative study of 7 fluorescent pseudomonad clinical isolates. Can J Microbiol 54:19-27.

36. Rossignol G, Sperandio D, Guerillon J, Duclairoir Poc C, Soum-Soutera E, Orange N, Feuilloley MG, Merieau A. 2009. Phenotypic variation in the Pseudomonas fluorescens clinical strain MFN1032. Res Microbiol 160:337-344.

37. Sperandio D, Decoin V, Latour X, Mijouin L, Hillion M, Feuilloley MG, Orange N, Merieau A. 2012. Virulence of the Pseudomonas fluorescens clinical strain MFN1032 towards Dictyostelium discoideum and macrophages in relation with type III secretion system. BMC Microbiol 12:223.

38. Madi A, Lakhdari O, Blottiere HM, Guyard-Nicodeme M, Le Roux K, Groboillot A, Svinareff P, Dore J, Orange N, Feuilloley MG, Connil N. 2010. The clinical Pseudomonas fluorescens MFN1032 strain exerts a cytotoxic effect on epithelial intestinal cells and induces Interleukin-8 via the AP-1 signaling pathway. BMC Microbiol 10:215.

39. Sperandio D, Rossignol G, Guerillon J, Connil N, Orange N, Feuilloley MG, Merieau A. 2010. Cell-associated hemolysis activity in the clinical strain of Pseudomonas fluorescens MFN1032. BMC Microbiol 10:124.

40. Gibaud M, Martin-Dupont P, Dominguez M, Laurentjoye P, Chassaing B, Leng B. 1984. [Pseudomonas fluorescens septicemia following transfusion of contaminated blood]. Presse Med 13:2583-2584.

41. Hsueh PR, Teng LJ, Pan HJ, Chen YC, Sun CC, Ho SW, Luh KT. 1998. Outbreak of Pseudomonas fluorescens bacteremia among oncology patients. J Clin Microbiol 36:2914-2917.

42. Khabbaz RF, Arnow PM, Highsmith AK, Herwaldt LA, Chou T, Jarvis WR, Lerche NW, Allen JR. 1984. Pseudomonas fluorescens bacteremia from blood transfusion. Am J Med 76:62-68.

43. Murray AE, Bartzokas CA, Shepherd AJ, Roberts FM. 1987. Blood transfusion-associated Pseudomonas fluorescens septicaemia: is this an increasing problem? J Hosp Infect 9:243-248.

44. Pittman M. 1953. A study of bacteria implicated in transfusion reactions and of bacteria isolated from blood products. J Lab Clin Med 42:273-288.

45. Scott J, Boulton FE, Govan JR, Miles RS, McClelland DB, Prowse CV. 1988. A fatal transfusion reaction associated with blood contaminated with Pseudomonas fluorescens. Vox Sang 54:201-204.

46. Benito N, Mirelis B, Luz Galvez M, Vila M, Lopez-Contreras J, Cotura A, Pomar V, March F, Navarro F, Coll P, Gurgui M. 2012. Outbreak of Pseudomonas fluorescens bloodstream infection in a coronary care unit. J Hosp Infect 82:286-289.

47. 2005. Pseudomonas bloodstream infections associated with a heparin/saline flush--Missouri, New York, Texas, and Michigan, 2004-2005. MMWR; Morbidity and Mortality Weekly Report 54:269-272.

48. 2006. Update: Delayed onset Pseudomonas fluorescens bloodstream infections after exposure to contaminated heparin flush--Michigan and South Dakota, 2005-2006. MMWR; Morbidity and Mortality Weekly Report 55:961-963.

49. Gershman MD, Kennedy DJ, Noble-Wang J, Kim C, Gullion J, Kacica M, Jensen B, Pascoe N, Saiman L, McHale J, Wilkins M, Schoonmaker-Bopp D, Clayton J, Arduino M, Srinivasan A. 2008. Multistate outbreak of Pseudomonas fluorescens bloodstream infection after exposure to contaminated heparinized saline flush prepared by a compounding pharmacy. Clin Infect Dis 47:1372-1379.

50. Landers CJ, Cohavy O, Misra R, Yang H, Lin YC, Braun J, Targan SR. 2002. Selected loss of tolerance evidenced by Crohn's disease-associated immune responses to auto- and microbial antigens. Gastroenterology 123:689-699.

51. Arnott ID, Landers CJ, Nimmo EJ, Drummond HE, Smith BK, Targan SR, Satsangi J. 2004. Sero-reactivity to microbial components in Crohn's disease is associated with disease severity and progression, but not NOD2/CARD15 genotype. Am J Gastroenterol 99:2376-2384.

52. Iltanen S, Tervo L, Halttunen T, Wei B, Braun J, Rantala I, Honkanen T, Kronenberg M, Cheroutre H, Turovskaya O, Autio V, Ashorn M. 2006. Elevated serum anti-I2 and anti-OmpW antibody levels in children with IBD. Inflamm Bowel Dis 12:389-394.

53. Mow WS, Vasiliauskas EA, Lin YC, Fleshner PR, Papadakis KA, Taylor KD, Landers CJ, Abreu-Martin MT, Rotter JI, Yang H, Targan SR. 2004. Association of antibody responses to microbial antigens and complications of small bowel Crohn's disease. Gastroenterology 126:414-424.

54. Spivak J, Landers CJ, Vasiliauskas EA, Abreu MT, Dubinsky MC, Papadakis KA, Ippoliti A, Targan SR, Fleshner PR. 2006. Antibodies to I2 predict clinical response to fecal diversion in Crohn's disease. Inflamm Bowel Dis 12:1122-1130.

55. Ashorn S, Raukola H, Valineva T, Ashorn M, Wei B, Braun J, Rantala I, Kaukinen K, Luukkaala T, Collin P, Maki M, Iltanen S. 2008. Elevated serum anti-Saccharomyces cerevisiae, anti-I2 and anti-OmpW antibody levels in patients with suspicion of celiac disease. J Clin Immunol 28:486-494.

56. Ashorn S, Valineva T, Kaukinen K, Ashorn M, Braun J, Raukola H, Rantala I, Collin P, Maki M, Luukkaala T, Iltanen S. 2009. Serological responses to microbial antigens in celiac disease patients during a gluten-free diet. J Clin Immunol 29:190-195.

57. Viitasalo L, Niemi L, Ashorn M, Ashorn S, Braun J, Huhtala H, Collin P, Maki M, Kaukinen K, Kurppa K, Iltanen S. 2014. Early microbial markers of celiac disease. J Clin Gastroenterol 48:620-624.

58. Yu JE, De Ravin SS, Uzel G, Landers C, Targan S, Malech HL, Holland SM, Cao W, Harpaz N, Mayer L, Cunningham-Rundles C. 2011. High levels of Crohn's disease-associated anti-microbial antibodies are present and independent of colitis in chronic granulomatous disease. Clin Immunol 138:14-22.

59. Mundwiler ML, Mei L, Landers CJ, Reveille JD, Targan S, Weisman MH. 2009. Inflammatory bowel disease serologies in ankylosing spondylitis patients: a pilot study. Arthritis Res Ther 11:R177.

60. Dickson RP, Erb-Downward JR, Freeman CM, Walker N, Scales BS, Beck JM, Martinez FJ, Curtis JL, Lama VN, Huffnagle GB. 2014. Changes in the lung microbiome following lung transplantation include the emergence of two distinct Pseudomonas species with distinct clinical associations. PLoS One 9:e97214.

61. Diaz PI, Hong BY, Frias-Lopez J, Dupuy AK, Angeloni M, Abusleme L, Terzi E, Ioannidou E, Strausbaugh LD, Dongari-Bagtzoglou A. 2013. Transplantation-associated long-term immunosuppression promotes oral colonization by potentially opportunistic pathogens without impacting other members of the salivary bacteriome. Clin Vaccine Immunol 20:920-930.

62. Patel SK, Pratap CB, Verma AK, Jain AK, Dixit VK, Nath G. 2013. Pseudomonas fluorescens-like bacteria from the stomach: a microbiological and molecular study. World J Gastroenterol 19:1056-1067.

63. Stenhouse MA, Milner LV. 1992. A survey of cold-growing gram-negative organisms isolated from the skin of prospective blood donors. Transfus Med 2:235-237.

64. Wagner J, Short K, Catto-Smith AG, Cameron DJ, Bishop RF, Kirkwood CD. 2008. Identification and characterisation of Pseudomonas 16S ribosomal DNA from ileal biopsies of children with Crohn's disease. PLoS One 3:e3578.

65. Loper JE, Hassan KA, Mavrodi DV, Davis EW, 2nd, Lim CK, Shaffer BT, Elbourne LD, Stockwell VO, Hartney SL, Breakwell K, Henkels MD, Tetu SG, Rangel LI, Kidarsa TA, Wilson NL, van de Mortel JE, Song C, Blumhagen R, Radune D, Hostetler JB, Brinkac LM, Durkin AS, Kluepfel DA, Wechter WP, Anderson AJ, Kim YC, Pierson LS, 3rd, Pierson EA, Lindow SE, Kobayashi DY, Raaijmakers JM, Weller DM, Thomashow LS, Allen AE, Paulsen IT. 2012. Comparative genomics of plant-associated Pseudomonas spp.: insights into diversity and inheritance of traits involved in multitrophic interactions. PLoS Genet 8:e1002784.

66. Mulet M, Lalucat J, Garcia-Valdes E. 2010. DNA sequence-based analysis of the Pseudomonas species. Environ Microbiol 12:1513-1530.

67. van Passel MW, Kuramae EE, Luyf AC, Bart A, Boekhout T. 2006. The reach of the genome signature in prokaryotes. BMC Evol Biol 6:84.

68. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol 57:81-91.

69. Barton MD, Petronio M, Giarrizzo JG, Bowling BV, Barton HA. 2013. The genome of Pseudomonas fluorescens strain R124 demonstrates phenotypic adaptation to the mineral environment. J Bacteriol 195:4793-4803.

70. Kim M, Oh HS, Park SC, Chun J. 2014. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 64:346-351.

71. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403-410.

72. Hohnadel D, Meyer JM. 1988. Specificity of pyoverdine-mediated iron uptake among fluorescent Pseudomonas strains. J Bacteriol 170:4865-4873.

73. Cornelis P, Matthijs S, Van Oeffelen L. 2009. Iron uptake regulation in Pseudomonas aeruginosa. Biometals 22:15-22.

74. Visca P, Imperi F, Lamont IL. 2007. Pyoverdine siderophores: from biogenesis to biosignificance. Trends Microbiol 15:22-30.

75. Wandersman C, Delepelaire P. 2004. Bacterial iron sources: from siderophores to hemophores. Annu Rev Microbiol 58:611-647.

76. Nielsen MN, Sorensen J, Fels J, Pedersen HC. 1998. Secondary metabolite- and endochitinase-dependent antagonism toward plant-pathogenic microfungi of pseudomonas fluorescens isolates from sugar beet rhizosphere. Appl Environ Microbiol 64:3563-3569.

77. Nikaidou N, Kamio Y, Izaki K. 1993. Molecular cloning and nucleotide sequence of the pectate lyase gene from Pseudomonas marginalis N6301. Biosci Biotechnol Biochem 57:957-960.

78. Anderson LM, Stockwell VO, Loper JE. 2004. An Extracellular Protease of Pseudomonas fluorescens Inactivates Antibiotics of Pantoea agglomerans. Phytopathology 94:1228-1234.

79. Stockwell VO, Johnson KB, Sugar D, Loper JE. 2011. Mechanistically compatible mixtures of bacterial antagonists improve biological control of fire blight of pear. Phytopathology 101:113-123.

80. Jousset A, Lara E, Wall LG, Valverde C. 2006. Secondary metabolites help biocontrol strain Pseudomonas fluorescens CHA0 to escape protozoan grazing. Appl Environ Microbiol 72:7083-7090.

81. Duong F, Soscia C, Lazdunski A, Murgier M. 1994. The Pseudomonas fluorescens lipase has a C-terminal secretion signal and is secreted by a three-component bacterial ABC-exporter system. Mol Microbiol 11:1117-1126.

82. Ryu CM, Farag MA, Hu CH, Reddy MS, Wei HX, Pare PW, Kloepper JW. 2003. Bacterial volatiles promote growth in Arabidopsis. Proc Natl Acad Sci U S A 100:4927-4932.

83. Tatusov RL, Koonin EV, Lipman DJ. 1997. A genomic perspective on protein families. Science 278:631-637.

84. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA. 2003. The COG database: an updated version includes eukaryotes. BMC Bioinformatics 4:41.

85. Cha JS, Cooksey DA. 1991. Copper resistance in Pseudomonas syringae mediated by periplasmic and outer membrane proteins. Proc Natl Acad Sci U S A 88:8915-8919.

86. Nies DH. 1995. The cobalt, zinc, and cadmium efflux system CzcABC from Alcaligenes eutrophus functions as a cation-proton antiporter in Escherichia coli. J Bacteriol 177:2707-2712.

87. Nies DH. 2003. Efflux-mediated heavy metal resistance in prokaryotes. FEMS Microbiol Rev 27:313-339.

88. Caille O, Rossier C, Perron K. 2007. A copper-activated two-component system interacts with zinc and imipenem resistance in Pseudomonas aeruginosa. J Bacteriol 189:4561-4568.

89. Delmar JA, Su CC, Yu EW. 2014. Bacterial multidrug efflux transporters. Annu Rev Biophys 43:93-117.

90. Gao F, Zhang CT. 2006. GC-Profile: a web-based tool for visualizing and analyzing the variation of GC content in genomic sequences. Nucleic Acids Res 34:W686-691.

91. Groisman EA, Ochman H. 1996. Pathogenicity islands: bacterial evolution in quantum leaps. Cell 87:791-794.

92. Koonin EV, Makarova KS, Aravind L. 2001. Horizontal gene transfer in prokaryotes: quantification and classification. Annu Rev Microbiol 55:709-742.

93. Juhas M, van der Meer JR, Gaillard M, Harding RM, Hood DW, Crook DW. 2009. Genomic islands: tools of bacterial horizontal gene transfer and evolution. FEMS Microbiol Rev 33:376-393.

94. Tomich M, Planet PJ, Figurski DH. 2007. The tad locus: postcards from the widespread colonization island. Nat Rev Microbiol 5:363-375.

95. Alfano JR, Charkowski AO, Deng WL, Badel JL, Petnicki-Ocwieja T, van Dijk K, Collmer A. 2000. The Pseudomonas syringae Hrp pathogenicity island has a tripartite mosaic structure composed of a cluster of type III secretion genes bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity in plants. Proc Natl Acad Sci U S A 97:4856-4861.

96. Carter MQ, Chen J, Lory S. 2010. The Pseudomonas aeruginosa pathogenicity island PAPI-1 is transferred via a novel type IV pilus. J Bacteriol 192:3249-3258.

97. Hudson CM, Bent ZW, Meagher RJ, Williams KP. 2014. Resistance determinants and mobile genetic elements of an NDM-1-encoding Klebsiella pneumoniae strain. PLoS One 9:e99209.

98. Gillings MR. 2013. Evolutionary consequences of antibiotic use for the resistome, mobilome and microbial pangenome. Front Microbiol 4:4.

99. Juhas M. 2015. Horizontal gene transfer in human pathogens. Crit Rev Microbiol 41:101-108.

100. Buchrieser C, Rusniok C, Kunst F, Cossart P, Glaser P. 2003. Comparison of the genome sequences of Listeria monocytogenes and Listeria innocua: clues for evolution and pathogenicity. FEMS Immunol Med Microbiol 35:207-213.

101. Kung VL, Ozer EA, Hauser AR. 2010. The accessory genome of Pseudomonas aeruginosa. Microbiol Mol Biol Rev 74:621-641.

102. Liu L, Chen H, Brecher MB, Li Z, Wei B, Nandi B, Zhang J, Ling H, Winslow G, Braun J, Li H. 2013. Pfit is a structurally novel Crohn's disease-associated superantigen. PLoS Pathog 9:e1003837.

103. Sutton CL, Kim J, Yamane A, Dalwadi H, Wei B, Landers C, Targan SR, Braun J. 2000. Identification of a novel bacterial sequence associated with Crohn's disease. Gastroenterology 119:23-31.

104. Wei B, Huang T, Dalwadi H, Sutton CL, Bruckner D, Braun J. 2002. Pseudomonas fluorescens encodes the Crohn's disease-associated I2 sequence and T-cell superantigen. Infect Immun 70:6567-6575.

105. Dalwadi H, Wei B, Kronenberg M, Sutton CL, Braun J. 2001. The Crohn's disease-associated bacterial protein I2 is a novel enteric t cell superantigen. Immunity 15:149-158.

106. Dubinsky MC, Lin YC, Dutridge D, Picornell Y, Landers CJ, Farrior S, Wrobel I, Quiros A, Vasiliauskas EA, Grill B, Israel D, Bahar R, Christie D, Wahbeh G, Silber G, Dallazadeh S, Shah P, Thomas D, Kelts D, Hershberg RM, Elson CO, Targan SR, Taylor KD, Rotter JI, Yang H. 2006. Serum immune responses predict rapid disease progression among children with Crohn's disease: immune responses predict disease progression. Am J Gastroenterol 101:360-367.

107. Suzuki H, Fukuda Y, Koizuka H, Tomita T, Hori K, Suzuki M, O'Morain C. 2008. Dietary antigens in Crohn's disease: antibodies to porcine pancreatic amylase. Am J Gastroenterol 103:656-664.

108. Ryan JD, Silverberg MS, Xu W, Graff LA, Targownik LE, Walker JR, Carr R, Clara I, Miller N, Rogala L, Bernstein CN. 2013. Predicting complicated Crohn's disease and

surgery: phenotypes, genetics, serology and psychological characteristics of a population-based cohort. Aliment Pharmacol Ther 38:274-283.

109. Xiong Y, Wang GZ, Zhou JQ, Xia BQ, Wang XY, Jiang B. 2014. Serum antibodies to microbial antigens for Crohn's disease progression: a meta-analysis. Eur J Gastroenterol Hepatol 26:733-742.

110. Eckburg PB, Relman DA. 2007. The role of microbes in Crohn's disease. Clin Infect Dis 44:256-262.

111. Madi A, Svinareff P, Orange N, Feuilloley MG, Connil N. 2010. Pseudomonas fluorescens alters epithelial permeability and translocates across Caco-2/TC7 intestinal cells. Gut Pathog 2:16.

112. Alnabhani Z, Montcuquet N, Biaggini K, Dussaillant M, Roy M, Ogier-Denis E, Madi A, Jallane A, Feuilloley M, Hugot JP, Connil N, Barreau F. 2015. Pseudomonas fluorescens Alters the Intestinal Barrier Function by Modulating IL-1beta Expression Through Hematopoietic NOD2 Signaling. Inflamm Bowel Dis 21:543-555.

113. Cianciotto NP. 2005. Type II secretion: a protein secretion system for all seasons. Trends Microbiol 13:581-588.

114. Peabody CR, Chung YJ, Yen MR, Vidal-Ingigliardi D, Pugsley AP, Saier MH, Jr. 2003. Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. Microbiology 149:3051-3072.

115. Sandkvist M. 2001. Type II secretion and pathogenesis. Infect Immun 69:3523-3535.

116. Filloux A. 2004. The underlying mechanisms of type II protein secretion. Biochim Biophys Acta 1694:163-179.

117. Ball G, Durand E, Lazdunski A, Filloux A. 2002. A novel type II secretion system in Pseudomonas aeruginosa. Mol Microbiol 43:475-485.

118. De Vrind J, De Groot A, Brouwers GJ, Tommassen J, De Vrind-De Jong E. 2003. Identification of a novel Gsp-related pathway required for secretion of the manganese-oxidizing factor of Pseudomonas putida strain GB-1. Mol Microbiol 47:993-1006.

119. da Silva AC, Ferro JA, Reinach FC, Farah CS, Furlan LR, Quaggio RB, Monteiro-Vitorello CB, Van Sluys MA, Almeida NF, Alves LM, do Amaral AM, Bertolini MC, Camargo LE, Camarotte G, Cannavan F, Cardozo J, Chambergo F, Ciapina LP, Cicarelli RM, Coutinho LL, Cursino-Santos JR, El-Dorry H, Faria JB, Ferreira AJ, Ferreira RC, Ferro MI, Formighieri EF, Franco MC, Greggio CC, Gruber A, Katsuyama AM, Kishi LT, Leite RP, Lemos EG, Lemos MV, Locali EC, Machado MA, Madeira AM, Martinez-Rossi NM, Martins EC, Meidanis J, Menck CF, Miyaki CY, Moon DH, Moreira LM, Novo MT, Okura VK, Oliveira MC, Oliveira VR, Pereira HA, Rossi A, Sena JA, Silva C, de Souza RF, Spinola LA, Takita MA, Tamura RE, Teixeira EC, Tezza RI, Trindade dos Santos M, Truffi D, Tsai SM, White FF, Setubal JC, Kitajima JP. 2002. Comparison of the genomes of two Xanthomonas pathogens with differing host specificities. Nature 417:459-463.

120. Parche S, Geissdorfer W, Hillen W. 1997. Identification and characterization of xcpR encoding a subunit of the general secretory pathway necessary for dodecane degradation in Acinetobacter calcoaceticus ADP1. J Bacteriol 179:4631-4634.

121. Salmond GP, Reeves PJ. 1993. Membrane traffic wardens and protein secretion in gram-negative bacteria. Trends Biochem Sci 18:7-12.

122. Lindgren PB, Peet RC, Panopoulos NJ. 1986. Gene cluster of Pseudomonas syringae pv. "phaseolicola" controls pathogenicity of bean plants and hypersensitivity of nonhost plants. J Bacteriol 168:512-522.

123. Marchi M, Boutin M, Gazengel K, Rispe C, Gauthier JP, Guillerm-Erckelboudt AY, Lebreton L, Barret M, Daval S, Sarniguet A. 2013. Genomic analysis of the biocontrol strain Pseudomonas fluorescens Pf29Arp with evidence of T3SS and T6SS gene expression on plant roots. Environ Microbiol Rep 5:393-403.

124. Preston GM, Bertrand N, Rainey PB. 2001. Type III secretion in plant growth-promoting Pseudomonas fluorescens SBW25. Mol Microbiol 41:999-1014.

125. Rezzonico F, Binder C, Defago G, Moenne-Loccoz Y. 2005. The type III secretion system of biocontrol Pseudomonas fluorescens KD targets the phytopathogenic Chromista Pythium ultimum and promotes cucumber protection. Mol Plant Microbe Interact 18:991-1001.

126. Redondo-Nieto M, Barret M, Morrisey JP, Germaine K, Martinez-Granero F, Barahona E, Navazo A, Sanchez-Contreras M, Moynihan JA, Giddens SR, Coppoolse ER, Muriel C, Stiekema WJ, Rainey PB, Dowling D, O'Gara F, Martin M, Rivilla R. 2012. Genome sequence of the biocontrol strain Pseudomonas fluorescens F113. J Bacteriol 194:1273-1274.

127. Ma Q, Zhai Y, Schneider JC, Ramseier TM, Saier MH, Jr. 2003. Protein secretion systems of Pseudomonas aeruginosa and P fluorescens. Biochim Biophys Acta 1611:223-233.

128. Neidig N, Paul RJ, Scheu S, Jousset A. 2011. Secondary metabolites of Pseudomonas fluorescens CHA0 drive complex non-trophic interactions with bacterivorous nematodes. Microb Ecol 61:853-859.

129. Barret M, Egan F, Fargier E, Morrissey JP, O'Gara F. 2011. Genomic analysis of the type VI secretion systems in Pseudomonas spp.: novel clusters and putative effectors uncovered. Microbiology 157:1726-1739.

130. Rainey PB. 1999. Adaptation of Pseudomonas fluorescens to the plant rhizosphere. Environ Microbiol 1:243-257.

131. Jackson RW, Preston GM, Rainey PB. 2005. Genetic characterization of Pseudomonas fluorescens SBW25 rsp gene expression in the phytosphere and in vitro. J Bacteriol 187:8477-8488.

132. Craig L, Li J. 2008. Type IV pili: paradoxes in form and function. Curr Opin Struct Biol 18:267-277.

133. Bradley DE. 1980. A function of Pseudomonas aeruginosa PAO polar pili: twitching motility. Can J Microbiol 26:146-154.

134. Henrichsen J. 1983. Twitching motility. Annu Rev Microbiol 37:81-93.

135. O'Toole GA, Kolter R. 1998. Flagellar and twitching motility are necessary for Pseudomonas aeruginosa biofilm development. Mol Microbiol 30:295-304.

136. Ward MJ, Zusman DR. 1997. Regulation of directed motility in Myxococcus xanthus. Mol Microbiol 24:885-893.

137. Watson AA, Alm RA, Mattick JS. 1996. Identification of a gene, pilF, required for type 4 fimbrial biogenesis and twitching motility in Pseudomonas aeruginosa. Gene 180:49-56.

138. Comolli JC, Hauser AR, Waite L, Whitchurch CB, Mattick JS, Engel JN. 1999. Pseudomonas aeruginosa gene products PilT and PilU are required for cytotoxicity in vitro and virulence in a mouse model of acute pneumonia. Infect Immun 67:3625-3630.

139. D'Argenio DA, Gallagher LA, Berg CA, Manoil C. 2001. Drosophila as a model host for Pseudomonas aeruginosa infection. J Bacteriol 183:1466-1471.

140. Singh PK, Schaefer AL, Parsek MR, Moninger TO, Welsh MJ, Greenberg EP. 2000. Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms. Nature 407:762-764.

141. Whitchurch CB, Tolker-Nielsen T, Ragas PC, Mattick JS. 2002. Extracellular DNA required for bacterial biofilm formation. Science 295:1487.

142. Compeau G, Al-Achi BJ, Platsouka E, Levy SB. 1988. Survival of rifampin-resistant mutants of Pseudomonas fluorescens and Pseudomonas putida in soil systems. Appl Environ Microbiol 54:2432-2438.

143. Mattick JS. 2002. Type IV pili and twitching motility. Annu Rev Microbiol 56:289-314.

144. Planet PJ, Kachlany SC, Fine DH, DeSalle R, Figurski DH. 2003. The Widespread Colonization Island of Actinobacillus actinomycetemcomitans. Nat Genet 34:193-198.

145. Wagner VE, Bushnell D, Passador L, Brooks AI, Iglewski BH. 2003. Microarray analysis of Pseudomonas aeruginosa quorum-sensing regulons: effects of growth phase and environment. J Bacteriol 185:2080-2095.

146. Schuster M, Lostroh CP, Ogi T, Greenberg EP. 2003. Identification, timing, and signal specificity of Pseudomonas aeruginosa quorum-controlled genes: a transcriptome analysis. J Bacteriol 185:2066-2079.

147. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, 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.

148. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res 42:D206-214.

149. Gerlach RG, Hensel M. 2007. Protein secretion systems and adhesins: the molecular armory of Gram-negative pathogens. Int J Med Microbiol 297:401-415.

150. Pukatzki S, Ma AT, Sturtevant D, Krastins B, Sarracino D, Nelson WC, Heidelberg JF, Mekalanos JJ. 2006. Identification of a conserved bacterial protein secretion system in Vibrio cholerae using the Dictyostelium host model system. Proc Natl Acad Sci U S A 103:1528-1533.

151. Mougous JD, Gifford CA, Ramsdell TL, Mekalanos JJ. 2007. Threonine phosphorylation post-translationally regulates protein secretion in Pseudomonas aeruginosa. Nat Cell Biol 9:797-803.

152. Pukatzki S, Ma AT, Revel AT, Sturtevant D, Mekalanos JJ. 2007. Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. Proc Natl Acad Sci U S A 104:15508-15513.

153. Suarez G, Sierra JC, Erova TE, Sha J, Horneman AJ, Chopra AK. 2010. A type VI secretion system effector protein, VgrG1, from Aeromonas hydrophila that induces host cell toxicity by ADP ribosylation of actin. J Bacteriol 192:155-168.

154. Hood RD, Singh P, Hsu F, Guvener T, Carl MA, Trinidad RR, Silverman JM, Ohlson BB, Hicks KG, Plemel RL, Li M, Schwarz S, Wang WY, Merz AJ, Goodlett DR, Mougous JD. 2010. A type VI secretion system of Pseudomonas aeruginosa targets a toxin to bacteria. Cell Host Microbe 7:25-37.

155. MacIntyre DL, Miyata ST, Kitaoka M, Pukatzki S. 2010. The Vibrio cholerae type VI secretion system displays antimicrobial properties. Proc Natl Acad Sci U S A 107:19520-19524.

156. Basler M, Pilhofer M, Henderson GP, Jensen GJ, Mekalanos JJ. 2012. Type VI secretion requires a dynamic contractile phage tail-like structure. Nature 483:182-186.

157. Basler M, Ho BT, Mekalanos JJ. 2013. Tit-for-tat: type VI secretion system counterattack during bacterial cell-cell interactions. Cell 152:884-894.

158. Boyer F, Fichant G, Berthod J, Vandenbrouck Y, Attree I. 2009. Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources? BMC Genomics 10:104.

159. Suarez G, Sierra JC, Sha J, Wang S, Erova TE, Fadl AA, Foltz SM, Horneman AJ, Chopra AK. 2008. Molecular characterization of a functional type VI secretion system from a clinical isolate of Aeromonas hydrophila. Microb Pathog 44:344-361.

160. Kitaoka M, Miyata ST, Brooks TM, Unterweger D, Pukatzki S. 2011. VasH is a transcriptional regulator of the type VI secretion system functional in endemic and pandemic Vibrio cholerae. J Bacteriol 193:6471-6482.

161. Bernard CS, Brunet YR, Gavioli M, Lloubes R, Cascales E. 2011. Regulation of type VI secretion gene clusters by sigma54 and cognate enhancer binding proteins. J Bacteriol 193:2158-2167.

162. Doucleff M, Pelton JG, Lee PS, Nixon BT, Wemmer DE. 2007. Structural basis of DNA recognition by the alternative sigma-factor, sigma54. J Mol Biol 369:1070-1078.

163. Potvin E, Sanschagrin F, Levesque RC. 2008. Sigma factors in Pseudomonas aeruginosa. FEMS Microbiol Rev 32:38-55.

164. Leung KY, Siame BA, Snowball H, Mok YK. 2011. Type VI secretion regulation: crosstalk and intracellular communication. Curr Opin Microbiol 14:9-15.

165. Hassan MT, van der Lelie D, Springael D, Romling U, Ahmed N, Mergeay M. 1999. Identification of a gene cluster, czr, involved in cadmium and zinc resistance in Pseudomonas aeruginosa. Gene 238:417-425.

166. Perron K, Caille O, Rossier C, Van Delden C, Dumas JL, Kohler T. 2004. CzcR-CzcS, a two-component system involved in heavy metal and carbapenem resistance in Pseudomonas aeruginosa. J Biol Chem 279:8761-8768.

167. Goldberg M, Pribyl T, Juhnke S, Nies DH. 1999. Energetics and topology of CzcA, a cation/proton antiporter of the resistance-nodulation-cell division protein family. J Biol Chem 274:26065-26070.

168. Blindauer CA, Harrison MD, Parkinson JA, Robinson AK, Cavet JS, Robinson NJ, Sadler PJ. 2001. A metallothionein containing a zinc finger within a four-metal cluster protects a bacterium from zinc toxicity. Proc Natl Acad Sci U S A 98:9593-9598.

169. Tamber S, Ochs MM, Hancock RE. 2006. Role of the novel OprD family of porins in nutrient uptake in Pseudomonas aeruginosa. J Bacteriol 188:45-54.

170. Alonso A, Rojo F, Martinez JL. 1999. Environmental and clinical isolates of Pseudomonas aeruginosa show pathogenic and biodegradative properties irrespective of their origin. Environ Microbiol 1:421-430.

171. Wolfgang MC, Kulasekara BR, Liang X, Boyd D, Wu K, Yang Q, Miyada CG, Lory S. 2003. Conservation of genome content and virulence determinants among clinical and environmental isolates of Pseudomonas aeruginosa. Proc Natl Acad Sci U S A 100:8484-8489.

172. McArthur JV, Tuckfield RC. 2000. Spatial patterns in antibiotic resistance among stream bacteria: effects of industrial pollution. Appl Environ Microbiol 66:3722-3726.

173. Rasmussen LD, Sorensen SJ. 1998. The effect of longterm exposure to mercury on the bacterial community in marine sediment. Curr Microbiol 36:291-297.

174. Stepanauskas R, Glenn TC, Jagoe CH, Tuckfield RC, Lindell AH, McArthur JV. 2005. Elevated microbial tolerance to metals and antibiotics in metal-contaminated industrial environments. Environ Sci Technol 39:3671-3678.

175. Flume PA, Mogayzel PJ, Jr., Robinson KA, Goss CH, Rosenblatt RL, Kuhn RJ, Marshall BC. 2009. Cystic fibrosis pulmonary guidelines: treatment of pulmonary exacerbations. Am J Respir Crit Care Med 180:802-808.

176. Khanbabaee G, Akbarizadeh M, Sayyari A, Ashayeri-Panah M, Abdollahgorji F, Sheibani K, Rezaeig N. 2012. A survey on pulmonary pathogens and their antibiotic susceptibility among cystic fibrosis patients. Braz J Infect Dis 16:122-128.

177. Cullen L, McClean S. 2015. Bacterial Adaptation during Chronic Respiratory Infections. Pathogens 4:66-89.

178. McFadden ER, Jr., Pichurko BM, Bowman HF, Ingenito E, Burns S, Dowling N, Solway J. 1985. Thermal mapping of the airways in humans. J Appl Physiol (1985) 58:564-570.

179. Feazel LM, Baumgartner LK, Peterson KL, Frank DN, Harris JK, Pace NR. 2009. Opportunistic pathogens enriched in showerhead biofilms. Proc Natl Acad Sci U S A 106:16393-16399.

180. Hirvonen MR, Huttunen K, Roponen M. 2005. Bacterial strains from moldy buildings are highly potent inducers of inflammatory and cytotoxic effects. Indoor Air 15 Suppl 9:65-70.

181. Bowers RM, Sullivan AP, Costello EK, Collett JL, Jr., Knight R, Fierer N. 2011. Sources of bacteria in outdoor air across cities in the midwestern United States. Appl Environ Microbiol 77:6350-6356.

182. Silby MW, Winstanley C, Godfrey SA, Levy SB, Jackson RW. 2011. Pseudomonas genomes: diverse and adaptable. FEMS Microbiol Rev 35:652-680.

183. Barbier M, Damron FH, Bielecki P, Suarez-Diez M, Puchalka J, Alberti S, Dos Santos VM, Goldberg JB. 2014. From the environment to the host: re-wiring of the transcriptome of Pseudomonas aeruginosa from 22 degrees C to 37 degrees C. PLoS One 9:e89941.

184. Shigemura K, Osawa K, Kato A, Tokimatsu I, Arakawa S, Shirakawa T, Fujisawa M. 2015. Association of overexpression of efflux pump genes with antibiotic resistance in Pseudomonas aeruginosa strains clinically isolated from urinary tract infection patients. J Antibiot (Tokyo).

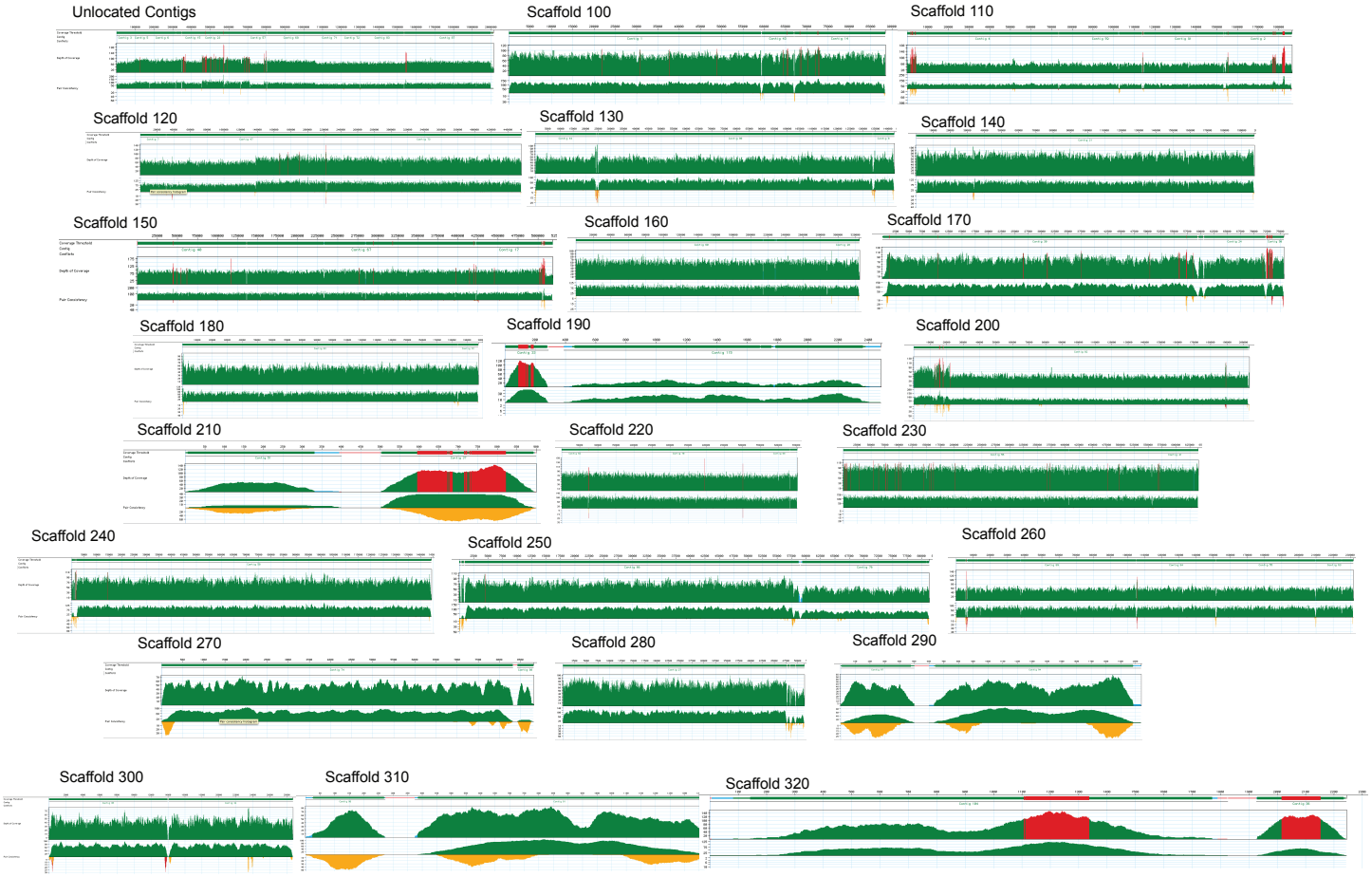
185. Ciofu O, Giwercman B, Pedersen SS, Hoiby N. 1994. Development of antibiotic resistance in Pseudomonas aeruginosa during two decades of antipseudomonal treatment at the Danish CF Center. Apmis 102:674-680.

186. Islam S, Oh H, Jalal S, Karpati F, Ciofu O, Hoiby N, Wretlind B. 2009. Chromosomal mechanisms of aminoglycoside resistance in Pseudomonas aeruginosa isolates from cystic fibrosis patients. Clin Microbiol Infect 15:60-66.

187. Jalal S, Ciofu O, Hoiby N, Gotoh N, Wretlind B. 2000. Molecular mechanisms of fluoroquinolone resistance in Pseudomonas aeruginosa isolates from cystic fibrosis patients. Antimicrob Agents Chemother 44:710-712.

188. Breidenstein EB, de la Fuente-Nunez C, Hancock RE. 2011. Pseudomonas aeruginosa: all roads lead to resistance. Trends Microbiol 19:419-426.

Appendix Figure IV.1: Coverage map of AU2989 assembled scaffolds and contigs.



Appendix Figure IV.1: Coverage map of AU2989 assembled scaffolds and contigs.
Genomic coverage across each scaffold in the de novo assembled genome of AU2989. Across the top x-axis is genomic coverage and across the bottom x-axis is pair consistency between the paired Illumina sequenced reads. Areas of thin red lines indicate areas where there was only one sequence pair covering that genomic area. Areas of orange indicate areas of pair inconsistency. Areas with no lines indicate gaps in coverage. Assembly performed by DNAstar SeqMan NGen. Assembly viewed by DNAstar SeqMan Pro.

Appendix Figure IV.2: The nucleotide sequence of the 16S rRNA sequences was used to query the NCBI nucleotide collection (nr/nt).

Megablast results based on A506 16S	Sequence Identity	Query Coverage	Gaps (%)	E value	Strand	Subject	
AU2989	98.56	90.04	0.29	0.00	Plus/Plus	116215	117597
	97.81	54.22	0.49	0.00	Plus/Plus	2282213	2283032
AU6026	99.61	67.61	0.00	0.00	Plus/Plus	61909809	6110833
	99.55	43.73	0.00	0.00	Plus/Minus	5361468	5360803
	99.67	39.97	0.00	0.00	Plus/Minus	4831512	4830907
	99.62	34.30	0.00	0.00	Plus/Plus	2689341	2689860
	100.00	11.74	0.00	0.00	Plus/Plus	6110834	61110111
AU10973	98.13	45.84	0.00	0.00	Plus/Plus	6127782	6128476
	97.69	37.14	1.07	0.00	Plus/Minus	5379218	5378659
AU11518	98.56	82.26	0.16	0.00	Plus/Plus	253190	254435
	97.98	49.08	0.27	0.00	Plus/Plus	2200107	2200849
	97.50	37.01	0.71	0.00	Plus/Plus	837685	838243
AU14440	99.90	66.62	0.00	0.00	Plus/Minus	72358	713449
	99.82	37.40	0.00	0.00	Plus/Minus	1100141	1099575
	99.82	36.28	0.00	0.00	Plus/Plus	283458	284007
	99.31	28.50	0.00	0.00	Plus/Minus	71248	70817
AU14705	98.73	99.01	0.00	0.00	Plus/Minus	515809	514309
	98.18	61.54	0.00	0.00	Plus/Minus	1479613	1478681
	98.09	51.78	0.00	0.00	Plus/Plus	4998483	4999267
	99.63	17.88	0.00	0.00	Plus/Plus	1329384	1329654
	93.90	14.05	0.94	0.00	Plus/Plus	1329755	1329966
AU14917	99.83	76.06	0.00	0.00	Plus/Minus	333395	332243
	99.91	69.53	0.00	0.00	Plus/Minus	848812	5719589
	100.00	18.54	0.00	0.00	Plus/Minus	3732400	3732120
P. fluorescens SBW25	98.81	100.00	0.13	0.00	Plus/Plus	119733	121248
	98.81	100.00	0.13	0.00	Plus/Plus	768545	770060
	98.81	100.00	0.13	0.00	Plus/Minus	5722246	5720731
	98.81	100.00	0.13	0.00	Plus/Minus	6073701	6072186

Appendix Figure IV.3: Average nucleotide identity (ANI).

	AU2989	AU6026	AU10973	AU11518	AU14440	AU14705	AU14917	SBW25	SBW25_UM	A506	SS101	BG33R	PAO1
AU2989	100												
AU6026	86.04	100											
AU10973	97.75	86.75	100										
AU11518	86.45	86.22	86.46	100									
AU14440	86.36	84.67	86.89	86.34	100								
AU14705	86.10	89.31	87.08	86.04	89.49	100							
AU14917	86.12	94.82	86.81	86.08	96.15	89.42	100						
SBW25	88.42	88.66	91.77	88.39	88.84	89.14	88.86	100					
SBW25 _UM	86.57	86.93	90.45	86.58	87.05	87.41	87.05	99.99	100				
A506	87.62	96.16	88.49	87.86	95.37	90.85	95.80	88.80	88.73	100			
SS101	86.00	94.67	85.99	86.84	96.09	89.51	98.89	88.92	87.14	95.78	100		
BG33R1	85.82	89.51	86.72	85.88	89.45	89.75	89.61	88.70	86.93	91.10	89.69	100	
PAO1	78.55	78.45	78.62	78.56	78.70	78.53	78.54	78.70	78.64	78.81	78.74	78.61	100

Appendix Figure IV.3: Average nucleotide identity (ANI).
Shared average nucleotide identity between each strain in subclade III was determined using the online tool at http://enve-omics.ce.gatech.edu/ani [68]. Values presented in % shared ANI and ≥95% shadowed in gray.

Appendix Figure IV.4: Secondary metabolite genes and reference organism for blast.

Gene	Source Organism	NCBI ID
DAPG		
phlD	Pseudomonas fluorescens F113	NC_016830.1
HCN		
hcnA	Pseudomonas fluorescens Pf01	NC_007492.2
hcnB	Pseudomonas fluorescens Pf01	NC_007492.2
hcnC	Pseudomonas fluorescens Pf01	NC_007492.2
Phenazine		
phzA	Pseudomonas fluorescens S2P5	AY960782.1
phzB	Pseudomonas fluorescens S2P5	AY960782.1
phzC	Pseudomonas fluorescens S2P5	AY960782.1
phzD	Pseudomonas fluorescens S2P5	AY960782.1
phzE	Pseudomonas fluorescens S2P5	AY960782.1
phzF	Pseudomonas fluorescens S2P5	AY960782.1
phzG	Pseudomonas fluorescens S2P5	AY960782.1
Pyrrolnitrin		
prnA	Pseudomonas fluorescens Pf5	NC_004129.6
prnB	Pseudomonas fluorescens Pf5	NC_004129.6
prnC	Pseudomonas fluorescens Pf5	NC_004129.6
prnD	Pseudomonas fluorescens Pf5	NC_004129.6
Rhizoxins		
rhiA	Ralstonia solanacearum CFBP2957	NC_014307.1
rhiB	Ralstonia solanacearum CFBP2957	NC_014307.1
rhiC	Ralstonia solanacearum CFBP2957	NC_014307.1
rhiD	Ralstonia solanacearum CFBP2957	NC_014307.1
rhiE	Ralstonia solanacearum CFBP2957	NC_014307.1
rhiF	Ralstonia solanacearum CFBP2957	NC_014307.1
rhiG	Ralstonia solanacearum CFBP2957	NC_014307.1
rhiH	Ralstonia solanacearum CFBP2957	NC_014307.1
rhiI	Ralstonia solanacearum CFBP2957	NC_014307.1
Pyoluteorin		
pltM	Pseudomonas fluorescens Pf5	NC_004129.6
pltR	Pseudomonas fluorescens Pf5	NC_004129.6
pltL	Pseudomonas fluorescens Pf5	NC_004129.6
pltA	Pseudomonas fluorescens Pf5	NC_004129.6
pltB	Pseudomonas fluorescens Pf5	NC_004129.6
pltC	Pseudomonas fluorescens Pf5	NC_004129.6
pltD	Pseudomonas fluorescens Pf5	NC_004129.6
pltE	Pseudomonas fluorescens Pf5	NC_004129.6
pltF	Pseudomonas fluorescens Pf5	NC_004129.6
pltG	Pseudomonas fluorescens Pf5	NC_004129.6
pltZ	Pseudomonas fluorescens Pf5	NC_004129.6
pltI	Pseudomonas fluorescens Pf5	NC_004129.6
pltJ	Pseudomonas fluorescens Pf5	NC_004129.6
pltK	Pseudomonas fluorescens Pf5	NC_004129.6
pltN	Pseudomonas fluorescens Pf5	NC_004129.6
pltO	Pseudomonas fluorescens Pf5	NC_004129.6
pltP	Pseudomonas fluorescens Pf5	NC_004129.6
Mupirocin		
mupZ	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupA	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mmpA	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupB	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mmpB	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mmpC	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mmpD	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupC	Pseudomoans fluorescens NCIMB 10586	AF318063.3
macpA	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupD	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupE	Pseudomoans fluorescens NCIMB 10586	AF318063.3
macpB	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupF	Pseudomoans fluorescens NCIMB 10586	AF318063.3

macpC	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupC	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupG	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupH	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupJ	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupK	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mmpE	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupL	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupM	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupN	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupO	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupP	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupQ	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupS	Pseudomoans fluorescens NCIMB 10586	AF318063.3
macpD	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mmpF	Pseudomoans fluorescens NCIMB 10586	AF318063.3
macpE	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupT	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupU	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupV	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupW	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupR	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupX	Pseudomoans fluorescens NCIMB 10586	AF318063.3
mupI	Pseudomoans fluorescens NCIMB 10586	AF318063.3
Pyoverdine		
ppyR	Pseudomonas aeruginosa PAO1	NC_002516.2
pvdN	Pseudomonas aeruginosa PAO1	NC_002516.2
pvdE	Pseudomonas aeruginosa PAO1	NC_002516.2
pvdR	Pseudomonas aeruginosa PAO1	NC_002516.2
pvdT	Pseudomonas aeruginosa PAO1	NC_002516.2
pvdF	Pseudomonas aeruginosa PAO1	NC_002516.2
pvdG	Pseudomonas aeruginosa PAO1	NC_002516.2
pvdP	Pseudomonas aeruginosa PAO1	NC_002516.2
pvdO	Pseudomonas aeruginosa PAO1	NC_002516.2
pvdD	Pseudomonas aeruginosa PAO1	NC_002516.2
pvdJ	Pseudomonas aeruginosa PAO1	NC_002516.2
pvdS	Pseudomonas aeruginosa PAO1	NC_002516.2
Pseudomonine		
pmsC	Pseudomonas fluorescens	NC_017911.1
pmsE	Pseudomonas fluorescens	NC_017911.1
pmsA	Pseudomonas fluorescens	NC_017911.1
pmsB	Pseudomonas fluorescens	NC_017911.1
Enantio-pyochelin		
pchC	Pseudomonas fluorescens Pf-5	NC_004129.6
pchF	Pseudomonas fluorescens Pf-5	NC_004129.6
pchE	Pseudomonas fluorescens Pf-5	NC_004129.6
fetA	Pseudomonas fluorescens Pf-5	NC_004129.6
Achromobactin		
acsD	Pseudomonas syringae pv. phaseolicola 1448A	NC_005773.3
acsC	Pseudomonas syringae pv. phaseolicola 1448A	NC_005773.3
acsB	Pseudomonas syringae pv. phaseolicola 1448A	NC_005773.3
acsA	Pseudomonas syringae pv. phaseolicola 1448A	NC_005773.3
cbrA	Pseudomonas syringae pv. phaseolicola 1448A	NC_005773.3
crbB	Pseudomonas syringae pv. phaseolicola 1448A	NC_005773.3
crbC	Pseudomonas syringae pv. phaseolicola 1448A	NC_005773.3
crbD	Pseudomonas syringae pv. phaseolicola 1448A	NC_005773.3
Hemophore		
hasA	Pseudomonas protegens CHA0	NC_021237.1
Chitanse		
chiC	Pseudomonas fluorescens A506	NC_017911.1
AprA		
aprA	Pseudomonas fluorescens A506	NC_017911.1
Pectate Lyase		
pectate lyase	Pseudomonas fluorescens A506	NC_017911.1

IAA Biosynthesis and Catabolism		
iaaM	Pseudomonas protegens CHA0	NC_021237.1
iaaH	Agrobacterium tumefaciens	NC_002377.1
iacA	Pseudomonas putida 1290	NC_002947.3
iacC	Pseudomonas putida 1290	NC_002947.3
iacD	Pseudomonas putida 1290	NC_002947.3
iacE	Pseudomonas putida 1290	NC_002947.3
iacF	Pseudomonas putida 1290	NC_002947.3
iacG	Pseudomonas putida 1290	NC_002947.3
iacR	Pseudomonas putida 1290	NC_002947.3
iacH	Pseudomonas putida 1290	NC_002947.3
iacI	Pseudomonas putida 1290	NC_002947.3
PAA catabolism		
ppaX	Pseudomonas protegens Pf-5	NC_004129.6
paaY	Pseudomonas protegens Pf-5	NC_004129.6
paaB	Pseudomonas protegens Pf-5	NC_004129.6
paaD	Pseudomonas protegens Pf-5	NC_004129.6
paaF	Pseudomonas protegens Pf-5	NC_004129.6
paaG	Pseudomonas protegens Pf-5	NC_004129.6
paaH	Pseudomonas protegens Pf-5	NC_004129.6
paaI	Pseudomonas protegens Pf-5	NC_004129.6
paaJ	Pseudomonas protegens Pf-5	NC_004129.6
paaK	Pseudomonas protegens Pf-5	NC_004129.6
paaW	Pseudomonas protegens Pf-5	NC_004129.6
paaL	Pseudomonas protegens Pf-5	NC_004129.6
paaN	Pseudomonas protegens Pf-5	NC_004129.6
ACC deaminase		
acdS	Pseudomonas fluorescens F113	NC_016830.1
budC	Pseudomonas fluorescens A506	NC_017911.1
ydjL	Pseudomonas aeruginosa NCGM2.S1	NC_017549.1
2,3-bd biosynthesis		
ilvB	Pseudomonas fluroescens A506	NC_017911.1
ilvN	Pseudomonas fluroescens A506	NC_017911.1
Acetoin catabolism (LP) and (DP)		
acoR	Pseudomonas fluroescens A506	NC_017911.1
acoA	Pseudomonas fluroescens A506	NC_017911.1
acoB	Pseudomonas fluroescens A506	NC_017911.1
acoC	Pseudomonas fluroescens A506	NC_017911.1
acoX	Pseudomonas protegens Pf-5	NC_004129.6
bdhA	Pseudomonas protegens Pf-5	NC_004129.6

Appendix Figure IV.4: Secondary metabolite genes and reference organism for blast.
Each gene involved in the production of the secondary metabolites analyzed in **Figure IV.2** is listed in the left column. The source organism of each gene's nucleotide sequence used for the blast search is in the middle column. The NCBI ID for each source organism is in the right column.

Appendix Figure IV.5: Blast results of selected GC Islands.

Isolate	GC Island	GC(%)	Start	End	Size (nt)	Top Hits	E value	NCBI ID
AU2989	1	46.61	102157	127677	25520	Pseudomonas sp. TKP	0.00E+00	CP006852.1
						Pseudomonas sp. WCS374	5.00E-173	CP007638.1
						Pseudomonas fluorescens A506	5.00E-173	CP003041.1
						Pseudomonas fluorescens SBW25	1.00E-168	AM181176.4
						Pseudomonas simiae strain WCS417	3.00E-155	CP007637.1
						Pseudomonas fluorescens strain UK4	1.00E-138	3.00E-105
Isolate	GC Island	GC(%)	Start	End	Size (nt)	Top Hits	E value	NCBI ID
AU6026	1	44.58	1911471	1920380	8911	Pseudomonas sp. VLB120	0.00E+00	CP003961.1
						Pseudomonas alkylphenolia strain KL28	0.00E+00	CP009048.1
						Pseudmonas simiae strain WCS417	0.00E+00	CP007637.1
						Pseudomonas sp. WCS374	0.00E+00	CP007638.1
						Pseudomonas fluorescens A506	0.00E+00	CP003041.1
						Pseudomonas fluorescens SBW25	0.00E+00	AM181176.4
AU6026	2	37.5	2438921	2445960	7040	Pseudomonas fluorescesn A506	0.00E+00	CP003041.1
						Pseudomonas sp. WCS374	0.00E+00	CP007638.1
						Pseudomonas sp. TKP	2.00E-166	CP006852.1
						Pseudomonas simiae strain WCS417	2.00E-151	CP007637.1
						Pseudomonas fluorescens SBW25	5.00E-148	AM181176.4
						Pseudomonas poae RE*1-1-14	2.00E-136	CP004045.1
Isolate	GC Island	GC (%)	Start	End	Size (nt)	Top Hits	E value	NCBI ID
AU11518	1	41.9	2352011	2357520	4511	Pseudomonas fluorescens SBW25	6.00E-51	AM181176.4
						Pseudomonas sp. TKP	2.00E-45	CP006852.1
						Pseudomonas poae RE*1-1-14	1.00E-43	CP004045.1
						Pseudomonas fluorescens A506	5.00E-42	CP003041.1
						Pseudomonas sp. WCS374	1.00E-32	CP007638.1
						Pandoraea apista strain TF81F4	1.00E-27	CP010518.1
AU11518	2	45.56	2876941	287690	2751	Pseudomonas sp. VLB120	0.00E+00	CP003961.1
						Pseudomonas stutzeri DSM 4166	0.00E+00	CP002622.1
						Pseudomonas aeruginosa PA38182	0.00E+00	HG530068.1
						Klebesiella pneumoniae plasmid pKPN_CZ	0.00E+00	JX424424.1
						Enterobacter cloacae strain ECIH5 plasmid pENT-22e	0.00E+00	CP009855.1
						Enterobacter cloacae ECR091 plasmid pENT-4bd	0.00E+00	CP008907.1
AU11518	3	38.07	4641231	4643981	2750	Pseudomonas sp. TKP	1.00E-86	CP006852.1
						Pseudomonas sp. WCS374	8.00E-26	CP007638.1
						Pseudomonas fluorescens A506	8.00E-26	CP003041.1
						Pseudomonas fluorescens SBW25	2.00E-08	AM181176.4
						Pseudomonas simiae strain WCS417	1.00E-04	CP007637.1
Isolate	GC Island	GC (%)	Start	End	Size (nt)	Top Hits	E value	NCBI ID
AU14440	1	42.59	373671	390170	16501	Pseudomonas fluorescens A506	0.00E+00	CP003041.1
						Burkholderia pseudomallei TSV 48 chromosome 2	0.00E+00	CP009160.1
						Burkholderia pseudomallei NAU35A-3 chromosome 2	0.00E+00	CP004378.1
						Pseudomonas aeruginosa YL84	0.00E+00	CP007147.1
						Pseudomonas putida S12	0.00E+00	CP009974.1
						Pseudomonas sp. WCS374	0.00E+00	CP007638.1
AU14440	2	43.07	419101	430541	11440	Pseudomonas aeruginosa plasmid pUM505	0.00E+00	HM56097.1
						Pseudomonas fluorescens A506	0.00E+00	CP003041.1
						Pseudomonas aeruginosa PA38182	0.00E+00	HG530068.1
						Pseudomonas fluorescens SBW25	0.00E+00	AM181176.4
						Pseudomonas plecoglossicida strain NyZ12	0.00E+00	CP010359.1
						Pseudomonas aeruginosa PA96	0.00E+00	CP007224.1
Isolate	GC Island	GC (%)	Start	End	Size (nt)	Top Hits	E value	NCBI ID
AU14705	1	41.91	3897300	3900490	3190	Pseudomonas poae RE*1-1-14	9.00E-134	CP004045.1
						Pseudomonas fluorescens SBW25	3.00E-89	AM181176.4
						Pseudomonas sp. TKP	6.00E-79	CP006852.1
						Pseudomonas fluorescens A506	3.00E-76	CP003041.1
						Pseudomonas sp. WCS374	4.00E-75	CP007638.1
						Pseudomonas simiae strain WCS417	8.00E-52	CP007637.1

Isolate	GC Island	GC (%)	Start	End	Size (nt)	Top Hits	E value	NCBI ID
AU14917	1	39.52	1	1870	1871	Burkholderia ambifaria MC40-6 chromosome 2	2.00E-07	CP001026.1
AU14917	2	43.06	2480830	2489410	8691	Pseudomonas sp. WCS374	0.00E+00	CP007638.1
						Pseudomonas fluorescens A506	0.00E+00	CP003041.1
						Pseudomonas poae RE*1-1-14	5.00E-180	CP004045.1
						Pseudomonas simiae strain WCS417	3.00E-176	CP007637.1
						Pseudomonas fluorescens SBW25	7.00E-172	AM181176.1
AU14917	3	40.81	5618140	5620230	2200	Pseudomonas sp. TKP	1.00E-73	CP006852.1
						Pseudomonas resinovorans NBRC 106553	1.00E-73	AP013068.1
						Pseudomonas stutzeri CCUG 29243	1.00E-73	CP003677.1
						Pseudomonas stutzeri strain 36N1 , insertion sequence ISPst11 TnpA (tnpA5) gene	1.00E-72	GQ221266.1
						Pseudomonas stutzeri strain	1.00E-72	EF648211.1
						Pseudomonas aeruginosa strain SG17M	1.00E-72	AF440524.1
AU14917	4	41.6	6256031	6262850	6820	Pseudomonas fluorescens A506	3.00E-143	CP003041.1
						Pseudomonas sp. WCS374	5.00E-134	CP007638.1
						Pseudomonas fluorescens strain UK4	4.00E-97	CP008896.1
						Pseudomonas putida 0CT plasmid alk genes cluster (alkBFGHJKL, alkN and alkST genes)	7.00E-12	NG_035191.1
						Pseudomonas mandelii JR-1	2.00E-06	CP005960.1
						Pseudomonas brassicacearum subsp. brassicacearum NFM421	7.00E-06	CP002585.1

Appendix Figure IV.5: Blast results of selected GC Islands.
The nucleotide sequence of selected GC Islands (indicated with arrows in **Figure IV.8**) were used to query the NCBI nucleotide collection (nr/nt). The top six results are displayed.

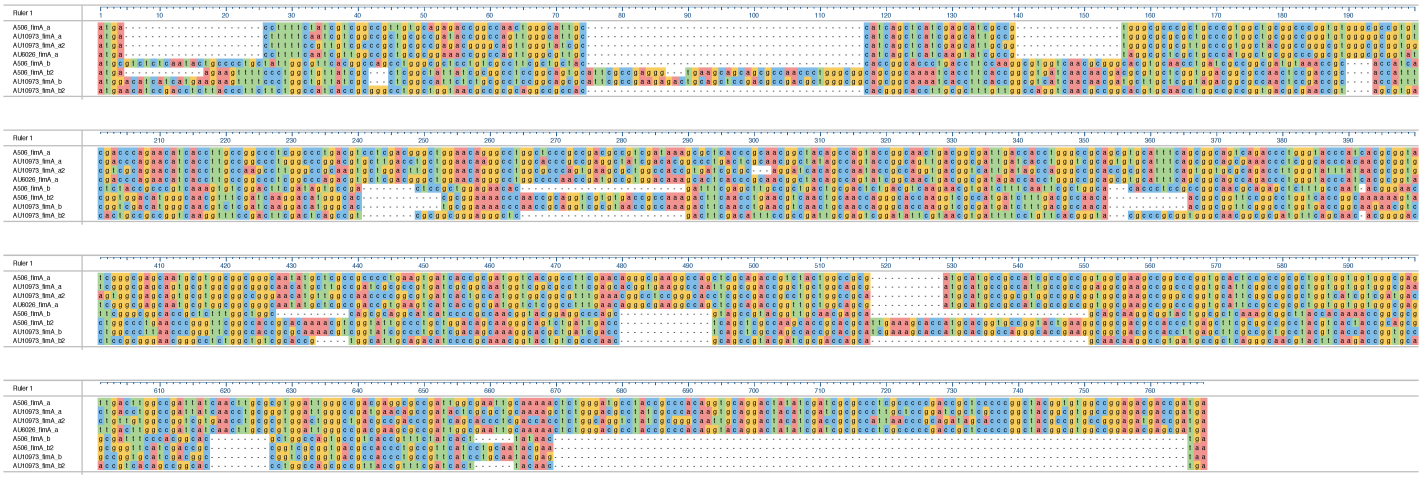
Appendix Figure IV.6: Full annotation of genes involved in secretion systems in subclade III strains.

	RAST Annotation
Type II Secretion	Type II Secretion
GspC	General Secretion Pathway protein C, GspC
GspD	General Secretion Pathway protein D, GspD
GspE	General Secretion Pathway protein E, GspE
GspF	General Secretion Pathway protein F, GspF
GspG	General Secretion Pathway protein G, GspG
GspH	General Secretion Pathway protein H, GspH
GspI	General Secretion Pathway protein I, GspI
GspJ	General Secretion Pathway protein J, GspJ
GspK	General Secretion Pathway protein K, GspK
GspL	General Secretion Pathway protein L, GspL
GspM	General Secretion Pathway protein M, GspM
GspN	General Secretion Pathway protein N, GspN
Type III Secretion (Hrp1)	Type III Secretion (Hrp1)
HrpA	Type III secretion pilin HrpA
HrpB	Type III secretion protein (Pto) HrpB
HrpD	Type III secretion protein HrpD
HrpG	Type II secretion protein HrpG
HrpJ	Type III secretion protein HrpJ
HrpP	Type III secretion protein HrpP
HrpQ	Type III secreiton component protein HrpQ
HrpT	Type III secretion protein HrpT
HrpV	Negative regulator of hrp expression HrpV
HopPmaJ	Type III effector HopPmaJ
Can Hop	Candidate Type III Hop effector
Type IV Secretion	Type IV Secretion
PIIA	Type IV pilin, PilA
PilB	Type II secretory pathway, ATPase PulE/Tfp pilus assembly pathway, ATPase PilB
PilC	Type IV fimbrial assembly protein, PilC
PilD	Type IV pilus biogenesis protein, PilD
PilE	Type IV pilus biogenesis protein, PilE
PilF	Type IV pilus biogenesis protein, PilF
PilG	Twitching motility protein, PilG
PilH	Twitching motility protein, PilH
PIII	Type IV pilus biogenesis protein, PilI
PilJ	Type IV pilus biogenesis protein, PilJ
PIIL	Type IV pilus biogenesis protein, PilL
PilM	Type IV pilus biogenesis protein, PIIM
PIIN	Type IV pilus biogenesis protein, PIIN
PilO	Type IV pilus biogenesis protein, PilO
PIIQ	Type IV pilus biogenesis, PilQ
PilR	Type IV pilus biogenesis, PilR
PilT	Twitching motility protein PilT
PilV	Type IV fimbrial biogenesis protein, PilV
PIIW	Type IV fimbrial biogenesis protein, PIIW
PilX	Type IV fimbrial biogenesis protein, PilX
Adhesin	Adhesin major subunit pilin, adhesion
FimA	Type IV fimbria major subunit, FimA
FimD	Outer membrane usher protein, FimD
FimT	Type IV fimbrial biogenesis protein, FimT
FimV	Probably type IV pilus assembly FimV-related transmembrane protein

Widspread Colonizing Island (WCI)	Widspread Colonizing Island (WCI)
TadV/CpaA	Type IV prepillin peptidase,TadV/CpaA
RcpC/CpaB	Flp pilus assembly protein RcpC/CpaB
RcpA/CpaC	Type II/IV secretion system secretin RcpA/CpaC, assocaited with Flp pilus assembly
TadZ/CpaE	Type II/IV secretion system ATPase TadZ/CpaE, associated with Flp pilus assembly
TadA/CpaF	Type II/IV secretion system ATP hydrolase TadA/VirB11/CpaF, TadA subfamily
TadB	Flp pilus assembly protein TadB
TadC	Type II/IV secretion system protein TadC, associated Flp pilus assembly
TadD	Flp pilus assembly protein TadD, contains TPR repeats
TadG	Flp pilus assembly protein TadG
Type VI Secretion System	Type VI Secretion System
ImpA	Uncharacterized protein ImpA
ImpB	Uncharacterized protein ImpB
ImpC	Uncharacterized protein ImpC
ImpD	Uncharacterized protein ImpD
ImpF	Uncharacterized protein ImpF
ImpG	Protein ImpG/VasA
ImpH	Uncharacterized protein ImpH/VasB
ImpI	Uncharacterized protein ImpI/VasC
ImpJ	Uncharacterized protein ImpJ/VasE
ImpK	Uncharacterized protein ImpK/VasF/OmpA/MotB domain
ImpM	Protein phosphatase ImpM
VasD	Type IV secretion lipoprotein/VasD
VasH	Sigma-54 dependent transcriptional regulator
PppA	Phosphoprotein phosphatase PppA
PpkA	Serine/threonine protein kinase (EC 2.7.11.1) PpkA
Hcp	Secreted protein Hcp
IcmF	IcmF-related protein
Pvc109	Uncharacterized protein similar to VCA0109
ClpB	ClpB protein
VgrG	VgrG protien

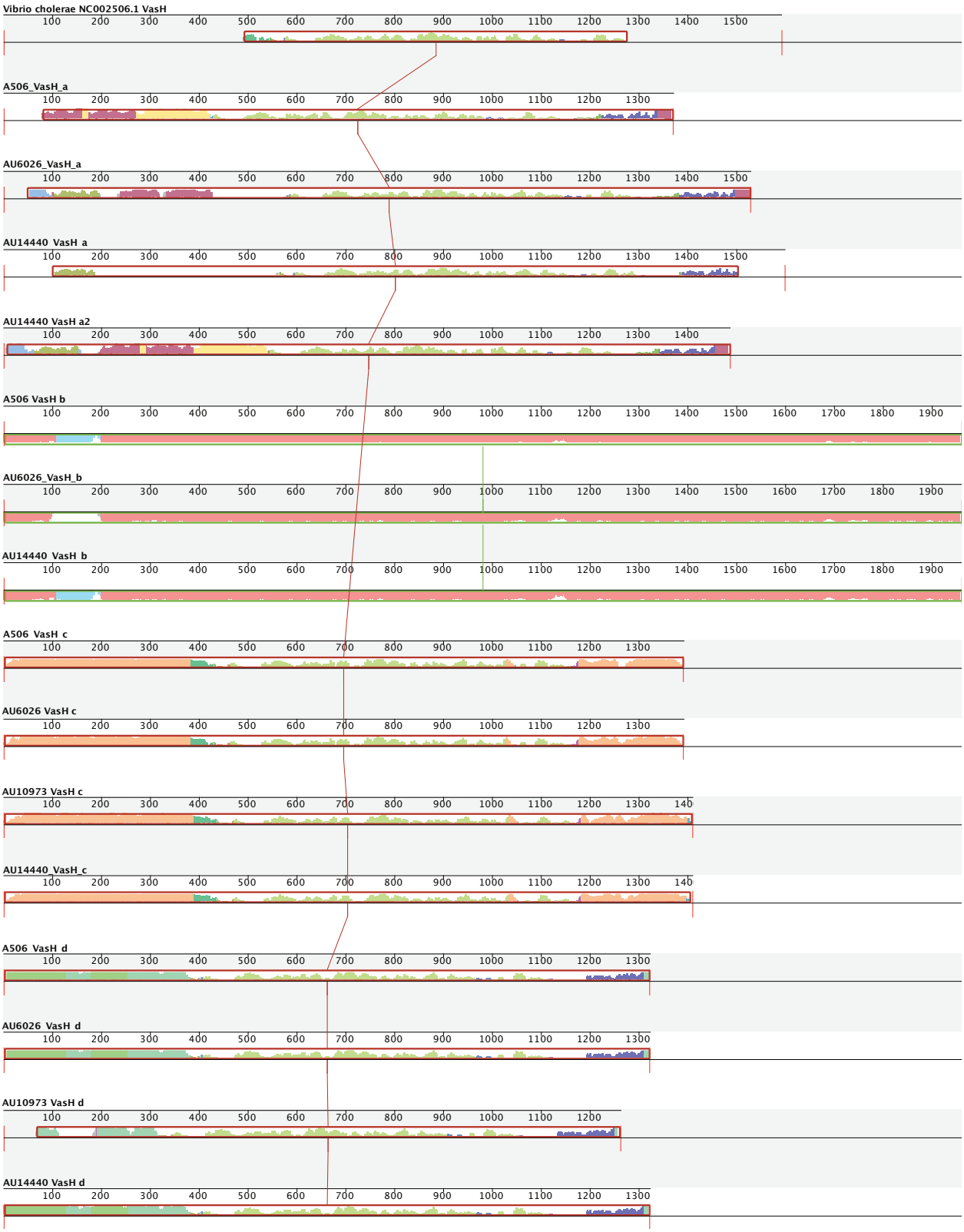
Appendix Figure IV.6: Full annotation of genes involved in secretion systems in subclade III strains. Abbreviations on left correspond to those found in Figure IV.7 of paper. Full annotated gene names are in the right column, as provided by the RASTannotation pipeline [148].

Appendix Figure IV.7: Alignment of FimA homologues.



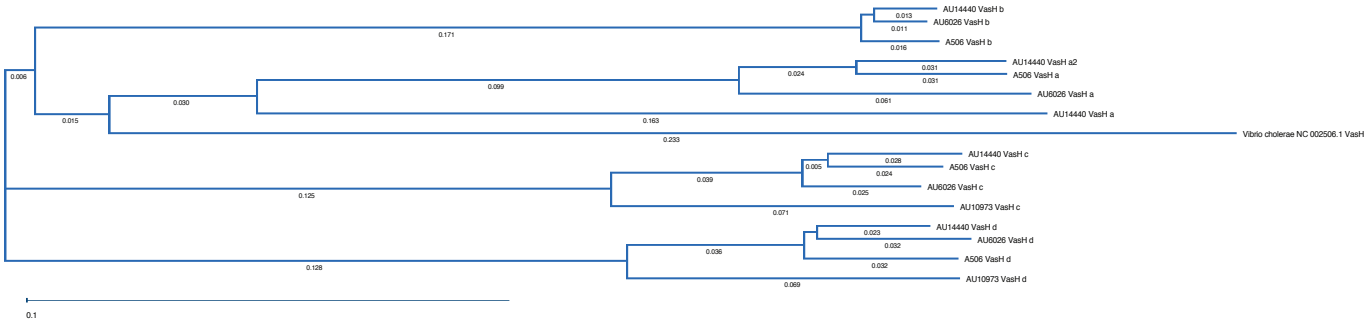
Appendix Figure IV.7: Alignment of FimA homologues.
Alignment of FimA homologues from select environmental and clinical subclade III strains. FimA homologues defined in RAST [149]. Nucleotide sequence of FimA homologues aligned with the MAFFT algorithm [194, 193] within DNAstar's MegAlign Pro software. A506 is a representative environmental strain; AU6026 and AU10973 are representative clinical strains.

Appendix Figure IV.8: Alignment of VasH homologues.



Appendix Figure IV.8: Alignment of VasH homologues.
Alignment of VasH homologues from select environmental and clinical subclade III strains. VasH homologues defined in RAST [149]. Nucleotide acid sequence of VasH homologues aligned in Mauve [198]. Colored regions refer to shared open-reading frames. A506 is a representative environmental strain; AU6026, AU10973, AU14440 are representative clinical strain.

Appendix Figure IV.9: Phylogenetic tree of VasH homologues.



Appendix Figure IV.9: Phylogenetic tree of VasH homologues.
Phylogenetic tree inferred from the nucleotide acid sequences of VasH homologues from representative environmental and clinical subclade III strains. Mauve algorithm used for alignment [198]. A506 is a representative clinical strain; AU6026, AU10973, AU14440 are representative clinical strains.