DISCOVERY OF A PRODIGIOSIN PRODUCING *VIBRIO*

BY
STEPHANIE MORGAN

A Thesis

Submitted to the Division of Natural Sciences

New College of Florida

In partial fulfillment of the requirements for the degree

Bachelor of Arts

Under the sponsorship of Dr. Katherine Walstrom

Sarasota, Florida

May, 2016

Acknowledgements

There are several people who helped me throughout this project that I'd like to express my gratitude to. First and foremost, to my committee members; Dr. Walstrom, Dr. Scudder and Dr. Ryba, who all personally helped me with this project. I'd also like to thank Dr. Steven Shipman for sponsoring the initial ISP and Tutorial where I was able to develop the project. A special thanks to Jon Reed for being incredibly helpful and allowing me to use the Mass Spectrometers at the Roskamp Institute. Finally, I'd like to thank Dr. Eric Warrick for his endless patience and for fostering my excitement in the project from the very beginning.

I'd also like to express my appreciation to the people who have kept me sane throughout this process. To my Dad for always supporting me and for never questioning my goals in life. To my wonderful partner Justin Munoz for reminding me to relax and pretending my jokes are funny. To Dr. Amy Clore for her kindness and advice throughout my time at New College. And to Michael Davis for keeping my head above the water at times when I felt I was drowning in thesis responsibilities.

Table of Contents

List of Figures

List of Tables

DISCOVERY OF A PRODIGIOSIN PRODUCING *VIBRIO*

Stephanie Morgan

New College of Florida, 2016

ABSTRACT

A soil sample taken from Leffis Key in Bradenton led to the discovery of a bacterial isolate capable of producing bright pink-pigmented colonies in culture. Identification of the isolate was done using a polyphasic approach that included biochemical characterization and multi-locus sequence analysis with genes found in the isolate genome. The isolate was identified as a member of the *Vibrio* genus within the Gazogenes clade.

Some of the pigment molecules responsible for the vibrant pink color observed in culture were purified via a series of steps including flash chromatography. Mass spectrometry analysis of the chemical extract identified the pigment molecules as related members of the prodiginine family. Additionally, a prodiginine biosynthesis gene cluster consisting of 13 genes with sequence homology to prosigiosin biosynthesis gene clusters in *Serratia* species was found in the genome of the isolate. Prodigiosins are biomolecules produced by a variety of bacterial organism as secondary metabolites. They are often brightly colored pigment molecules with unique chemical properties. These compounds have been

studied for their antibacterial, anti-malarial and anti-cancer properties. Prodigiosins also have potential applications in industry, as their brightly colored appearance makes them ideal candidates for natural dyes.

_____

Katherine Walstrom
Natural Sciences

Introduction

Bacterial organisms are often overlooked by the general public as only contributors of disease; however, prokaryotes are an incredibly diverse group of organisms serving many important functions necessary to sustain life. More practically, prokaryotes have long been investigated as the source of many natural products useful in medicine and industry. Recently, focus has been put on the discovery of diverse microbes in marine environments (Xiong et al., 2013). It was under this pretext that a bacterial organism appearing as vibrant pink colonies was isolated and was nick-named "PINK" (Figure 1).

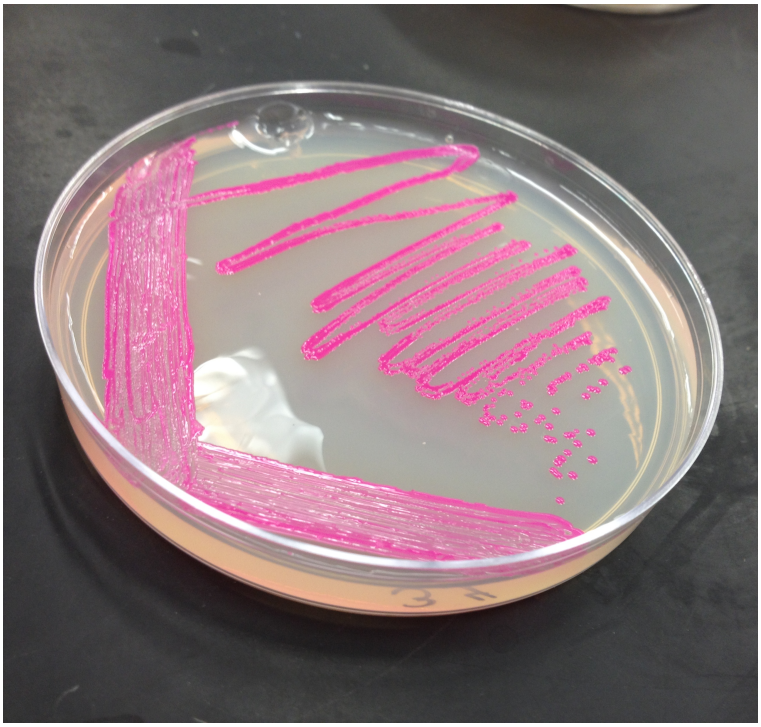

Figure 1: PINK in Culture
PINK streaked on modified PDA media shows a colony morphology consisting of small circular bright pink colonies. The bright pink color is thought to be a result of the production of a prodiginine pigment compound.

Prokaryotic Species Classification

Traditionally bacteria have been identified using phenotypic characteristics including pure culture bacterial morphology and biochemical tests. In 1923 the first edition of "Bergy's Manual of Determinative Bacteriology" was published and included a series of phenotypic characteristics used to identify and classify bacteria (as reviewed by Roselló-Mora and Amann, 2001). The rational for this type of classification system is that phenotypic observations are directly related to the genotype; however, the problem with this approach is that it is difficult to determine whether phenotypic differences or similarities are directly related to the same genotypes. Regardless, phenotypic classification is still widely used in clinical settings where commercial products have been manufactured to test a wide array of biochemical characteristics in a short amount of time. An example of this is the analytical profile index (API®) strip 20 E, which uses a set of biochemical test capsules in a small portable strip. 24 hours after inoculation of a pure culture, the strip can be read as a set of positive and negative results and entered into a database for recognizing *Enterobacteriaceae* (API 20 E user manual).

In the 1960's, analysis of DNA first became popular for classifying bacteria but was limited to relative values such as base composition and DNA-DNA hybridization (DDH) studies. In DDH studies the genomes of two organisms are allowed to denature to single strands and then re-hybridize to double-stranded DNA molecules. The degree to which hybridization occurs is then measured using relative binding ratios or melting temperature (T_m) to determine the relative similarity between the two organisms. A 70% or higher DDH value is necessary for species identification. This technique was later expanded on in DNA-rRNA hybridization experiments as the conservation of ribosomal

RNAs was shown to be a good marker for identifying related organisms (as reviewed by Roselló-Mora and Amann, 2001).

As genetic sequence information became available, the use of more direct sequence comparison has been used to show relatedness between many organisms (Chun and Rainey, 2014). The 16S rRNA gene was the first gene widely used to determine bacterial relatedness and is generally sequenced as a preliminary identifier of an isolated organism (as reviewed by Roselló-Mora and Amann, 2001). While the 16S rRNA gene is sufficient for identifying an organism at higher taxa, it has poor resolution at the species level because the sequence similarity is too high (Pascual et al., 2010; Zeigler, 2003).

The introduction of second generation sequencing techniques greatly advanced bacterial taxonomy by significantly reducing the labor and cost required for whole genome sequencing (Chun and Rainey, 2014). Rather than experimentally determining genome relatedness using DDH techniques, a direct comparison of genomes could be done by determining the average nucleotide identity (ANI). This value is the average similarity between various sets of aligned regions in a pair of genomes, and an ANI value of 96% or higher correlates to a 70% DDH value and thus defines a species (Konstantinidis and Tiedje, 2005). This method is limited in that there are still many un-sequenced bacterial genomes, and it is not always practical or cost effective to have an entire bacterial genome sequenced.

In overall genome relatedness studies, only the conserved regions of genomes contain phylogenetic information and therefore can be reduced to comparing sets of genes (Chun and Rainey, 2014). Zeigler suggested a set of rules for determining appropriate genes to use to determine phylogenies that are as follows:

"First, the genes must be widely distributed among genomes, with orthologous sequences appearing in most, if not all, free-living bacteria. Secondly, because the occurrence of gene families could make sequencing and alignment technically difficult, each of the 'prediction set' genes must be unique within a given genome, without close paralogues that could confuse analysis. Thirdly, individual gene sequences must be long enough to contain phylogenetically useful information but short enough to be sequenced economically with a small set of primers. Finally, the sequences must predict whole-genome relationships with acceptable precision and accuracy."

(Zeigler 2003).

Multilocus sequence analysis (MLSA) is used to determine relatedness between organisms based on the sequence similarity of several conserved genes and is a more practical and reliable alternative to DDH studies (Sawabe et al., 2013). This is done by concatenating a set of several conserved gene sequences for each species in question and performing a multi-sequence alignment. MLSA has the benefit of contributing phylogenetic information from multiple different genes while remaining less resource intensive than whole genome sequencing. The use of several conserved genes reduces inconsistencies caused by horizontal gene transfer and increases resolution when analyzing a large group of bacteria because different genes within the set will have different resolving power for different groups of bacteria within the whole (Thompson et al, 2005; Pascual et al., 2010).

The Vibrionaceae family is a diverse group of bacteria including *Aliivibrio*, *Echinimonas*, *Enterovibrio*, *Grimontia*, *Photobacterium*, *Salinivibrio* and *Vibrio* genera (Sawabe et al., 2013). Vibrionaceae are Gram-negative, facultatively anaerobic and generally shaped like curved rods. Modern characterization and phylogenetic analysis of Vibrionceae family is done using MLSA (Thompson et al., 2005; Sawabe et al., 2013; Gabriel et al., 2014; Pascual et al., 2010). An 8-gene MLSA on Vibrionaceae further characterized the family into 22 distinct clades (Sawabe et al., 2013). Genes used in one

or more of these studies are: *recA*, *rpoA*, *pyrH*, *gapA*, *gyrB*, *topA*, *mreB*, *ftsZ*, *rctB*, *toxR* and 16S rRNA. However, the phylogenetic study of *Vibrios* utilizing an 8-gene scheme showed that addition of the 16S rRNA gene did not have a significant impact on the overall phylogeny (Sawabe et al., 2013).

The current consensus on bacterial taxonomy is that a polyphasic approach be employed when identifying a new species, this includes both sequence analysis and traditional biochemical and phenotypic characterization (Van damme et al., 1996; Roselló-Mora and Amann, 2001). It should also be noted that biochemical and phenotypic characterization is still widely used to roughly identify previously known species (Noguerola and Blanch, 2008).

Prodiginine Compounds in Bacterial Species

Prodiginine compounds are often vibrant pigment compounds produced as secondary metabolites by a variety of different bacteria. Of the prodiginine compounds, prodigiosin has been the most extensively studied since originally sparking the interest of scientists in 1902 as the bright red compound produced by *Serratia marcescens*; however, it was not until 1960 that the correct structure for prodigiosin was determined (as reviewed in Gerber, 1975). Since then many other prodiginine pigments have been discovered as secondary metabolites in a variety of different bacterial organisms, including species of the genera *Actinomadura*, *Hahella*, *Psuedomonas*, *Streptomyces*, *Vibrio* and others (as reviewed by Darshan and Manonmani, 2015). Prodigiosin is described as a secondary metabolite with no direct benefit to cellular growth. Pure culture mutant non-pigmented strains of *S. marcescens* show no significant difference in viability from pigmented strains (Williams, 1973). However, in the environment prodigiosins may

serve a purpose in bacterial defense, as many have antibacterial properties (as reviewed in Fürstner, 2003). Interestingly, some have also been shown to have other activities such as anti-malarial and even anti-cancer properties (Singh and Shekhawat, 2012; Williamson et al., 2007).

Structure and Chemical properties of Prodiginine Compounds

Prodiginine compounds consist of a pyrolylpyrromethene core containing 3 linearly arranged pyrrole rings (A, B and C) with a methoxy group on pyrrole B and a variety of substituent groups on the pyrrole rings (Figure 2) (as reviewed in Gerber, 1975). The structure for the original prodigiosin consists of a prodiginine core with a methoxy group on pyrrole B and methyl and pentyl alkyl groups on ring C. Similar compounds produced in some species as minor byproducts vary only in the number of carbons in the alkyl substituent groups on pyrrole C. Alternatively, undecylprodigiosin produced as the major product in *Streptomyces coelicolor* contains the same methoxy tripyrrole core with only one 11 carbon akyl chain substituent group on pyrrole C (Cerdeño et al., 2001).

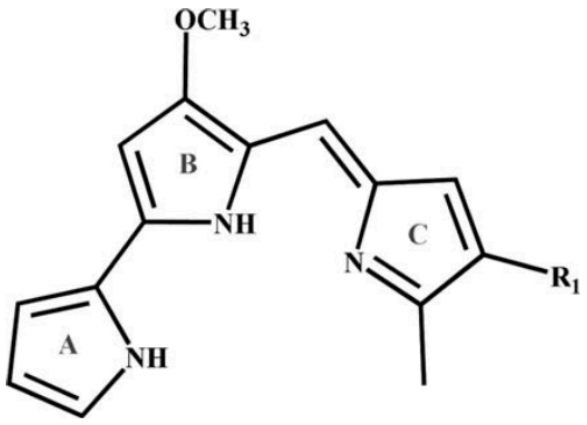

Figure 2: Prodiginine Core Structure
Core structure of prodiginine pigments showing linear tripyrrole arrangement with a methoxy substituent group and alkyl groups on pyrrole C similar to prodigiosin, where R1 is n-pentyl (Alihosseini et al., 2009).

Other variations occur as substituent groups on pyrrole rings A and B. Norprodigiosin is produced as a minor byproduct in *Serratia marcescens* and contains a hydroxyl group in place of the methoxy group typical for prodigiosin (Hearn et al., 1964). The structure of a prodiginine compound produced by *Pseudoalteromonas rubra* shows a p-hydroxybenzene substiuent group on pyrrole A (Fehér et al., 2008). Finally several cyclic prodiginine derivatives are also abundantly produced among different species, such as cycloprodigiosin, nonylprodigiosin, streptorubin B and many others (as reviewed in Fürstner, 2003).

Interestingly, prodigiosin can also act as a pH indicator as a change in color occurs at a pKa of approximately 8.25. This results in UV-Vis spectra showing two peaks, one at 537 nm in acidic conditions and another at 470 nm in basic conditions (Hearn et al., 1968). The color change is due to two conformational isomers that have an equilibrium depending on pH. The protonated form of prodigiosin is more prominently in the α-conformation, due to hydrogen bond stabilization between the protonated nitrogen and the oxygen of the methoxy group. The unprotonated form prefers the β-conformation in which the C pyrrole ring is rotated 180 degrees and no hydrogen bond stabilization occurs (Figure 3) (Rizzo et al., 1998).

As mentioned earlier, prodiginine pigments have been studied for their antibiotic activity. It is thought that this ability is due to the nuclease activity of the prodiginine core when forming a complex between the nitrogens of the prodiginine core and a copper metal cation. While it is not completely clear how this mechanism works, in the absence of copper nuclease activity was not detected (as reviewed by Fürstner, 2003).

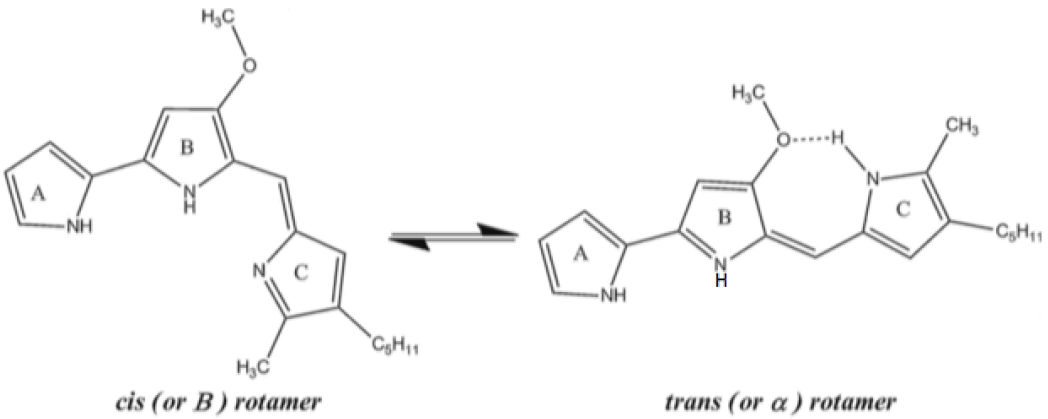

Figure 3: Cis and Trans Prodigiosin Rotamers
The structure of prodigiosin as cis and trans rotamers. The protonated trans rotamer is more prominent at acidic pH and appears red or pink, while the cis rotamer is more prominent in basic conditions and appears yellow (Darshan and Manonmani , 2015).

Prodigiosin Biosynthesis

Although prodiginine pigments have been found in many different organisms, prodigiosin biosynthesis has only been studied in *Serratia marcescens* strain ATCC 274, *Serratia* spp 39006, *Hahella chejuensis* strain KCTC 2396 and some *Streptomyces* strains (Figure 4). In *S. marcescens* strain ATCC 274 (*Sma* 274), the prodigiosin biosynthesis (Pig) cluster consists of 14 protein coding genes, abbreviated as *pigA* through *pigN*, that are transcribed as a single 14-gene polycistronic mRNA. The Pig cluster in *Serratia* strain ATCC 39006 includes an additional gene, *pigO*, although it has no putative function in prodigiosin biosynthesis and knockout experiments yielded no difference in the pigments produced (Harris et al., 2004). The *H. chejuensis* KCTC 2396 genome contains a set of 14 prodiginine biosynthesis genes, *hapA* through N, similar in the overall gene

layout as well as sequence homology that is within a 26 gene cluster. Of the additional genes in the Hap cluster, two serve regulatory functions and the function of the other 10 are unknown (Kim et al., 2006). In contrast, the Red cluster is responsible for prodiginine biosynthesis in *Streptomyces coelicolor* and contains 23 genes (Cerdeño et al., 2001). The structure of this cluster is significantly different than that of Pig or Hap clusters and is transcribed into 4 separate mRNA strands. Despite this, there are 12 genes between the *Streptomyces* strain and *Serratia* strain prodiginine biosynthesis clusters that are homologous, suggesting similar biosynthesis pathways (Harris et al., 2004). A similarly structured cluster in other *Streptomyces* species, the Mar cluster, is responsible for the production of cyclic prodiginines called marineosins (Salem et al., 2014).

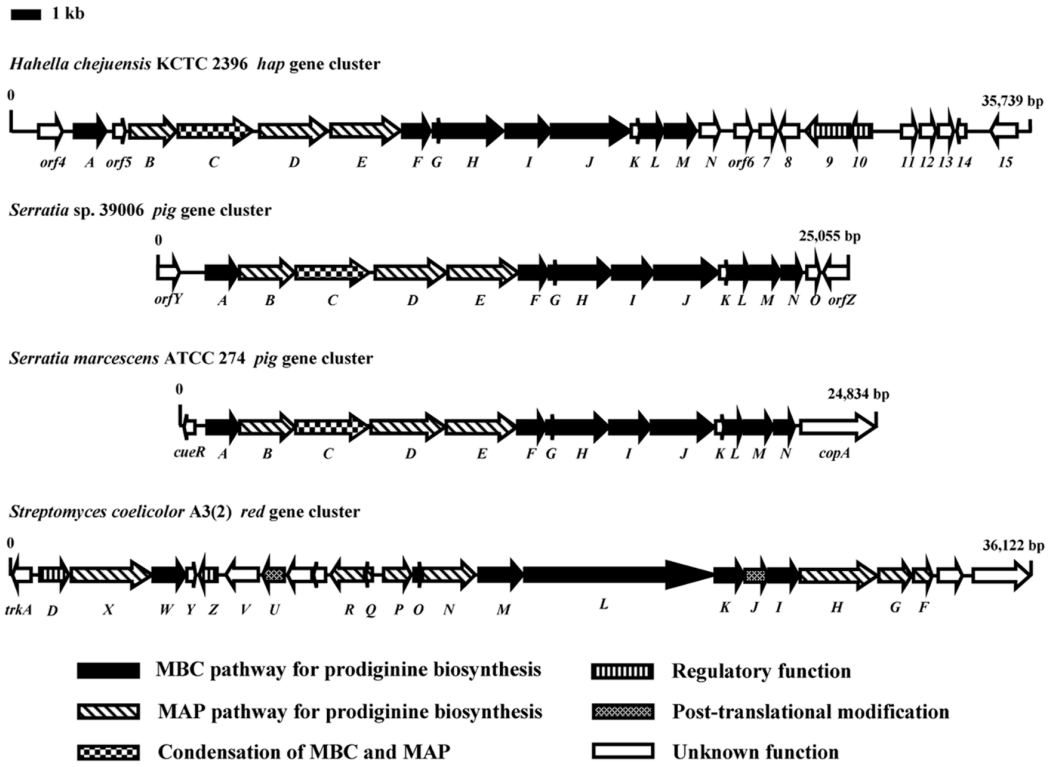

Figure 4: Prodiginine Biosynthesis Clusters in different organisms
The structure of prodiginine biosynthesis gene clusters is very similar between *Serratia* species and *Hahella chejuensis*, which both contain 14 homologous genes with similar putative function. In contrast the gene cluster in *Streptomyces coelicolor* is different in both overall structure and homology between genes (Kim et al., 2006).

In early experiments to elucidate the prodigiosin biosynthesis pathway, UV-induced mutant strains of *S. marcescens* were used in a set of experiments where syntrophic interactions between specific mutants with the ability to produce one precursor but not the other were observed. The results revealed that the Pig pathway is bifurcated with the production of two main precursors: 2-methyl-3-N-amyl-pyrrole (MAP) and 4-methoxy-2,2-bipyrrole-5-carboxyaldehyde (MBC). These are combined enzymatically in a condensation reaction in the final step of prodigiosin production (Morrison, 1966).

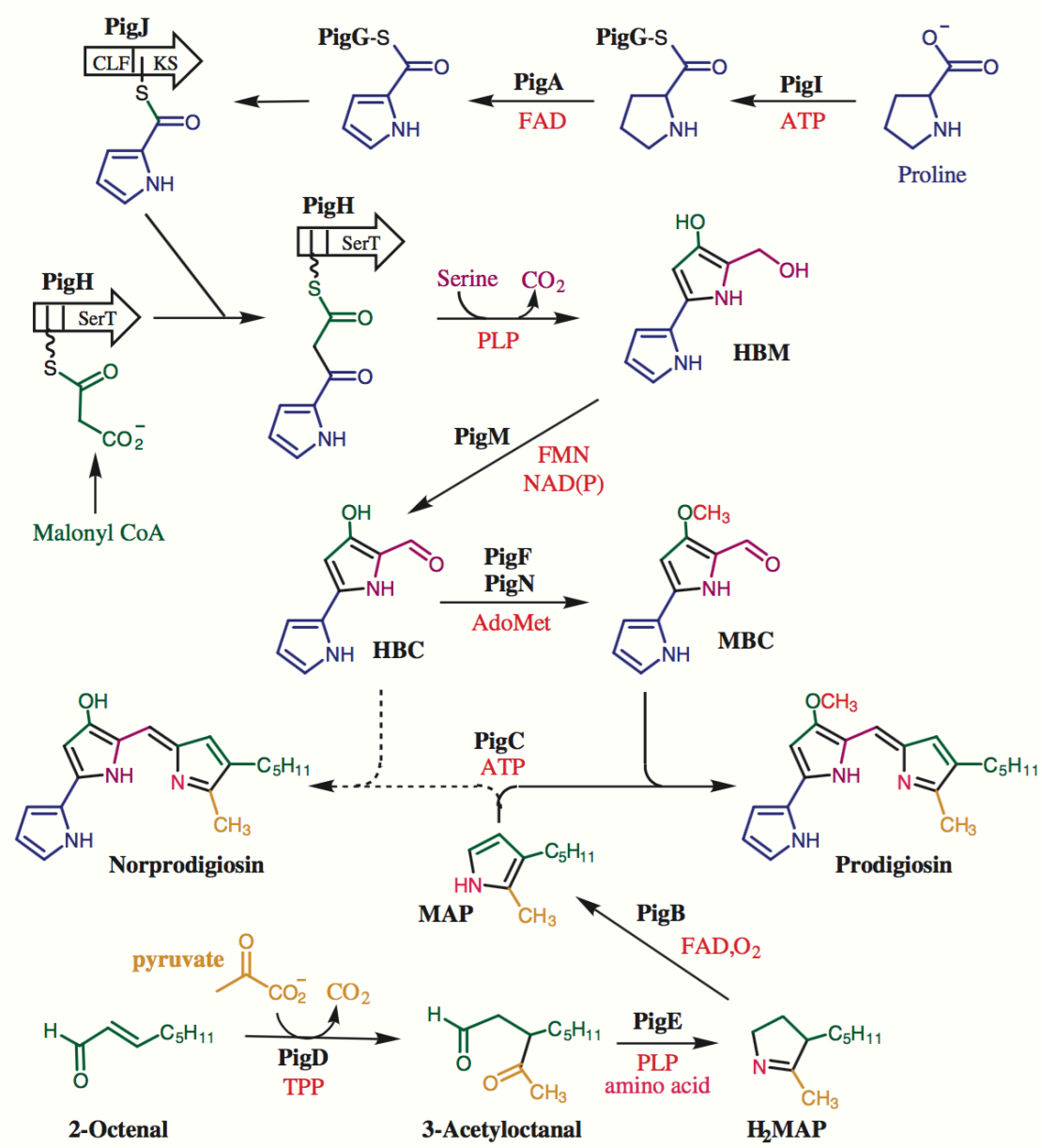

Figure 5: Proposed Prodigiosin Biosynthesis Pathway in *Serratia marcescens*
Diagram showing the bifurcated prodigiosin biosynthesis pathway showing all intermediate compounds and the association with each enzyme in the pathway. The use of atoms from precursor molecules, proline, malonyl CoA, serine, 2-octenal and pyruvate, are shown via color-coding throughout the pathway (Williamson et al., 2005).

In a set of similar syntrophy experiments, selective Pig mutants were created in *Serratia* strain 39006 using in-frame insertions and deletions for each enzyme within the Pig cluster. When used in combination with high-resolution mass spectrometry, this allowed for the correlation of each enzyme to the intermediate

compounds within the molecular pathway (Williamson et al., 2005). The entire prodigiosin biosynthesis pathway was elucidated in this way and is shown in Figure 5. Enzymes PigA, PigF, PigG, PigH, PigI, PigJ, PigM and PigN were implicated in biosynthesis of the precursor molecule MBC, while enzymes PigB, PigD and PigE were implicated in biosynthesis of the precursor MAP. Additionally, PigC was recognized as the enzyme responsible for condensation of MAP and MBC (Williamson et al., 2005).

Discovery of PINK

PINK was originally discovered during a Small World Initiative Microbiology lab course taken as an independent study credit at State College of Florida, Manatee-Sarasota. The Small World Initiative is a program started in 2012 by Yale University that outsources the screening of bacteria from the environment for antibacterial production by employing the program at different schools both nationally and internationally (http://www.smallworldinitiative.org/).

The environmental sample containing PINK was taken from a salt flat at Leffis Key in Bradenton, Florida. The initial interest in PINK surrounded the antibacterial properties PINK exhibited. Inhibition assays using a crude chemical extract from cultured PINK cells showed growth inhibition activity against Gram-positive organisms *Staphylococcus epidermidis* and *Bacillus subtilis,* and Gram-negative organisms *Enterobacter aerogenes*, *Citrobacter freundii* and *Chromohalobacter salexigen* (Figure 6). This suggested that PINK was capable of producing a broad-spectrum antibiotic or multiple antibiotics.

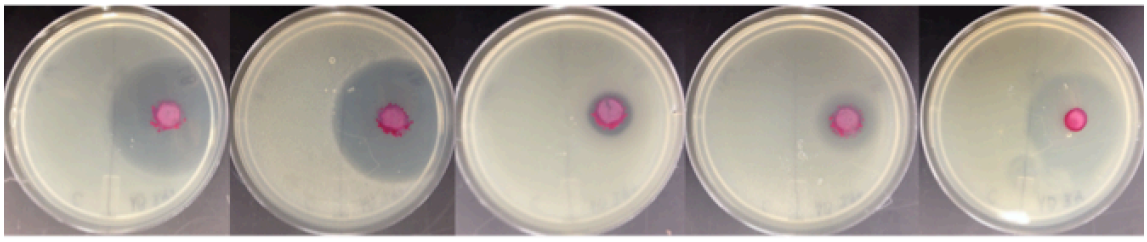

Figure 6: Antibacterial activity of crude extract from PINK
Growth inhibition assays against *S. epidermidis, B. subtilis, E. aerogenes, C. freundii* and *C. salexigens* on LB and 0.5 M NaCl LB (*C. salexigens*) agar plates. Zones of inhibition (clear areas) of varying thickness are visible around bright pink spots on the right side of the plate. Zones are more prominent against Gram positive organisms. No zone of inhibition was seen on the left side of the plate containing an ethyl acetate spotted control.

One aim of this study is to confirm the identity of PINK as a member of the *Vibrio* genus using a polyphasic approach that includes both sequence analysis and biochemical characterization. Another goal of this study is to isolate and identify the pigment produced by PINK that is suspected to be a prodiginine compound. In addition to this, a prodiginine biosynthesis gene cluster found within the PINK genome was also analyzed for homology with other known prodigiosin biosynthesis clusters.

Materials and Methods

Isolation and Phenotypic Characterization

PINK was initially isolated from an environmental sample taken from a salt flat located on Leffis Key in Bradenton, FL. The sample was serially diluted in 0.5 M NaCl to 10^{-5} grams soil/mL and grown on 4% NaCl modified potato dextrose agar (PDA) medium containing tryptone, MOPS salts and D-glucose in addition to the standard PDA recipe (see Appendix). To control for fungal growth from environmental samples 100 μg/L cyclohexamide was added on the initial isolation plates. After initial isolation of the PINK colony, the organism was maintained on 4% NaCl modified PDA agar and 4% NaCl LB agar (see Appendix A) and kept at 30°C. It should be noted that as a safety precaution the PINK isolate was kept on solid agar media until whole colony PCR of the16S rRNA gene, followed by sequencing and BLAST analysis could roughly identify the organism as a non-pathogenic organism.

In determining suitable halophillic conditions for PINK, the organism was streaked on a series of modified PDA and LB agar plates with different NaCl concentrations at 0%, 2%, 4%, 6%, 8%, 10% and 12%. Plates were incubated at 30°C and allowed to grow for 24 to 48 hours. Biochemical testing for PINK was done by inoculating Biomerieux analytical profile index (API®) strips (see Appendix for manual) with freshly streaked colonies suspended in 0.5 M NaCl sterile saline. The strips were incubated at 30°C for approximately 24 hours before interpretation using the kit guidelines.

Genomic DNA from PINK was extracted at Florida Southern University, and the procedure was acquired through communications with Dr. Eric Warrick (see Appendix). Sequencing was conducted using HiSeq 2500 technology at Purdue University. The genomic sequence from PINK was analyzed using the sequence viewer and annotator tool, Artemis (Rutherford et al., 2000). The program has a feature that enables it to mark all putative coding DNA sequences (CDS). The CDS region includes all nucleotide bases that are translated into protein as indicated by start and stop codons. This includes the intact protein sequence, as prokaryotic genes lack the introns found in eukaryotic organisms. The CDS detected were translated into the amino acid sequences by Artemis using a standard codon table. Each CDS was identified manually using the National Center for Biological Information (NCBI) protein basic local alignment tool (Gish and States, 1993).

Phylogenetic Analysis by MLSA

MLSA is a less intensive alternative to overall genome studies in order to identify organisms using a set of genes that are aligned and then concatenated. The 5-gene and 8-gene MLSA schemes used in this study were based on previous analyses done by Gabriel et al., 2014 and Sawabe et al., 2013 respectively. All nucleotide sequences for *recA*, *rpoA*, *gapA*, *gyrB* and *ftsZ* genes from 62 different organisms in the *Vibrio genus* were obtained from the GenBank database and compared to sequences acquired from the PINK genome (see Appendix C for accession numbers) <http://www.ncbi.nlm.nih.gov/genbank/>. *Vibrio* organisms lacking a full set of the 5 MLSA genes were excluded. Additionally, organisms with gene sequences significantly smaller (50 bp or more difference) than all other aligned sequences for one or more genes

were also excluded from this study. Sequences were aligned separately using ClustalW (Larkin et al., 2007; Goujon et al., 2010). Aligned regions of the genes were used while the extensions of individual genes beyond the aligned regions were removed. Aligned sequences from the listed genes were then concatenated to create a 5-gene chain. Similarly, a separate 8-gene chain of concatenated sequences including *recA*, *rpoA*, *gapA*, *gyrB*, *ftsZ*, *pyrH*, *topA*, and *mreB* genes was constructed for PINK and each *Vibrio* organisms within the Gazogenes clade defined by Sawabe et al., 2013. This clade includes *V. aerogenes*, *V. rhizosphaerae*, *V. gazogenes* and *V. ruber*.

Phylogentic analysis was done for each gene set using MEGA7 (Kumar et al., 2015). Maximum likelihood trees with a phylogenetic test of 100 bootstrap replications were created using parameters suggested in Gabriel et al., 2014. Nucleotide percent identity values were acquired for both the *Vibrio* genus 5-gene sequence alignments and the Gazogenes clade 8-gene sequence alignments using a percent identity matrix in Clustal Omega (Sievers et al., 2011; Goujon et al., 2010).

Analysis of Prodigiosin Gene Cluster

Prodigiosin cluster genes were identified using NCBI BLASTP (Gish and States, 1993). Comparison of prodigiosin biosynthesis genes from PINK and other known prodiginine compound producing organisms *Serratia marcescens* ATCC 274, *Serratia* spp. 39006, *Hahella chejunsis* KCTC 2396, and *Streptomyces coelicolor* A3(2) was done by aligning amino acid sequences using Clustal Omega for each gene within the cluster (Sievers et al., 2011). Gene sequences were acquired using the NCBI's GenBank database, and a list of accession numbers used can be found in the appendix.

Crude Extract Growth Inhibition Assays

PINK was heavily streaked on 4% NaCl modified PDA and left to grow over 24 to 48 hours at 30°C to create a bacterial lawn. The agar was then cut into pieces and placed in a container, and 20 mL of ethyl acetate was added. The sample was then placed at -80°C for 30 minutes before being placed on an orbital shaker at room temperature for 12 hours. The organic layer was then removed, placed in a small conical vial and air evaporation of ethyl acetate took place in a hood over 12 hours. Extracts were re-suspended in 1mL of ethyl acetate. A total of 80 µL of extract was spotted onto a single region of an LB agar plate and allowed to dry. Ethyl acetate was used as a control on another region of the plate. The organism being tested against was inoculated in top agar and poured over the plate evenly. Organisms tested against were Gram-positive organisms, *Staphylococcus epidermidis* and *Bacillus subtilis* and Gram-negative organisms *Chromohalobacter salexigenes, Enterobacter aerogenes, and Citrobacter freundii.* Plates were left to incubate for 24 hours and then analyzed for clear zones of inhibition to indicate antibacterial activity.

Purification of Prodigiosin

Overnight cultures of PINK in 4% NaCl LB were centrifuged at 8000 rpm for 10 minutes. The supernatant was discarded, and approximately 2 ml of a methanol and 2 N HCl mixture (24:1) was used to resuspend the pellets. The addition of acid was necessary to break down a suspected prodigiosin associated protein that may sequester the pigment molecule (Gerber, 1975; Kobayashi and Ichikawa, 1989). All portions were combined in 35 ml scintillation vials wrapped in tin foil and left for 12 to 18 hours on an orbital

rotator. Extracts were then centrifuged at 8000 rpm for 10 minutes to remove any additional cell mass. The supernatant was collected, and the solvent was evaporated using a rotary evaporator. Hydrophilic impurities were removed by a series of chloroform-water liquid-liquid extractions that were repeated on the organic layer until the water layer no longer appeared cloudy (Alihosseini et al., 2009). The chloroform was evaporated via rotary evaporation, and pigment was redissolved in acetonitrile. Flash chromatography was performed using silica gel as the stationary phase and acetonitrile as the mobile phase. Fractions from flash TLC correlating to a pigmented smear at 0.82 to 0.63 Rf regions on silica TLC plates with acetonitrile as the mobile phase were combined and concentrated by rotary evaporation. This particular smear was chosen because it appeared as bright pink on the TLC and was able to be separated from other impurities present in the solvent front.

UV-Vis pH Assay

The pH of methanol solvents was adjusted to 1.8, 5.0, 7.0, 8.0, 11.1 and 12.0 using solutions of 1 M HCl and 1 M NaOH. The pH of each solvent was confirmed using a pH probe and pH paper. 100 μL of purified prodigiosin pigment extract was suspended in solvent, and absorbance spectra were measured from 350 nm to 700 nm at 0.25 nm intervals. A baseline correction was performed for all samples presented.

Mass Spectrometry

Electrospray Ionization (ESI) tandem mass spectrometry (MS/MS or MS^2) was performed using Fourier transform mass spectrometry (FTMS) mode on an LTQ-Orbitrap with helium used as the collision gas. Relative collision energy (rCE) ranged from 30-35

in the ion trap component of the instrument. High-resolution mass spectra were obtained with full width half maximum resolving power of 100,000 at 400 m/z in profile mode. Full width at half maximum is the width of the curve measured between the two points at which the y-axis is at half the maximum value.

Results

Identification of PINK

As a preliminary experiment to determine the identity of the PINK isolate, PCR was performed using primers specific for the 16S rRNA gene. Entering the 16S rDNA sequence into the National Center for Biotechnology Information (NCBI) basic local alignment search tool (BLAST) found that the sequence most closely aligned with sequences within the genus *Vibrio.* Additionally results from BLAST indicated that PINK is most closely related to *Vibrio ruber, Vibrio gazogenes* and *Vibrio rhizosphaerae*, all members of the Gazogenes clade within the Vibrionaceae family (Sawabe et al., 2013). All three are obligate halophiles isolated from marine environments, have a single polar flagellum and appear as bright red-pigmented circular colonies in culture (Harwood, 1978; Shieh et al., 2003; Ramesh kumar and Nair, 2007).

In this study, sequences located in the genome of the PINK isolate were included in a phylogenetic analysis of the *Vibrio* genus using a 5-gene (*recA*, *rpoA*, *gapA*, *gyrB* and *ftsZ*) MLSA scheme as suggested in a 2014 paper by Gabriel et al. Further analysis was performed using sequences for each species within the Gazogenes clade in an 8-gene MLSA scheme with the addition of *mreB*, *topA* and *pyrH* as suggested in by Sawabe et al 2013. Biochemical characterization of PINK was also done using a bioMeurieux API® strip and compared to the known characteristics of *Vibrio* in the Gazogenes clade (Noguerola and Blanch, 2008).

Phenotypic Characterization of Pink

PINK is a halophillic bacterium with the ability to grow on media with 2% to 12% NaCl concentration but will not grow on media containing no NaCl (data not shown). Further growth media experiments revealed that PINK grew on media supplemented with Na_2SO_4 but not KCl, indicating that only sodium ions are necessary for the growth of PINK (data not shown). Interestingly, the color of PINK colonies in culture varies depending on type of media and the NaCl concentration (Figure 7). Colonies on modified PDA media generally appear as a light pink, while colonies on LB appear as a red. Additionally, higher concentration of NaCl generally correlates with less vibrant color.

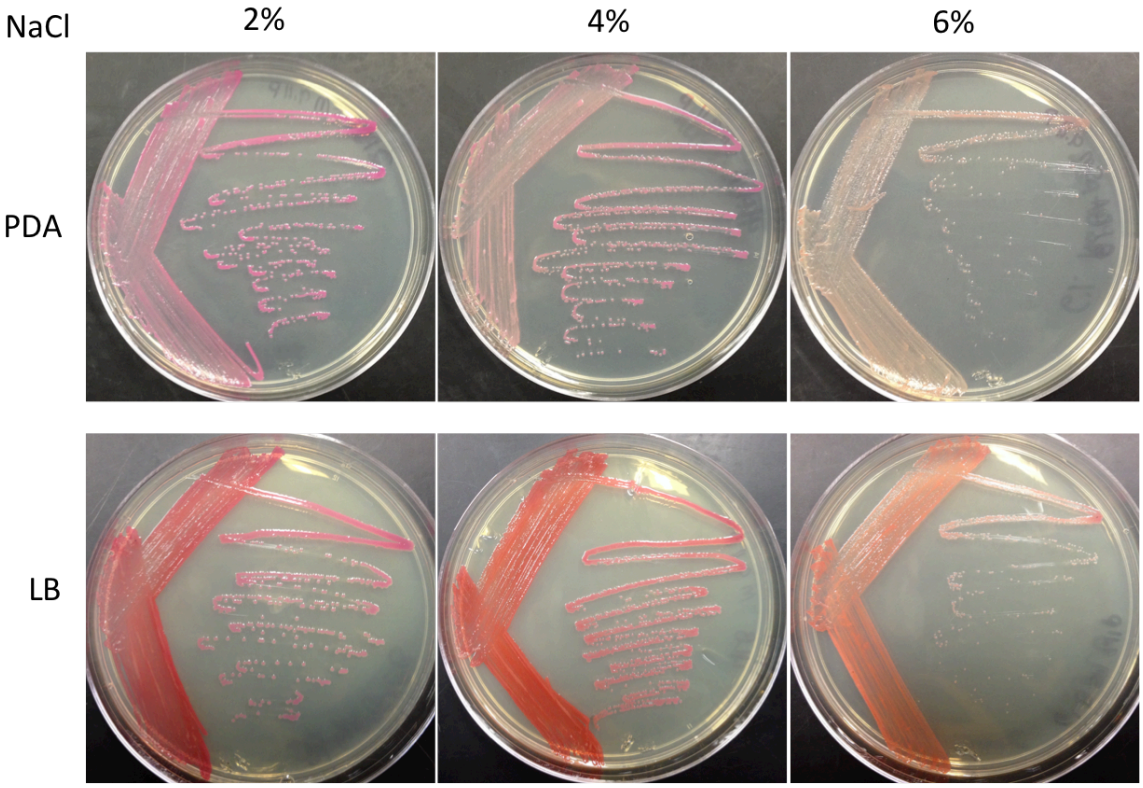

Figure 7: Appearance of PINK on Different Media
PINK streaked on two different types of media, LB and modified PDA, with different concentrations of NaCl appear as different colors.

Biochemical Test		*V. aerogenes*	*V. ruber*	*V. gazogenes*	*V. rhizosphaerae*	PINK
Halophilic Growth Conditions	0% NaCl	-	-	-	-	-
	3% NaCl	+	+	+	+	+
	6% NaCl	+	+	+	+	+
	8% NaCl	+	+	+	+	+
	10% NaCl	+	+	+	+	+
Acid from Fermentation of:	Inositol	+	-	-	+	-
	Manitol	+	+	+	+	+
	Melibiose	-	+	v	+	-
	Rhamnose	ND	ND	-	ND	-
	Sorbitol	-	ND	v	+	-
	Sucrose	+	+	+	+	+
Amygdaline		ND	ND	v	ND	+
Urease		ND	ND	-	ND	-
Voges-Proskauer		-	-	-	+	-
NO_2		+	+	-	-	-
ONPG		+	ND	v	ND	+
Oxidase		-	-	-	-	-
Citrate utilization		+	+	+	+	-
Gelatinase		+	+	+	+	+
Indole Production		+	-	-	-	-
Arginine dihydrolase		+	-	-	-	-
Lysine decarboxylase		-	-	-	-	-
Ornithine decarboxylase		-	-	-	-	-

Table 1: Biochemcial Characterization of PINK Compared to the Gazogenes Clade
Biochemical characteristics for all organisms other than PINK were found in Noguerola and Blanch, 2008.

Biochemical characterization of PINK compared to other members of the Gazogenes clade is summarized in Table 1. Characterization of PINK most closely aligns

with that of *V. gazogenes,* which is similar to that of PINK in every test except the ability to utilize citrate. However, lack of data for some tests for *V. ruber*, *V. rhizosphaerae* and *V. aerogenes* significantly hinders this comparison. Despite the homology between members of the Gazogenes clade, there are many characteristics in which these organisms differ, especially between *V. aerogenes* and other members of the clade. This variation exists even between different strains of *V. gazogenes*. It can be argued that lack of testing among different strains of other represented species is cause for false identification via biochemical analysis. It is interesting to note that PINK was not able to utilize citrate while all other members of the Gazogenes clade were able to. This suggests that this characteristic has been lost in PINK.

Phylogenetic Analysis of PINK using MLSA

The relationship between several organisms can be described by the homology between conserved genetic sequences. Maximum likelihood methods create a topology for related organisms based on the statistical likelihood of each pairwise alignment for the organisms analyzed (Mount, 2008). This is based on sequence similarity as well as the likelihood that other genetic modifications, such as substitution, have occurred. Maximum likelihood methods applied to the conserved 5-gene concatenated sequences; *recA*, *rpoA*, *gapA*, *gyrB* and *ftsZ*, for the 62 *Vibrio* genera used in this study produce a phylogenetic tree topology show in Figure 7. The topology of the tree produced can be organized into previously defined clades. Comparing the placement of organisms in a given clade is necessary for this study in order to find the most closely related organisms to PINK (Sawabe et al., 2013; Gabriel et al., 2014). It also is used as a measure of the reliability of the methods and sequences used to generate the phylogenetic tree.

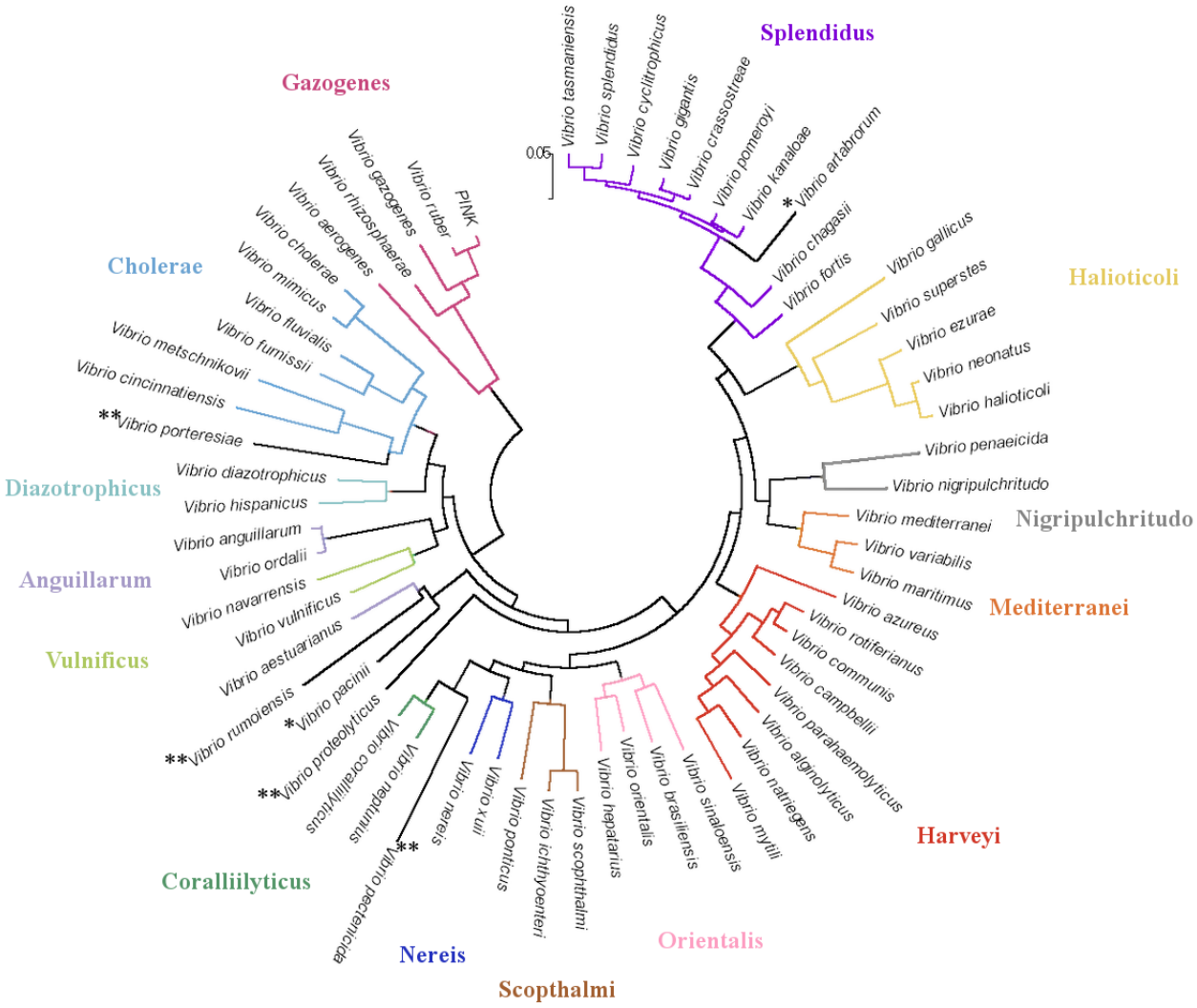

Figure 8: *Vibrio* genera Maximum Likelihood Phylogenetic Tree including PINK
Maximum likelihood tree from 5-gene MLSA of 62 *Vibrio* genera species generated using MEGA7. Clades as defined in Sawabe et al., 2013 are color-coded. PINK is represented as a member of the clade Gazogenes. *Species not previously placed in Clade via Sawabe et al., 2013. ** Orphan clade or only represented species of Clade in study, clades Porteresiae, Rumoiensis, Proteolyticus and Pectenicida.

Bootstrapping is a statistical test used to measure the likelihood that a result occurred by chance by randomly resampling and replacing data to see if the altered data yield similar results. As a test of phylogeny, bootstrapping is a resampling of the

characteristics used to determine the overall topology of the tree. The bootstrap value represents the reliability of each node, and nodes with values above 70% can be considered accurate 95% of the time (Hillis and Bull, 1993). Bootstrap values for clades represented in the phylogenetic tree generally showed good support, with the exception of clades Orientalis and Cholerae at 61% and 67% respectively (Table 2). The low bootstrap value associated with the Cholerae clade may be due to the inclusion of the Porteresiae clade, which was previously defined as an orphan clade (Sawabe et al., 2013).

Clade Node	Bootstrap value
Splendidus	100
Halioticoli	100
Nigripulchritudo	100
Mediterranei	100
Harveyi	100
Orientalis	61
Scopthalmi	93
Nereis	95
Coralliilyticus	93
Vulnificus	84
Anguillarum*	NA
Diazotrophicus	100
Cholerae	67
Gazogenes	100

Table 2: Bootstrap values for nodes of defined clades in ML Phylogenetic Tree List of bootstrap values for nodes containing organisms within a clade correlating to the maximum likelihood tree in Figure 8. Values were calculated by MEGA7 with 100 bootstrap replications. *Anguillarum is not represented by a single node and cannot be defined.

Although bootstrap values are not directly comparable between different phylogenetic trees, the clear representation of each clade in this study supports the previously defined phylogeny in Sawabe et al., 2013 with the exception of the Anguillarum clade and the orphan or single member clades. However, the purpose of this study was not to re-establish the overall phylogeny of the *Vibrio* genus but instead to provide data to test the placement of PINK within the Gazogenes clade. It should be noted that two *recA* genes are present within the PINK genome, indicating that a gene

duplication event has occurred. In the 5-gene MLSA scheme, both segments of the two genes used were identical, while in the Gazogenes clade 8-gene MLSA scheme the two segments of *recA* used shared 91% nucleotide sequence identity. Each *recA* sequence was tested separately and yielded the same position in the Gazogenes clade with only slight differences in the overall sequence identity to *V. ruber*.

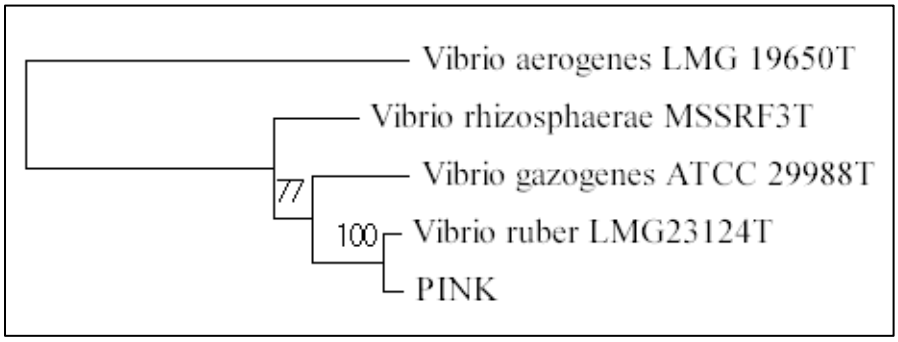

Figure 9: Maximum likelihood tree of Gazogenes Clade and PINK
Maximum likelihood tree from 8-gene MLSA of species within the Gazogenes clade generated using MEGA7. Numbers next to nodes indicate the bootstrap values for that particular node.

PINK was observed in the Gazogenes clade directly next to *V. ruber* and showed strong bootstrap support for both of these associations (Figure 8, Table 2, and Figure 9). Percent identity values for the 8-gene concatenated sequences in this clade showed 98% identity between PINK and *V. ruber,* with only 83%, 92%, and 92.5% sequence identity values between PINK and *V. aerogenes, V. rhizosphaerae* and *V. gazogenes* respectively. The species definition between two organisms is >98% sequence identity in MLSA studies (Sawabe et al., 2013). Data acquired in this study indicated that PINK is a strain of *V. ruber*.

Prodiginine Production in PINK

Prodigiosin Biosynthesis Cluster

The PINK prodiginine biosynthesis gene cluster is approximately 20.3 kilobases in length and consists of 13 genes (Figure 10). The overall structure of the PINK gene cluster was highly similar to prodiginine biosynthesis clusters in both in the arrangement of protein homologs as well as in the relative sizes of proteins and gaps within the cluster. Many of the open reading frames overlapped with one another on the ends, but a significant gap of 177 base pairs was found between *pigC* and *pigD*. This was comparable with the structure of the Pig clusters in *S. marcescens* ATCC 274 and Serratia 39006 as well as the Hap cluster in *H. chejuensis* (Figure 4) (Harris et al., 2004; Kim et al., 2006). This structure is very different from Red and Mar clusters found in *Streptomyces* species that produce undecyl prodigiosin and other cyclic prodiginines (Cerdeño et al., 2001; Salem et al., 2014). The only major difference between the structure of the prodigiosin cluster in PINK and *Sma* 274 is the absence of *pigA* in PINK. A nucleotide alignment of bases proceeding *pigB* with the coding nucleotide sequence for *pigA* in *S. marcescens* ATCC 274 showed short regions of homology (data not shown). This was an indicator that a PigA homolog may have been present at one time but has since been lost. Additionally another protein was found in the genome with homology that suggested it may be able to replace PigA (Table 3). This is highly speculative and further studies would be needed to confirm this.

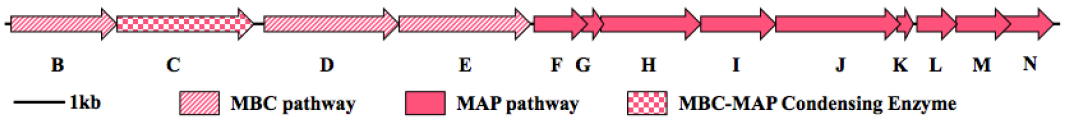

Figure 10: PINK Prodigiosin Biosynthesis Gene Cluster
Representation of prodigiosin gene cluster in PINK. The cluster is approximately 20.3 kb in length and contains 13 genes. Arrows show the directionality of open reading frames. Letters for ORFs correspond to the PigA-N homologs. Different arrow patterns indicate the putative role of the enzyme in the bifurcated prodiginine biosynthesis pathway.

In addition to the structural homology of the cluster, there was also consistent homology between genes in the PINK cluster and other genes relevant in the prodiginine pathway (Table 3). The amino acid identities (AAI), which is the ratio of similarity between two sequences based on alignment, between PINK proteins and Pig homologs typically ranged from 50% to 80%. Exceptions to this can be seen in PigL and PigN homologs, which had very low AAI; however, the lack of conservation of these two proteins was expected, as neither were required for the production of prodigiosin (Williamson et al., 2005). This divergence was also seen between PigL and HapL homologs (Dockyu et al., 2006).

PINK	*Sma* 274/*Serratia* 39006			*H. chejuensis*		*S. coelicolor* A3(2)	
	Homolog	AAI	AAI	Homolog	AAI	Homolog	AAI
A*	PigA	55	58	HapA	54	RedW	43
B	PigB	57	58	HapB	38	RedS	35
C	PigC	69	72	HapC	54	RedH	40
D	PigD	73	77	HapD	51	-	-
E	PigE	80	77	HapE	61	-	-
F	PigF	73	75	HapF	58	RedI	19
G	PigG	64	67	HapG	45	RedO	21
H	PigH	72	71	HapH	58	RedN	59
I	PigI	58	62	HapI	43	RedM	40
J	PigJ	61	59	HapJ	33	RedX	29
K	PigK	68	63	HapK	47	RedY	45
L	PigL	17	12	HapL	19	RedU	19
M	PigM	48	50	HapM	33	RedV	26
N	PigN	17	14	HapN	14	RedF	15

Table 3: Amino acid identity (AAI) of PINK Pig cluster Homologs
List of gene homologs from the PINK prodigiosin biosynthesis cluster with other known prodiginine biosynthesis genes from *S. marcescens* ATCC 274 (Sma 274), *Serratia* sp. ATCC 39006, *H. chejuensis* KCTC 2396 and *S. coelicolor* A3(2). Amino acid identities with PINK homologs were calculated using Clustal Omega. PigD and PigE homologs in *S. coelicolor* have not been found. *PigA homolog is not found within the PINK Pig cluster.

Amino acid identity is likely an improper method of comparison for protein function as sequence alignments reveal several conserved domains between PINK, Pig and Hap homologs. Conservation of specific domains is a better determinant as to whether a group of proteins will serve the same function. Although *H. chejuensis* and *S. marcescens* both produced prodigiosin, the amino acid identity between Pig and Hap homologs generally ranged from only 30% to 55% (Dockyu et al., 2008). Despite the

seemingly low amino acid identities between homologs, these enzymes served the same function in prodiginine biosynthesis (Williamson et al., 2005; Kim et al., 2008).

A better comparison of protein function can be done by comparing structural domains or specific amino acid residues that have been found essential to the functionality of the protein. The final reaction of the prodigiosin biosynthesis pathway is performed by PigC, which combines the MBC and MAP precursors. As expected, *pigC* knockout strains completely lost the ability to produce the prodigiosin pigment (Williamson et al., 2005).

The mechanism for PigC is similar to that of pyruvate phosphate dikinase (PPDK) enzymes in that they phosphorylate at a carbon-oxygen double bond by facilitating the transfer of a phosphoryl group from ATP using a histidine residue in a phosphoryl transfer domain (PTD). In PigC phosphorylation occurs at the aldehyde group of MBC. The addition of the phosphoryl group activates the adjacent carbon such that the MAP substrate binds to MBC, replacing the phosphoryl group in the process (Figure 11). Selective mutagenesis experiments on PigC in *Serratia* spp. 39006 showed that replacement of His840 with an alanine residue completely eliminated activity. Similarly, residues predicted to be important for ATP binding, Glu281 and Arg295, also showed significant reduction in activity when replaced with alanine residues (Chawrai et al., 2012). According to the NCBI protein basic local alignment tool, the PINK homolog of PigC contained a domain belonging to the PPDK superfamily. Closer analysis of the amino acid sequence in the PINK homolog also revealed residues His841, Glu282 and Arg296 that are analogous to key residues in PigC. Given this comparison it is likely that PigC and the PINK protein homolog function identically to one another.

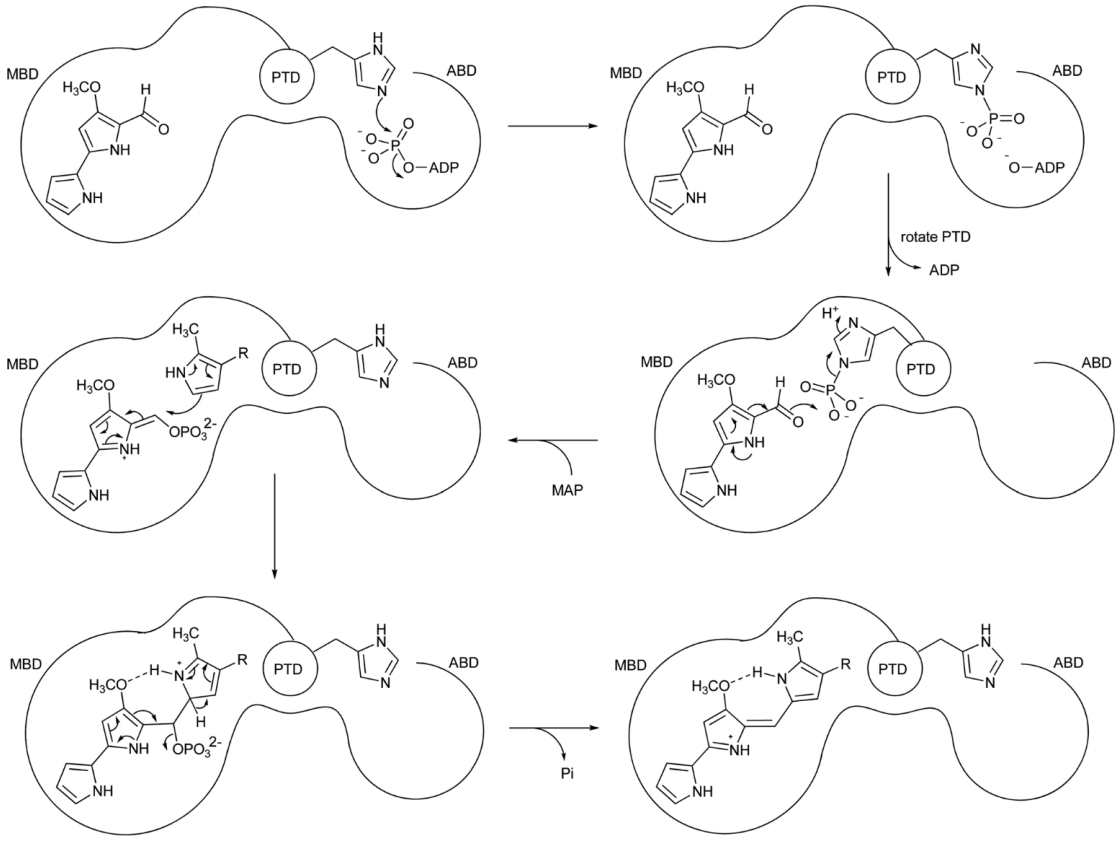

Figure 11: Mechanism for condensation of MBC and MAP by PigC
The mechanism for PigC utilizes energy from ATP to drive the condensation of MAP and MBC. A histidine residue on the phosphoryl transfer domain (PTD) facilitates the transfer of a phosphoryl group from an ATP molecule in the ATP binding domain (ABD). MAP is then able to replace the phosphoryl group and bind to MBC in the active site (Chawrai et al., 2011).

The sequence homology between all putative prodiginine biosynthesis genes in PINK and their homologs as well as the justification that a functional MBC/MAP condensing enzyme is conserved within the PINK gene cluster supports the idea that other proteins within the cluster maintain similar function to their Pig homologs (Table 4).

Protein	Pathway	Function
PigA	MBC	L-prolyl-PCP dehydrogenase
PigB	MAP	H_2MAP oxidase/dehydrogenase
PigC	Terminal step	Condensing enzyme
PigD	MAP	2-Acetyloctanal synthase
PigE	MAP	2-Acetyloctanal aminotransferase
PigF	MBC	HBC *O*-methyl transferase
PigG	MBC	Peptidyl carrier protein (PCP)
PigH	MBC	HBM synthase (seryl transferase)
PigI	MBC	L-prolyl-AMP ligase
PigJ	MBC	Pyrrolyl-β-ketoacyl ACP synthase
PigK	None	Unknown
PigL	MBC	4′-Phosphopantetheinyl transferase
PigM	MBC	HBM oxidase/dehydrogenase
PigN	MBC	Unknown

Table 4: Putative function of Prodigiosin Biosynthesis Genes
Proposed functions for enzymes in the prodigiosin biosynthesis gene cluster of *S. marcescens* strain ATCC 274 as determined with the use of knockout mutant strains for each enzyme and mass spectrometry to identify the products produced (modified from Williamson et al., 2005).

Structure and Characterization of Prodiginine Pigment in PINK

UV-Vis pH Spectra

The UV-Vis spectra of PINK pigment extract showed two main peaks that depend on the pH of the solvent (Figure 12). This suggests that the compound exists as isomers between protonated and non-protonated forms. The peak absorbance for what was presumed to be the protonated form occurred at a wavelength of approximately 535 nm and the non-protonated peak absorbance occurred at approximately 470 nm. These results were consistent with UV-Vis spectra for prodigiosin (Hearn et al., 1968).

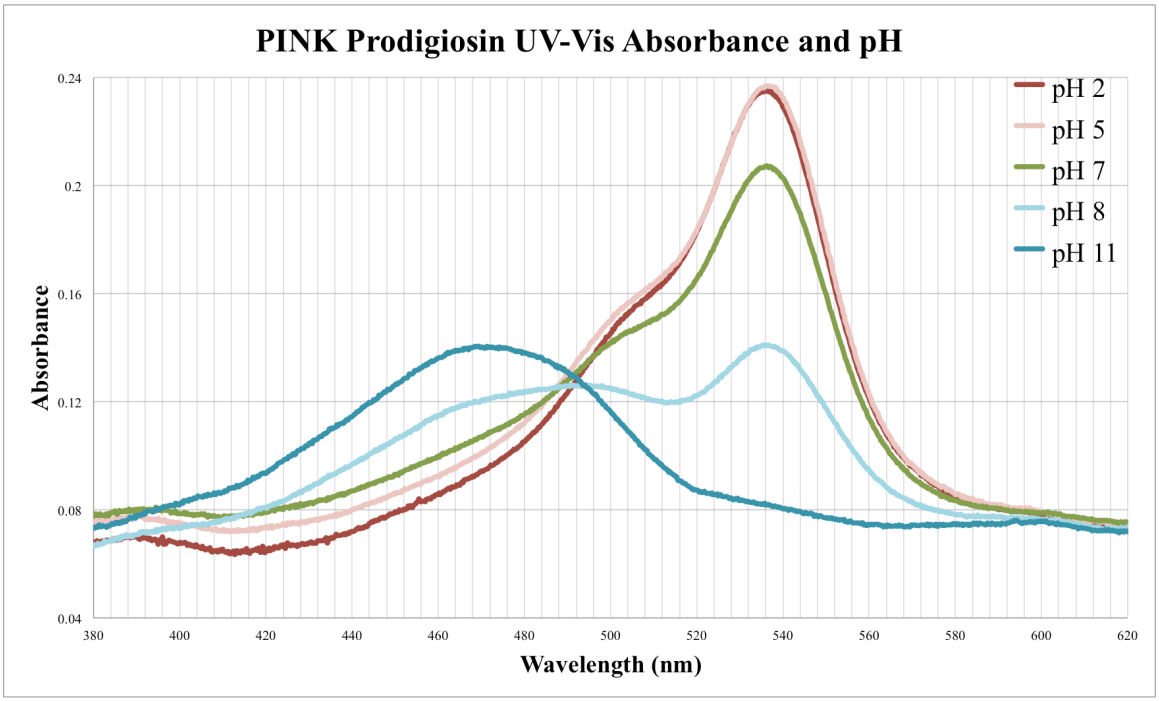

Figure 12: UV-Vis Spectra of PINK pigment extract at various pHs
The graph shows the absorbance from wavelengths ranging from 380 nm to 620 nm of PINK pigment extract in aqueous methanol solvents at varying pHs. Two maxima can be observed at 470 nm and 535 nm. Absorbance at 540 nm is greater at lower pH while absorbance at 470 nm is greater at basic pH. An isosbestic point can be observed at 490 nm.

It is known that prodigiosin exists as a pair of α (trans) and β (cis) conformational isomers due to the rotation about the carbon-carbon double bond between pyrrole rings B and C. At low pH, the protonation of nitrogen on pyrrole ring C stabilizes a 180° rotation of the ring by facilitating hydrogen bond interactions between the oxygen of the methoxy group and the hydrogen on the nitrogen of pyrrole C to form the α (trans) isomer. Likewise, when protonation does not occur, the β (cis) isomer is more predominant (Rizzo et al., 1998). The spectra for the PINK pigment extract revealed an isosbestic point at a wavelength of approximately 490 nm. This is the wavelength in which the absorbance remains constant as a conformational change occurs due to the changes in pH of the aqueous methanol. Additionally, the data suggested that the pH at which the

absorbance between the two peaks would be equal was just above pH 8, which is consistent with the known pKa value of 8.25 for prodigiosin in acidified ethanol (Hearn et al., 1968).

Prodigiosin Extraction Efficacy

The efficacy of the prodigiosin extraction protocol used in this study can be measured by the results from the mass spectra (shown below) as well as by performing thin layer chromatography (TLC) on aliquots taken throughout the extraction procedure.

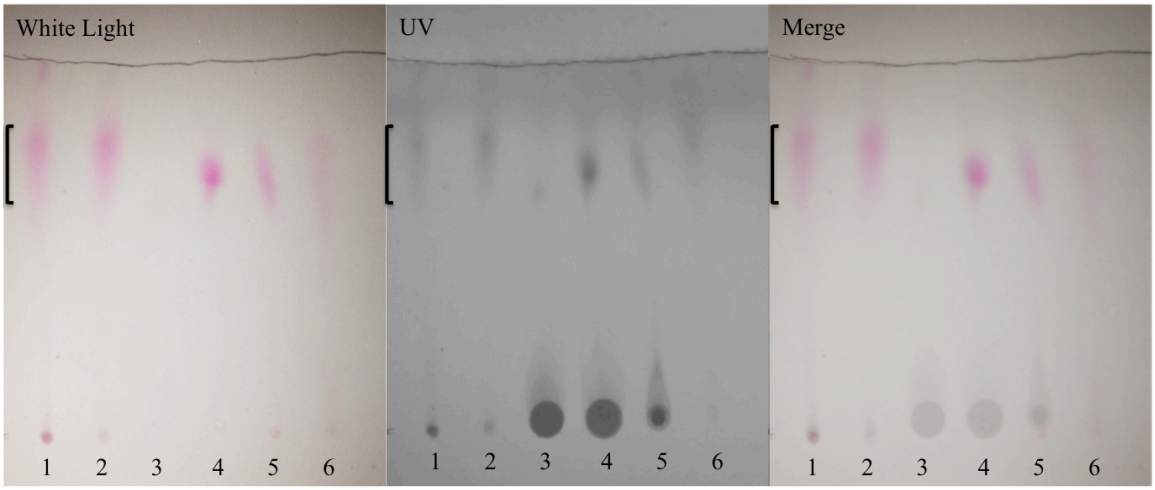

Figure 13: TLC of Stepwise Extraction Aliquots
Images from TLC run with silica as the stationary phase and acetonitrile as the mobile phase. White light showed the presence of pink pigment within the extract sample. Images shot with a UV light source were made grayscale for clarity. Spots 1 and 2 were from the chloroform layer of a chloroform-water liquid-liquid extraction. Spot 1 was 2 months old while spot 2 was only a week old. Spot 3 was the water layer from the chloroform-water liquid-liquid extraction. Spot 4 was an aliquot from before the chloroform-water extraction. Spot 5 was an initial aliquot taken from pelleted cells directly from culture. Spot 6 was post flash-chromatography. The bracket shows the target of the purification at an Rf of 0.6 to 0.82.

TLC of aliquots taken throughout the extraction procedure helped illustrate the progression from the crude extract to the purified product (Figure 13). However, it was difficult to truly compare the relative abundance of impurities within each aliquot because they were not at a fixed concentration. The appearance of a smear rather than a single spot could be caused by several different factors. One possible factor is that the

compound may start to degrade as it moves up the silica plate. Prodigiosin is known to degrade after exposure to light (as reviewed by Fürstner, 2003). Another factor might be the difference in relative acidity between the stationary phases. Prodiginine interactions with the silica stationary phase, which is slightly acidic, would cause prodigiosin to favor the α conformation while solvent interactions with acetonitrile would cause prodigiosin to prefer the β conformation (Rizzo et al., 1998). This conformation change may affect the solubility in the mobile phase as the molecules move up the plate.

Another interesting thing to note is the appearance of additional pink spots around Rf 1 and at the origin that may be additional prodiginine compounds. Spots at the origin represent highly polar molecules that are not soluble in acetonitrile. Therefore, it is suggested that these compounds may be oxidized prodiginine derivatives.

Mass Spectra of Purified Pigment Extract

Mass spectrometry is a powerful tool that can be used to help elucidate the structure of a compound by detecting the mass of a compound based on the mass to charge ratio of an ion at the detector. In tandem mass spectrometry (MS/MS or MS^2), ions coming from an ion source are detected after running through a mass spectrometer and are then trapped in a linear ion trap. From there, selected ions of a particular mass or range of masses can be run through a second mass spectrometer where a gas (He in this case) bombards ions to cause collision-induced dissociation (CID). The resulting fragments are then read by a second detector to produce mass spectra.

Mass spectra of purified PINK pigment extract revealed a mixture of 4 different compounds in different abundances. A compound with a mass of 324 g/mol was the main

component of the mixture, and compounds with masses of 338 g/mol, 352 g/mol and 371 g/mol were present in the extract as only minor components (Figure 14). Tandem mass spectra of compounds with a mass of 324, 338 and 352 exhibited similar fragmentation patterns that were consistent with prodiginine compounds (Figure 14, 15 & 16). This was evident in that all 3 contain a peak at 252 m/z, which can be explained by the loss of the alkyl substituent group to yield the 2-methoxy prodiginine core. Additionally, all three showed an ion at a mass consistent with the loss of the methyl group from the methoxy. Compounds 324, 338 and 352 can be represented by compounds with varying lengths of alkyl chains, as there is a mass difference of 14 that can be represented by the addition of a CH_2. Prodiginine compounds varying in only the length of the alkyl chain substituent group on pyrrole C are well documented and have been produced as minor byproducts in other species (Gerber et al., 1975; Alihousseini et al., 2009**;** Lee et al., 2011).

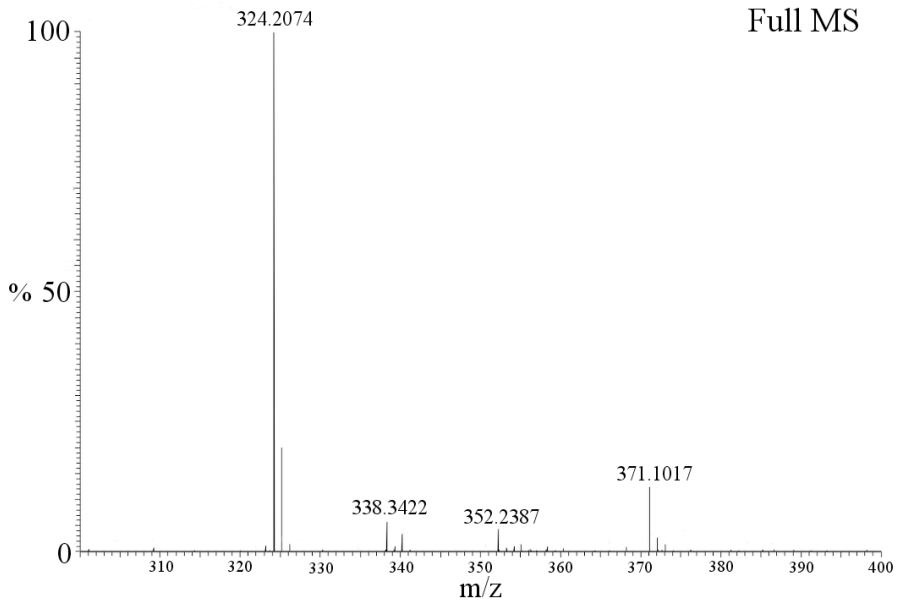

Figure 14: Electrospray ionization mass spectra of PINK pigment extract
Peaks represent ions of a particular mass (x-axis) with a relative abundance (y-axis). The peak at 324 is the most prominent while peaks at 338, 352 and 371 appear as minor components.

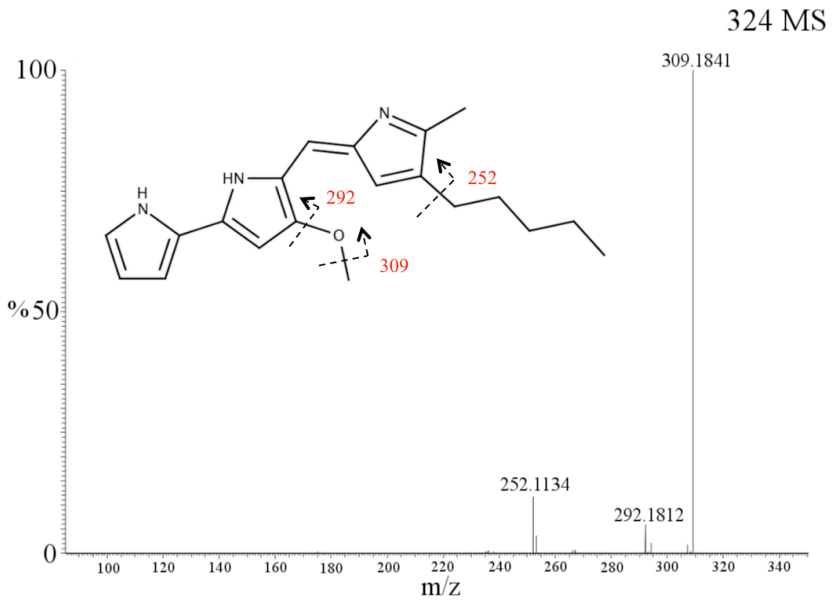

Figure 15: Tandem mass spectra of a molecule with a molecular mass of 324 with the proposed structure, prodigiosin

Peaks represent ions produced as a result of gas-induced dissociation (CID) of the parent molecule. The most stable ion produced appears at m/z 309 and is the ion produced by cleavage of the methyl on the methoxy group. The peak at 252 represents an ion created as a result of the alkyl chain being cleaved from the molecule and 292 is the loss of the oxygen. These fragmentation patterns are shown with the dashed arrows.

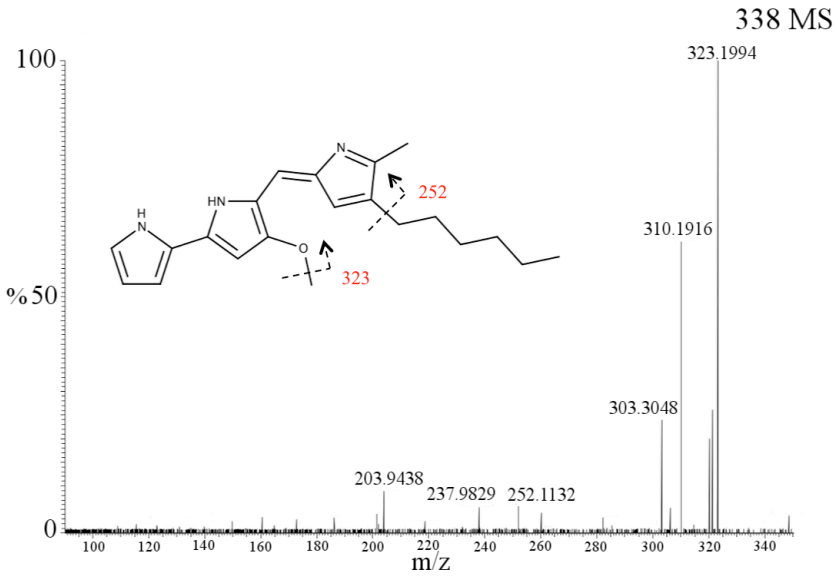

Figure 16: Tandem mass spectra of a molecule with a molecular mass of 338 with the proposed structure, 2-methyl-3-hexyl prodiginine

Peaks represent ions produced as a result of gas-induced dissociation (CID) of the parent molecule. The most stable ion produced appears at m/z 323 and is the ion produced by cleavage of the methyl on the methoxy group. The peak at 252 represents an ion created as a result of the alkyl chain being cleaved from the molecule, and the peak at 310 is likely loss of an ethene. Other peaks represented are not readily explainable given the proposed structure.

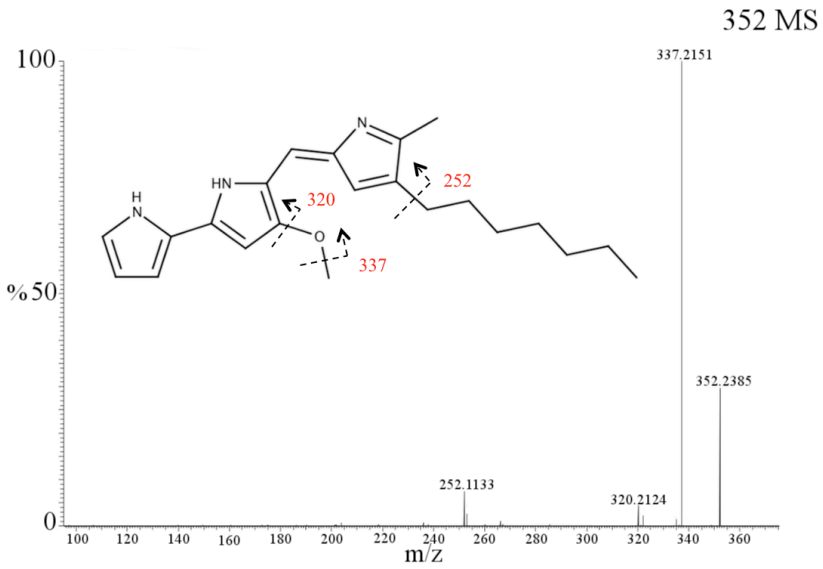

Figure 17: Tandem mass spectra of a molecule with a molecular mass of 352 with the proposed structure, prodigiosin

Peaks represent ions produced as a result of gas-induced dissociation (CID) of the parent molecule. The most stable ion produced appears at m/z 337 and is the ion produced by cleavage of the methyl on the methoxy group. The peak at 252 represents an ion created as a result of the alkyl chain being cleaved from the molecule, and 320 is produced after the loss of the oxygen.

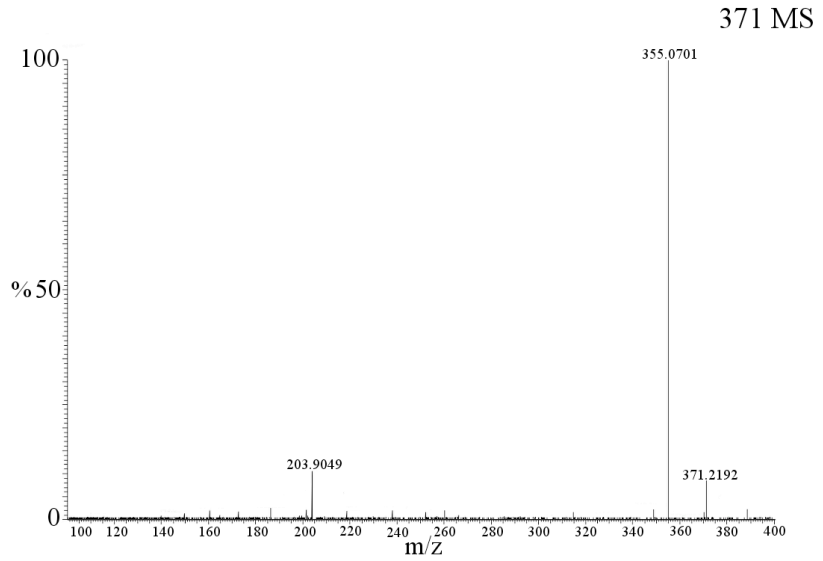

Figure 18: Tandem mass spectra of a molecule with a molecular mass of 371

Peaks represent ions produced as a result of gas-induced dissociation (CID) of the parent molecule. Lack of a peak at 252 suggests that the parent molecule likely does not contain a prodiginine core and is an unrelated compound.

The compound with a mass of 271 appeared to be unrelated as it did not contain the prodiginine core 252 peak; however, further analysis would be required to fully

elucidate the structure. It should also be noted that it is likely that PINK produces other prodiginine derivatives, as not all pigment compounds observed could be purified and analyzed (see discussion).

Discussion

Bacterial Taxonomy is Problematic

For bacteria, the term "species" does not have the same biological merit as for eukaryotic organisms. Instead the term is loosely defined based on the differences between organisms as determined by a polyphasic approach that considers both genomic divergence and phenotypic characterization (Rossello-Mora and Amann, 2001). As of July 2013, over 10,000 different bacterial species were recognized with a validly published name, but only about 25% have a strain with the entire sequenced genome available (Chun and Rainey, 2014). This number is far from a true representation of prokaryotic diversity, as many microbial environments remain largely unexplored. Additionally, the number of prokaryotes that are able to be cultured in the lab is drastically less than what is present based on analyses of DNA extracts taken from environmental soil samples (Curtis et al., 2006). With such a largely diverse and expansive group of organisms, it calls into question whether a species classification system is really practical.

The definition for species compared using an 8-gene MLSA scheme is a nucleotide sequence identity above 98% (Sawabe et al., 2013). By this definition, PINK would be considered a strain of *V. ruber*. Naturally, the next step to confirm these results would be to perform overall genome comparison analyses between PINK and *V. ruber*; however, the genomic sequence for *V. ruber* is not present in the NCBI database. Conflicting with the MLSA results are the observed differences in biochemical

characteristics between PINK and *V. ruber* (Table 1). The conflicting results found in this study echo the general difficulties with bacterial taxonomy.

V. rhizosphaerae, *V. gazogenes* and *V. ruber* are all marine, prodigiosin producing organisms with little variation in biochemical characteristics (Ramesh kumar and Nair, 2007; Shieh et al., 2003; Harwood, 1978). Biochemical characteristics of *V. gazogenes* is also referenced as being variable among different strains of that organism, making species classification based on biochemical characterization difficult (Noguerola and Blanch, 2008). The rationale for a polyphasic approach in which both phenotypic and genomic factors are considered is that a species must show phenotypic consistency (Van damme et al., 1996). The notion that multiple, very similar species can be defined based on genomic differences that do not reflect appreciable differences in the phenotype appears trivial. On the other hand, the phenotype of an organism can also appear quite different with very little overall divergence in the genome. This divide between a species identification system based on genomic divergence as measured by conserved housekeeping genes and phenotypic characterization makes a polyphasic approach impractical.

A recent study of the *Vibrio* genus used a computational approach to locate genes responsible for given phenotypes in a species (Amaral et al., 2014). Integrating these particular genes in genomic divergence studies and phenotyping may lead to a more consistent taxonomic system. Other suggestions call for a larger focus on differences in genomic data using methods like MLSA (Konstantinidis and Tiedje, 2005; Thompson et al., 2015). Overall, the threshold for what constitutes a new bacterial species should be reconsidered for practicality. The implementation of numbered or phenotypic description

based strain names for a group of highly similar bacteria might be more reasonable than using several different species names.

Prodigiosin Biosynthesis Clusters in *Vibrio* Organisms

Although prodiginine compounds have been studied for a long time, there is limited research on prodigiosin biosynthesis clusters in different genera. To date prodigiosin biosynthesis clusters have only been examined in *Serratia* spp. 39006, *S. marcescens* ATCC 274, *H. chejuensis* KCTC 2396 and some *Streptomyces* species (Harris et al, 2004; Williamson et al., 2005, Kim et al., 2008; Cerdeño et al., 2001, Salem et al., 2014). The addition of another prodiginine biosynthesis cluster in *Vibrio* species will add more information on the potential mechanism for each enzyme by revealing key residues in conserved regions of each gene.

NCBI blastp results show high sequence similarity between PINK Pig gene homologs and proteins annotated in both the *V. rhizosphaerae* and *V. gazogenes* genomes (data not shown). The recognition of these genes as Pig homologs could help to guide future studies on the prodiginine biosynthesis genes in *Vibrio* organisms. For example, cycloprodigiosin was found to be produced by *V. gazogenes* (Alihosseini et al., 2008). The biosynthesis of cycloprodigiosin has yet to be elucidated, likely due to the lack of information on prodiginine gene clusters capable of producing cycloprodigiosin. The genome of *Pseudoalteromonas rubra,* another cycloprodigiosin producer, contains a prodiginine gene cluster similar to *H. chejuensis,* although no further studies have been done with this cluster (Xie et al., 2012).

Another study used non-specific chemical mutagenesis to create prodiginine mutants in *V. gazogenes*. The study found that certain mutants were able to produce

higher amounts of prodigiosin as well as different proportions of specific prodiginine derivatives, such as cycloprodigiosin (Alihosseini et al., 2009). Unfortunately, no research was done on the gene sequences of the prodiginine biosynthesis genes in wild type or mutant strains. A comparative analysis between Pig clusters in both wildytpe and mutant *V. gazogenes* strains*, P. rubra* and other similar clusters incapable of producing cycloprodigiosin could point to the key proteins necessary for producing cycloprodigiosin.

Currently, putative prodiginine biosynthesis genes are being cloned into plasmids and transformed into mutants of *Serratia marcescens* with that gene knocked out (personal communication, Brittany Gasper of FL Southern College). This will confirm the function of each gene in the PINK cluster with the Pig gene homologs in *Serratia marcescens*.

Prodiginine Compounds Produced by PINK

The prodiginine compounds isolated and identified in this study have been previously recognized and documented in several other species (Lee et al., 2011; Alihosseini et al., 2009; Kim et al., 2006). Given the homology between Pig, Hap and PINK prodiginine gene clusters and the lack of homology with the Red cluster, it makes sense that PINK would produce similar prodiginine derivatives to those found in *Serratia* species and *H. chejuensis* rather than the prodiginine derivatives produced by some *Streptomyces* species. *Streptomyces* species containing Red or Mar clusters have been found to produce undecylprodigiosin as well as cyclic prodiginines such as marineosins (Salem et al., 2014). The difference in PINK appearance on the modified PDA and the

LB plates may be due to the amount of overall prodiginine production and the ratio of different prodiginine compounds produced on each media. This ratio has been previously correlated with color change in *V. gazogenes* (Alihosseini et al., 2009).

Additional pigment spots on the TLC indicated that there may be additional prodiginine compounds produced by PINK (Figure 13). The isolation of these compounds may be achieved using a chromatography procedure with higher resolution. In other studies, HPLC was used with a reversed-phase C-18 column to purify prodiginines (Montaner et al., 2000; Alihosseini et al., 2009). Preparative TLC has also been used to purify different prodiginines (Casullo de Araújo et al., 2010; Nakashima et al., 2005). Additionally, use of acetonitrile as a solvent for flash chromatography may have impeded separation by not stabilizing a single conformation during separation techniques. Prodigiosins prefer the β conformer in acetonitrile and would likely prefer the α conformer in response to interactions with the silica, which is slightly acidic (Rizzo et al., 1998). A similar separation was achieved using ratios of chloroform, methanol and ethanol but was not extensively tested due to the convenience of using a single solvent; however, a better separation might be achieved by experimenting with the ratios further.

Finally, prodiginine compounds have been well documented as possessing antibacterial capabilities against Gram negative and Gram positive organisms (as reviewed by Darshan and Manomanni, 2015). The initial antibacterial properties observed in PINK may be due to the production of prodiginine compounds (Figure 6). Antibacterial growth inhibition assays using purified prodiginine extract from PINK must be repeated to confirm this. It is not clear whether another antibiotic is produced by PINK or not. Bioautography is a technique typically used to test for antibiotic compounds in

plant tissues and combines TLC with agar overlay growth inhibition assays (Dewanjee et al., 2014). Using bioautography techniques on crude extract could show the production of more than one antibiotic.

Conclusion

From the discovery and characterization of PINK some conclusions can be made. Multi-locus sequence analysis using 5 conserved genes places PINK as a member of the Gazogenes clade within the Vibrionacaea family. An additional MLSA using 8 conserved genes shows high sequence similarity to *Vibrio ruber,* indicating that PINK is a strain of *V. ruber.* However, biochemical characterization of PINK does not agree with the MLSA results, raising several questions about the definition of a species in bacterial taxonomy.

Analysis of the PINK genome revealed a Pig cluster containing 13 genes with significant homology to previously defined clusters in *Serratia* species and *H. chejuensis.* The inclusion of a *Vibrio* prodiginine biosynthesis cluster in other Pig studies will provide a greater understanding on the biosynthesis of prodiginine compounds.

Prodiginine compounds isolated from PINK include prodigiosin derivatives with substituent group alkyl chains C5, C6 and C7. However, TLC results suggest that other pigment compounds remain to be examined. The antibacterial activity of PINK initially detected may be due to the production of prodiginine derivatives or the production of other antibiotics in PINK, and further testing is required to determine this.

Appendices

Appendix A: Media Components

LB Agar with 4% NaCl

Ingredient	Concentration
Tryptone	10 g/L
Yeast Extract	5 g/L
NaCl	40 g/L
Agar	20 g/L

Modified PDA with 4% NaCl

Ingredient	Concentration
Tryptone	10 g/L
Potato Starch	4 g/L
NaCl	40 g/L
Glucose*	5 mMol
10X Neidhardt MOPS salts*[+]	0.1X
Agar	20 g/L

*Must be added by filter sterilization after autoclaving
[+]Neidhart MOPS salts components (Neidhardt et al., 1974)
<https://www.genome.wisc.edu/resources/protocols/mopsminimal.htm>

Appendix B: PINK DNA Extraction Procedure

Performed at Florida Southern University by Brittany Gasper and Eric Warrick
Obtained through personal communication with Eric Warrick.

An aliquot of 5 ml of an overnight culture of PINK grown in LB were harvested by centrifugation at 8,000 rcf for 10 minutes. The LB was poured off and the cell pellet was resuspended in 5 ml TE buffer at pH 8.0. The cells were pelleted by centrifugation at 8,000g for 10 minutes. This washing step was repeated a total of 3 times. After the final wash, the cell pellet was resuspended in 2.5 ml of TE buffer. 250 µl of a 10% SDS

solution was added to the cell suspension and mixed by inversion. Then 9 µl of a solution consisting of Proteinase K at 20 mg/ml and RNAse A at 10 mg/ml was added to the cell suspension. After mixing by inversion, the cell suspension was place at 37°C for 1 hour. Then a wash step was performed by adding 500 µl of room temperature buffer phenol at pH 8.0, 500 µl of chloroform and 50 µl of isoamyl alcohol and mixing by inversion. The mixture was centrifuged at 10,000 rcf for 5 minutes, the aqueous layer was transferred to a fresh tube. This wash step was repeated two more times. The aqueous layer was then transferred to a fresh tube and 750 µl of chloroform was added, mixed by inversion and centrifuged at 10,000 rcf for 5 minutes. The aqueous layer was transferred to a fresh tube and this step was repeated. Then the aqueous layer was transferred to a fresh tube and placed on ice as 1/10 volume of 3M NaOAc was added to the aqueous solution and 2.5x volume of ice cold 100% EtOH was added and gently mixed by inversion. The mixture underwent centrifugation at 10,000 rcf for 10 minutes. All of the liquid was carefully poured off and the tube was allowed to dry using a vacuumed desiccator. After drying, 500 µl of molecular grade water was used to resuspend the DNA. Quantification was conducted using a nanodrop 2000.

Appendix C: Genbank Accession Numbers for Sequences Used

Genbank Gene Accession Numbers for genes used in 5-gene MLSA scheme

Vibrio Species	Gene Accession Number				
	recA	*rpoA*	*gapA*	*ftsZ*	*gyrB*
V. aerogenes	AJ842368	AJ842553	DQ907270	DQ907330	KF697270
V. aestuarianus	AJ842369	AJ842554	DQ907271	DQ907331	AB298200
V. alginolyticus	AJ842373	AJ842558	DQ907274	EF027344	AF007288
V. anguillarum	AJ580852	AJ842561	DQ907275	DQ907334	AB298203
V. azureus	FR669656	AB465316	AB428907	AB428903	AB428899
V. brasiliensis	AJ842376	HM771384	DQ449613	DQ449613	AB298204

V. campbellii	EU652261	AJ842564	DQ449614	JF836300	EF596590
V. chagasii	AJ842385	AJ842572	DQ481611	DQ996590	AB298206
V. cholerae	AM942078	HE805630	HE805629	HE805627	FM202624
V. cincinnatiensis	AJ842397	AJ842582	DQ907278	DQ907340	AB298208
V. communis (owensii)	GU078693	GU078697	AB609125	AB609124	GU078680
V. corallilyticus	AJ842402	AJ842587	DQ907279	DQ907341	AB298210
V. crassostreae	EU541594	EU541574	DQ481612	DQ481624	AB298209
V. cyclitrophicus	AJ842405	AJ842592	DQ481613	DQ481625	DQ164536
V. diazotrophicus	AJ842411	AJ842598	DQ907280	DQ907342	AB298212
V. ezurae	AJ842413	AJ842600	AY546645	DQ907343	AB298213
V. fluvialis	AJ842419	AJ842606	DQ907281	DQ907345	AB298215
V. fortis	AJ842422	AJ842609	DQ907282	DQ907346	AB298216
V. furnissii	AJ842427	AJ842614	DQ907283	EF027345	AB298217
V. gallicus	AJ842428	EU541581	AY546654	KF697254	AB298218
V. gazogenes	KF697297	KF697306	DQ907284	DQ907348	AB298258
V. gigantis	EU541593	EU541573	DQ481617	DQ481629	AB298219
V. halioticoli	AJ842430	AJ842617	AY546638	DQ907349	AB298220
V. hepatarius	AJ842444	AJ842631	DQ907285	DQ907352	AB298222
V. hispanicus	AJ842445	AJ842632	DQ907286	DQ907353	AB298223
V. ichthyoenteri	HM771380	HM771385	HM771360	HM771355	HM771365
V. kanaloae	AJ842450	AJ842637	DQ481619	DQ481631	AM162563
V. maritimus	GU929935	GU929937	KF666681	GU929927	GU929929
V. mediterranei	AJ842459	AJ842644	DQ907290	DQ907356	EF380258
V. metschnikovii	AJ842466	AJ842650	DQ907291	EF027346	AB298229
V. mimicus	EF643485	EF643486	DQ907292	DQ907357	EF380259
V. mytili	HQ455540	AJ842657	DQ907293	DQ907358	AB298231
V. natriegens	AJ842473	AJ842658	DQ907294	DQ907359	AB298232
V. navarrensis	AJ842474	AJ842659	DQ907295	DQ907360	AB298233
V. neonatus	KF697298	KF697307	AY546644	DQ907363	AB298257
V. neptunius	AJ842478	AJ842663	DQ907296	DQ907361	AB298234
V. nereis	AJ842479	AJ842666	DQ449617	DQ907362	AB298235
V. nigripulchritudo	AJ842480	AJ842667	DQ907297	EF027347	AB298236
V. ordalii	AJ842482	AJ842669	DQ907298	DQ907364	AB298237
V. orientalis	EU130528	AJ842672	DQ907299	DQ907365	EF380260
V. pacinii	AJ842486	AJ842674	DQ907300	DQ907366	AB298264
V. parahaemolyticus	AJ842490	AJ842677	DQ449618	DQ907367	EU130502
V. pectenicida	AJ842491	AJ842678	DQ907301	DQ907368	AB298240
V. penaeicida	AJ842496	AJ842683	DQ907303	DQ907370	AB298260
V. pomeroyi	AJ842497	AJ842684	DQ481621	DQ481634	AJ577822
V. ponticus	JQ308799	JQ308806	DSM16217	DQ907371	AB298243
V. porteresiae	EF547199	EF547200	KF697266	KF697257	KF697274
V. proteolyticus	AJ842499	AJ842686	DQ907305	EF114210	AB298261

V. rhizosphaerae	KF697300	KF697309	EU713846	FJ876000	EU713847
V. rotiferianus	AJ842501	AJ842688	DQ449619	DQ907372	AB298244
V. ruber	KF697301	KF697310	DQ907306	DQ907373	AB298267
V. rumoiensis	AJ842503	AJ842690	AB298245	DQ907374	AB298245
V. scopthalmi	HM771381	HM771386	DQ907309	DQ907376	AB298247
V. sinaloensis	HM771382	HM771387	HM771362	HM771357	HM771367
V. splendidus	EU130529	AJ842725	DQ481622	DQ481635	EF380261
V. superstes	KF697303	EU541580	AY546650	DQ907378	AB298256
V. tasmaniensis	AJ842515	AJ842731	DQ481623	DQ481636	AB298250
V. variabilis	GU929934	GU929936	KF666686	GU929926	GU929928
V. vulnificus	AJ842523	AJ842737	DQ907313	DQ907382	AY705491
V. xuii	AJ842529	AJ842742	DQ907315	DQ907384	AB298254

Genbank Gene Accession Numbers for genes used in 8-gene MLSA scheme

Vibrio Species		*V. aerogenes*	*V. gazogenes*	*V. rhizosphaerae*	*V. ruber*
Strain		LMG 19650T	ATCC 29988T	MSSRF3T	LMG23124
Gene Accession Number	*recA*	AJ842368	KF697297	KF697300	KF697301
	rpoA	AJ842553	KF697306	KF697309	KF697310
	gapA	DQ907270	DQ907284	EU713846	DQ907306
	ftsZ	DQ907330	DQ907348	FJ876000	DQ907373
	mreB	DQ907401	DQ907420	KF697283	EF114211
	topA	EF114212	KF697315	KF697318	EF114214
	gyrB	KF697270	AB298258	EU713847	AB298267
	pyrH	KF697286	KF697288	KF697291	KF697292

Prodiginine Biosynthesis Cluster Genbank

Organism	Genbank Accession #
Serratia marcescens ATCC 274	AJ833002
Serratia 39006	AJ833001
Hahella chejuensis KCTC 2396	DQ266254
Streptomyces coelicolor	AL645882

Works Cited

Alihosseini F, Lango J, Ju K, Hammock BD, Sun G. 2009. Mutation of bacterium *Vibrio gazogenes* for selective preparation of colorants. Biotechnol Prog 26(2):352-60.

Alihosseini F, Ju K, Lango J, Hammock BD, Sun G. 2008. Antibacterial colorants: Characterization of prodiginines and their applications on textile materials. Biotechnol Prog 24(3):742-7.

Amaral GRS, Dias GM, Wellington-Oguri M, Chimetto L, Campeão ME, Thompson FL, Thompson CC. 2014. Genotype to phenotype: Identification of diagnostic *Vibrio* phenotypes using whole genome sequences. Int J Syst Evol Microbiol 64(2):357-65.

Casullo de Araújo, Helvia W, Fukushima K, Takaki GMC. 2010. Prodigiosin production by *Serratia marcescens* UCP 1549 using renewable-resources as a low cost substrate. Molecules 15(10):6931-40.

Cerdeño AM, Bibb MJ, Challis GL. 2001. Analysis of the prodiginine biosynthesis gene cluster of *Streptomyces coelicolor* A3 (2): New mechanisms for chain initiation and termination in modular multienzymes. Chem Biol 8(8):817-29.

Chawrai SR, Williamson NR, Mahendiran T, Salmond GP, Leeper FJ. 2012. Characterisation of PigC and HapC, the prodigiosin synthetases from *Serratia* sp. and *Hahella chejuensis* with potential for biocatalytic production of anticancer agents. Chemical Science 3(2):447-54.

Chun J and Rainey FA. 2014. Integrating genomics into the taxonomy and systematics of the bacteria and archaea. Int J Syst Evol Microbiol 64(2):316-24.

Curtis TP, Head IM, Lunn M, Woodcock S, Schloss PD, Sloan WT. 2006. What is the extent of prokaryotic diversity? Philos Trans R Soc Lond B Biol Sci 361(1475):2023-37.

Darshan N and Manonmani H. 2015. Prodigiosin and its potential applications. Journal of Food Science and Technology :1-15.

Dewanjee S, Gangopadhyay M, Bhattacharya N, Khanra R, Dua TK. 2015. Bioautography and its scope in the field of natural product chemistry. Journal of Pharmaceutical Analysis 5(2):75-84.

Fehér D, Barlow RS, Lorenzo PS, Hemscheidt TK. 2008. A 2-substituted prodiginine, 2-(p-hydroxybenzyl) prodigiosin, from *Pseudoalteromonas rubra*. J Nat Prod 71(11):1970-2.

Fürstner A. 2003. Chemistry and biology of roseophilin and the prodigiosin alkaloids: A survey of the last 2500 years. Angewandte Chemie International Edition 42(31):3582-603.

Gabriel MW, Matsui GY, Friedman R, Lovell CR. 2014. Optimization of multilocus sequence analysis for identification of species in the genus *Vibrio*. Appl Environ Microbiol 80(17):5359-65.

Gerber NN. 1975. Prodigiosin-like pigments. CRC Crit Rev Microbiol 3(4):469-85.

Gish W and States DJ. 1993. Identification of protein coding regions by database similarity search. Nat Genet 3(3):266-72.

Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R. 2010. A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res 38(Web Server issue):W695-9.

Harris AK, Williamson NR, Slater H, Cox A, Abbasi S, Foulds I, Simonsen HT, Leeper FJ, Salmond GP. 2004. The *Serratia* gene cluster encoding biosynthesis of the red antibiotic, prodigiosin, shows species-and strain-dependent genome context variation. Microbiology 150(11):3547-60.

Harwood C. 1978. *Beneckea gazogenes* sp. nov., a red, facultatively anaerobic, marine bacterium. Curr Microbiol 1(4):233-8.

Hearn WR, Medina J, Elson MK. 1968. Colour change of prodigiosin. .

Hearn WR, Worthington R, Burgus RC, Williams RP. 1964. Norprodigiosin: Occurrence in a mutant of *Serratia marcescens*. Biochem Biophys Res Commun 17(5):517-20.

Hillis DM and Bull JJ. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst Biol 42(2):182-92.

Kim D, Park YK, Lee JS, Kim JF, Jeong H, Kim BS, Lee CH. 2006. Analysis of a prodigiosin biosynthetic gene cluster from the marine bacterium *Hahella chejuensis* KCTC 2396. Journal of Microbiology and Biotechnology 16(12):1912.

Kim D, Kim JF, Yim JH, Kwon SK, Lee CH, Lee HK. 2008. Red to red - the marine bacterium *Hahella chejuensis* and its product prodigiosin for mitigation of harmful algal blooms. J Microbiol Biotechnol 18(10):1621-9.

Kobayashi N and Ichikawa Y. 1989. A protein associated with prodigiosin formation in *Serratia marcescens*. Microbiol Immunol 33(4):257-63.

Konstantinidis KT and Tiedje JM. 2005. Genomic insights that advance the species definition for prokaryotes. Proc Natl Acad Sci U S A 102(7):2567-72.

Kumar S, Stecher G, Tamura K. 2015. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol .

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al. 2007. Clustal W and clustal X version 2.0. Bioinformatics 23(21):2947-8.

Lee JS, Kim YS, Park S, Kim J, Kang SJ, Lee MH, Ryu S, Choi JM, Oh TK, Yoon JH. 2011. Exceptional production of both prodigiosin and cycloprodigiosin as major metabolic constituents by a novel marine bacterium, *Zooshikella rubidus* S1-1. Appl Environ Microbiol 77(14):4967-73.

Montaner B, Navarro S, Piqué M, Vilaseca M, Martinell M, Giralt E, Gil J, Pérez-Tomás R. 2000. Prodigiosin from the supernatant of *Serratia marcescens* induces apoptosis in haematopoietic cancer cell lines. Br J Pharmacol 131(3):585-93.

Morrison DA. 1966. Prodigiosin synthesis in mutants of *Serratia marcesens*. J Bacteriol 91(4):1599-604.

Mount DW. 2008. Choosing a method for phylogenetic prediction. CSH Protoc 2008:pdb.ip49.

Nair S and N. Ramesh kumar. 2007. *Vibrio rhizosphaerae* sp. nov., a red-pigmented bacterium that antagonizes phytopathogenic bacteria. Int J Syst Evol Microbiol 57(10):2241-6.

Nakashima T, Kurachi M, Kato Y, Yamaguchi K, Oda T. 2005. Characterization of bacterium isolated from the sediment at coastal area of omura bay in japan and several biological activities of pigment produced by this isolate. Microbiol Immunol 49(5):407-15.

Noguerola I and Blanch A. 2008. Identification of *Vibrio* spp. with a set of dichotomous keys. J Appl Microbiol 105(1):175-85.

Pascual J, Macián MC, Arahal DR, Garay E, Pujalte MJ. 2010. Multilocus sequence analysis of the central clade of the genus vibrio by using the 16S rRNA, *recA*, *pyrH*, *rpoD*, *gyrB*, *rctB* and *toxR* genes. Int J Syst Evol Microbiol 60(1):154-65.

Rizzo V, Morelli A, Pinciroli V, Sciangula D, D'Alessio R. 1998. Equilibrium and kinetics of rotamer interconversion in immunosuppressant prodigiosin derivatives in solution. J Pharm Sci 88(1):73-8.

Rossello-Mora R and Amann R. 2001. The species concept for prokaryotes. FEMS Microbiol Rev 25(1):39-67.

Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. 2000. Artemis: Sequence visualization and annotation. Bioinformatics 16(10):944-5.

Salem SM, Kancharla P, Florova G, Gupta S, Lu W, Reynolds KA. 2014. Elucidation of final steps of the marineosins biosynthetic pathway through identification and characterization of the corresponding gene cluster. J Am Chem Soc 136(12):4565-74.

Sawabe T, Ogura Y, Matsumura Y, Feng G, Amin AR, Mino S, Nakagawa S, Sawabe T, Kumar R, Fukui Y. 2013. Updating the *Vibrio* clades defined by multilocus sequence phylogeny: Proposal of eight new clades, and the description of *Vibrio tritonius* sp. nov. Vibrio Ecology, Pathogenesis and Evolution :7.

Shieh WY, Chen Y, Chaw S, Chiu H. 2003. *Vibrio ruber* sp. nov., a red, facultatively anaerobic, marine bacterium isolated from sea water. Int J Syst Evol Microbiol 53(2):479-84.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J. 2011. Fast, scalable generation of high‐quality protein multiple sequence alignments using clustal omega. Molecular Systems Biology 7(1).

Singh P and Shekhawat N. 2012. Chemometric descriptors in the rationale of antimalarial activity of natural and synthetic prodiginines. J.Curr.Chem.Pharm.Sci 40:244-60.

Thompson CC, Amaral GR, Campeão M, Edwards RA, Polz MF, Dutilh BE, Ussery DW, Sawabe T, Swings J, Thompson FL. 2015. Microbial taxonomy in the post-genomic era: Rebuilding from scratch? Arch Microbiol 197(3):359-70.

Thompson FL, Gevers D, Thompson CC, Dawyndt P, Naser S, Hoste B, Munn CB, Swings J. 2005. Phylogeny and molecular identification of *Vibrios* on the basis of multilocus sequence analysis. Appl Environ Microbiol 71(9):5107-15.

Vandamme P, Pot B, Gillis M, de Vos P, Kersters K, Swings J. 1996. Polyphasic taxonomy, a consensus approach to bacterial systematics. Microbiol Rev 60(2):407-38.

Williams RP. 1973. Biosynthesis of prodigiosin, a secondary metabolite of *Serratia marcescens*. Appl Microbiol 25(3):396-402.

Williamson NR, Fineran PC, Leeper FJ, Salmond GP. 2005. The biosynthesis and regulation of bacterial prodiginines. Nature Reviews Microbiology 4(12):887-99.

Williamson NR, Fineran PC, Gristwood T, Chawrai SR, Leeper FJ, Salmond GP. 2007. Anticancer and immunosuppressive properties of bacterial prodiginines. .

Xie BB, Shu YL, Qin QL, Rong JC, Zhang XY, Chen XL, Zhou BC, Zhang YZ. 2012. Genome sequence of the cycloprodigiosin-producing bacterial strain *Pseudoalteromonas rubra* ATCC 29570(T). J Bacteriol 194(6):1637-8.

Xiong Z, Wang J, Hao Y, Wang Y. 2013. Recent advances in the discovery and development of marine microbial natural products. Marine Drugs 11(3):700-17.

Zeigler DR. 2003. Gene sequences useful for predicting relatedness of whole genomes in bacteria. Int J Syst Evol Microbiol 53(6):1893-900.