1-1-2017

Genomics and Molecular Approaches to Delineate Pathogenesis of Aeromonas Hydrophila, Aeromonas Veronii, and Edwardsiella Piscicida Infections in Fish

Hasan Cihad Tekedar

Genomics and molecular approaches to delineate pathogenesis of *Aeromonas hydrophila*, *Aeromonas veronii*, and *Edwardsiella piscicida* infections in fish

By

Hasan Cihad Tekedar

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Veterinary Medical Sciences
in the College of Veterinary Medicine

Mississippi State, Mississippi

December 2017

Genomics and molecular approaches to delineate pathogenesis of *Aeromonas hydrophila*, *Aeromonas veronii*, and *Edwardsiella piscicida* infections in fish

By

Hasan Cihad Tekedar

Approved:

_____

Mark L. Lawrence
(Major Professor)

_____

Attila Karsi
(Co-Major Professor)

_____

Daniel G. Peterson
(Committee Member)

_____

Andy D. Perkins
(Committee Member)

_____

Larry A. Hanson
(Graduate Coordinator)

_____

Kent H. Hoblet
Dean
College of Veterinary Medicine

Name: Hasan Cihad Tekedar

Date of Degree: December 8, 2017

Institution: Mississippi State University

Major Field: Veterinary Medical Sciences

Major Professor: Dr. Mark L. Lawrence

Title of Study: Genomics and molecular approaches to delineate pathogenesis of *Aeromonas hydrophila*, *Aeromonas veronii*, and *Edwardsiella piscicida* infections in fish

Pages in Study 208

Candidate for Degree of Doctor of Philosophy

The U.S. aquaculture industry has become well established in the last three decades, and channel catfish aquaculture is the most significant component of this industry. Virulent *Aeromonas hydrophila* has been a serious disease problem since 2009 in the U.S. catfish aquaculture, and *Aeromonas veronii* and *Edwardsiella piscicida* are emerging pathogens of catfish. Therefore, this study aims to address fundamental questions on virulence mechanisms of these three fish pathogens, which I expect to support the development of control measures for preventing these diseases.

In this study, *E. piscicida* and virulent *Aeromonas hydrophila* (vAh) genomes were sequenced, and comparative analyses were conducted using the genome sequences. Average nucleotide identity (ANI) calculations showed that *E. piscicida* strains share high sequence identity, yet they are from diverse host species and geographic regions. vAh isolates share very high sequence identity, while the other *A. hydrophila* genomes are more distantly related to this clonal group. We applied several comparative genomics approaches to evaluate *E. piscicida* genomes and *E. ictaluri* genomes, providing valuable

information about unique and shared features of these two important pathogens in the *Edwardsiella* genus.

Comprehensive secretion system analysis of 55 *A. hydrophila* genomes and deletion of *tssD* and *tssI* core elements of T6SS from vAh isolate ML09-119 has provided new knowledge. We sequenced the genome of virulent *Aeromonas veronii* strain ML09-123 from catfish indicated that it was highly similar to an *A. veronii* strain from China. Evaluation of all 41 *A. veronii* genomes available in the National Center for Biotechnology Information (NCBI) provides a base platform to investigate in detail the molecular mechanism of *A. veronii* biology and virulence.

Lastly, we constructed deletion mutants *vAhΔsia*, *vAhΔent*, *vAhΔcol*, *vAhΔhfq1*, *vAhΔhfq2*, and *vAhΔhfq1Δhfq2* to determine roles of *A. hydrophila* secreted proteins and regulatory proteins on virulence in catfish. Results showed that sialidase (*vAhΔsia*) and enterotoxin (*vAhΔent*) mutants were significantly attenuated.

DEDICATION

Dedicated to my family and all the scientists, who have been zealously studying to advance our understanding of science...

ACKNOWLEDGEMENTS

I thank my advisor Dr. Mark L. Lawrence for the opportunity to work in his laboratory and I am very grateful for the insight and guidance he provided throughout my studies.

I am privileged to have worked with my co-advisor, Dr. Attila Karsi, and grateful for the guidance and encouragement during my years as a graduate student.

I also thank to my Ph.D. committee members, Drs. Daniel G. Peterson, and Andy D. Perkins for their constructive comments on my research project. Their knowledge and expertise helped to expand my scientific knowledge in different aspects of the science.

I would like to thank my lab-mates and friends for their friendship and support.

Lastly, my most profound gratitude goes to my parents and sisters. Their continuous support and encouragement allowed me to complete this work.

TABLE OF CONTENTS

DEDICATION ii

ACKNOWLEDGEMENTS iii

LIST OF TABLES ix

LIST OF FIGURES x

CHAPTER

I. GENERAL INTRODUCTION AND RELATED REVIEW OF LITERATURE 1

1.1 Aquaculture 1
1.2 Some of the diseases affecting channel catfish industry 3
1.2.1 *Edwardsiella piscicida* 3
1.2.2 *Aeromonas hydrophila* 6
1.2.3 *Aeromonas veronii* 9
1.3 Some of the virulence factors in Gram-negative bacteria 11
1.3.1 Sialidase, Collagenase, Enterotoxin, and Hfq genes 11
1.3.2 Secretion systems in Gram-negative bacteria 14
1.3.3 Mobilome elements 16
1.4 Genomics 20
1.4.2 Genome sequencing 22
1.4.3 Comparative genomics 24
1.5 Significance of this study and objectives 25
1.5.1 Specific aim #1. Genome sequencing of *E. piscicida* and *A. hydrophila* isolates 26
1.5.2 Specific aim #2. Comparative genome analysis of *E. piscicida* and *E. ictaluri* genomes 27
1.5.3 Specific aim #3. Comparative analysis of secretion systems of virulent *Aeromonas hydrophila* strains and mutation of *tssD* and *tssI* genes from T6SS. 27
1.5.4 Specific aim #4. Comparative genomics insights into pathogenicity of *A. veronii* 27

1.5.5 Specific aim #5. Mutate selected *A. hydrophila* genes encoding secreted proteins (sialidase and collagenase), enterotoxin, and master regulator (Hfq) to determine their role in *A. hydrophila* virulence. 27
1.6 References 29

II. COMPLETE AND DRAFT GENOME SEQUENCES OF *EDWARDSIELLA PISCICIDA* AND *AEROMONAS HYDROPHILA* ISOLATES RECOVERED FROM OUTBREAKS....46

2.1 *Edwardsiella piscicida* 46
2.2 *Aeromonas hydrophila* 47
2.3 Material and methods 48
2.3.1 Sequencing 48
2.3.1.1 *Edwardsiella piscicida* C07-087 48
2.3.1.2 *Aeromonas hydrophila* ML09-119 49
2.3.1.3 *Aeromonas hydrophila* AL06-06....49
2.3.1.4 *Aeromonas hydrophila* TN97-08....50
2.3.1.5 *Aeromonas hydrophila* AL10-121, AL09-79, ML09-121, ML09-122....50
2.3.1.6 *Aeromonas hydrophila* Arkansas 2010, AL97-91, MN98-04 50
2.3.2 Plasmid sequencing 51
2.3.3 Genome annotation and relatedness 51
2.4 Results 53
2.4.1 *Edwardsiella piscicida* 53
2.4.2 *Aeromonas hydrophila* 53
2.5 Discussion....57
2.6 References 63

III. COMPLETE GENOME SEQUENCING AND COMPARATIVE GENOMICS OF THE FISH PATHOGENS *EDWARDSIELLA ICTALURI* 93-146 AND *EDWARDSIELLA PISCICIDA* C07-087 68

3.1 Abstract....68
3.2 Introduction 69
3.3 Material and Methods....71
3.3.1 Bacterial strains and *Edwardsiella* genomes....71
3.3.2 Sequencing, assembly, and annotation....73
3.3.3 ANI calculation 73
3.3.4 Phylogenetic tree creation 73
3.3.5 Pan-Core genome analysis 73
3.3.6 Singleton analysis....74
3.3.7 Orthologous gene clusters analysis 74
3.3.8 Genome structure variation 75
3.3.9 Protein Secretion Systems 75

3.3.10 Insertion elements....76
3.3.11 Genomic Islands (GI)....76
3.3.12 Prophages....76
3.3.13 Integron identification....77
3.3.14 CRISPR and Cas elements analysis....77
3.3.15 Virulence factors....78
3.4 Results....78
3.4.1 Genome Features....78
3.4.2 ANI and phylogenetic analysis....78
3.4.3 Pan-Core genome analysis....80
3.4.4 Singleton analysis....81
3.4.5 Orthology analysis....81
3.4.6 Gene variation in the thirteen *E. ictaluri* and *piscicida* genomes....83
3.4.7 Secretion Systems....84
3.4.8 Insertion sequences....87
3.4.9 Genomic islands....88
3.4.10 Phage sequences....89
3.4.11 Integron identification....89
3.4.12 CRISPR (Clustered regularly interspaced short palindromic repeats) and Cas elements....90
3.5 Discussion....90
3.6 References....98

IV. COMPARATIVE ANALYSIS OF SECRETION SYSTEMS OF VIRULENT *AEROMONAS HYDROPHILA* STRAINS AND MUTATION OF *tssD* AND *tssI* GENES FROM T6SS....107

4.1 Abstract....107
4.2 Introduction....108
4.3 Materials and methods....110
4.3.1 Genome sequencing, genomes from sequence database and annotation....110
4.3.2 Phylogenetic tree creation....112
4.3.3 ANI and AAI calculation....113
4.3.4 Secretion systems identification....113
4.3.5 Virulence database....113
4.3.6 Bacterial strains and plasmids....114
4.3.7 In-frame deletion of *A. hydrophila* genes....115
4.3.8 Virulence and efficacy of the vAh mutants in catfish fingerlings....117
4.3.9 Statistical analysis....118
4.3.10 Host-pathogen interaction network....118
4.4 Results....118
4.4.1 Genome features....118
4.4.2 ANI and phylogenetic tree creation....119

4.4.3 Secretion systems in *Aeromonas hydrophila* genomes 121
4.4.4 Virulence database comparison 124
4.4.5 Construction and virulence of T6SS effector mutant strains.....127
4.4.6 Host-pathogen interaction 129
4.5 Discussion.....129
4.6 References 137

V. COMPARATIVE GENOME ANALYSIS OF THE *AEROMONAS VERONII* STRAIN ML09-123 PROVIDES INSIGHTS INTO ITS PATHOGENICITY OF ORIGIN 147

5.1 Abstract.....147
5.2 Introduction 148
5.3 Material and methods 150
5.3.1 Bacterial strains and data source for comparative genome analysis 150
5.3.2 Sequencing, assembly, and annotation 151
5.3.3 Phylogenetic tree and ANI calculation.....152
5.3.4 Pan-Core genome 152
5.3.5 Subsystems coverage.....152
5.3.6 Secretion systems 153
5.3.7 Insertion elements.....153
5.3.8 Phage elements 153
5.3.9 Integron identification 154
5.3.10 CRISPR (Clustered regularly interspaced short palindromic repeats) and Cas (CRISPR associated genes) elements analysis 154
5.3.11 Antibiotic resistance genes 154
5.3.12 Fish challenge.....154
5.4 Results 155
5.4.1 Genome features.....155
5.4.2 Phylogenetic tree and ANI calculation.....155
5.4.3 Pan-Core genome analyses.....157
5.4.4 Subsystems coverage.....158
5.4.5 Secretion systems, Flagellum, TAD, T4P 160
5.4.6 Insertion elements.....161
5.4.7 Phage elements 162
5.4.8 Integron elements 162
5.4.9 CRISPR-Cas elements.....163
5.4.10 Antibiotic resistance genes 164
5.4.11 Fish challenge.....165
5.5 Discussion.....166
5.6 References 176

VI. CONSTRUCTION AND EVALUATION OF SIX *AEROMONAS HYDROPHILA* IN-FRAME DELETION MUTANTS.....182

6.1 Abstract....182
6.2 Introduction....183
6.3 Materials and methods....185
6.3.1 Bacterial strains and plasmids....185
6.3.2 Genomic DNA and plasmid DNA isolation....186
6.3.3 Preparation of competent cells....187
6.3.4 Selection of potential genes in *A. hydrophila*....187
6.3.5 In-frame gene deletion in vitro....188
6.3.6 Construction of double mutant....192
6.3.7 Growth kinetics....193
6.3.8 Virulence assays....193
6.3.9 Statistical Analysis....194
6.4 Results....194
6.4.1 Construction of the *A. hydrophila* in-frame deletion mutants....194
6.4.2 Growth kinetics....195
6.4.3 Virulence and efficacy of *A. hydrophila* ML09-119 in-frame deletion mutants....197
6.5 Discussion....197
6.6 References....201

VII. CONCLUSION....205

7.1 Summary of results....205

LIST OF TABLES

2.1 Features of *Edwardsiella piscicida* genomes used for analysis. 52

2.2 Features of *Aeromonas hydrophila* genomes used in this analysis 54

3.1 *Edwardsiella* genomes used in the comparative genomic analyses. 71

3.2 Type Four Secretion System genes 86

3.3 Insertion elements distribution in the *E. ictaluri* and *E. piscicida* genomes. 88

4.1 *A. hydrophila* genomes used in comparative genomic analyses. 111

4.2 Bacterial strains and plasmids used in the present study. 115

4.3 Primers used to generate and verify in-frame deletion of vAh genes. 116

4.4 Identified secretion systems elements and related components searched against MvirDB 124

4.5 The sizes of upstream (USF), downstream (DSF), and in-frame fused fragments (FF), deleted region (DR), and undeleted region (UD) by base pair (bp). 127

5.1 *A. veronii* genomes used in comparative genomic analyses. 150

5.2 General features of CRISPR-Cas loci in 41 *A. veronii* genomes. 164

6.1 Bacterial strains and plasmid used in this study. 186

6.2 Primers used to generate and verify in-frame deletion of vAh genes. 189

6.3 The sizes of upstream (USF), downstream (DSF), and in-frame fused fragments (FF), deleted region (DR), and undeleted region (UD) by base pair (bp). 195

LIST OF FIGURES

1.1 World capture fisheries and aquaculture production (FAO, 2016). 2

1.2 Common clinical motile *Aeromonas* species and their differentiation (Woo et al., 2011). 7

1.3 Schematic representation of the structure of the secretion systems in Gram-negative bacteria. 14

1.4 Representation of an integron element identified by IntegronFinder. 17

1.5 General features of genomic islands. 18

1.6 Major contributions to phage research in fish and shellfish research (Gon Choudhury et al., 2017). 19

1.7 DNA Sequencing costs reduction over the years (NHGRI, 2017). 21

1.8 Development of bacterial genome sequencing (Loman et al., 2015). 23

1.9 Potential effects caused by bacterial genome dynamics (Pallen et al., 2007). 24

2.1 Average nucleotide identities (ANI) for *E. piscicida* genomes (analysis conducted on 9/13/2017). 58

2.2 Sequence identity of *E. piscicida* genomes using *E. piscicida* C07-87 as the reference. 59

2.3 Average nucleotide identities (ANI) for *Aeromonas hydrophila* genomes (analysis conducted on 9/13/2017). 60

2.4 Sequence identity of *A. hydrophila* genomes using *A. hydrophila* ML09-119 as the reference. 61

3.1 Average nucleotide identities (ANI) of *Edwardsiella* genomes. 79

3.2 Phylogenetic tree of *Edwardsiella* genomes based on core genomes. 80

3.3 Pan vs. Core Plot analysis of *Edwardsiella* genus members. 81

3.4 Putative functions of orthologs and unique clusters for the *E. piscicida* and *E. ictaluri* genomes 83

3.5 A- T6SS Type-I subtype I4B and I2 in the *E. ictaluri* and *E. piscicida* genomes B-T4SS in *E. ictaluri* 93-146. 85

3.6 Genomic islands in *Edwardsiella* genomes.....89

4.1 Average nucleotide identities (ANI) of *A. hydrophila* genomes and phylogenetic tree based on core genome.....120

4.2 Core and accessory proteins of secretion systems, T4P, Tad, and flagella encoded in *A. hydrophila* genomes. 123

4.3 Virulence of vAh T6SS mutants in channel catfish fingerlings. 128

5.1 Average nucleotide identity (ANI) values of *A. veronii* genomes and phylogenetic tree analysis based on the core genomes. 157

5.2 Pan vs. Core genome development plot for *A. veronii* genomes.....158

5.3 SEED subsystem category for *A. veronii* genomes. 159

5.4 Core components of secretion systems elements, T4P, Tad and flagella availability in *A. veronii* genomes. 161

5.5 Percent mortalities in catfish challenged with *A. hydrophila* ML09-119 and *A. veronii* ML09-123. 166

6.1 *A. hydrophila* in-frame deletion process. 190

6.2 Growth kinetics in wild-type *A. hydrophila* ML09-119 and mutants. 196

6.3 Percent mortalities of channel catfish fingerlings challenged with the 5 single mutants (*vAhΔsia*, *vAhΔcol*, *vAhΔent*, *vAhΔhfq*, *vAhΔhfq2*) and 1 double mutant (*vAhΔhfq1Δhfq2*) in-frame deletion mutants and wild type. 197

CHAPTER I

GENERAL INTRODUCTION AND RELATED REVIEW OF LITERATURE

1.1 Aquaculture

The human population is rising exponentially, reaching 7.3 billion in 2017. According to the United Nations, it is expected to reach 8.5 billion by 2030. This growth leads to several major challenges for developed and developing countries, particularly in food security. To overcome this problem, humankind has been developing different methods to increase the yield of aquatic products for food consumption. Aquaculture is a manmade system for controlling environmental conditions to raise fish for food, and it has the potential to provide a vital meat protein source.

Aquaculture began approximately 4000 years ago, Fan li mentioned in his writings that Asian emperors maintained their favorite fish in ponds (Overturf, 2009). It is one of the fastest growing agricultural sectors providing almost half of the animal food for human consumption. Therefore, it could be the best solution for food production (Bostock et al., 2010; FAO, 2017a). Currently, 567 aquatic species are farmed by individual producers and multinational companies (FAO, 2017b). According to a FAO report, worldwide captured fish production has remained relatively static since the early

1990s, while worldwide aquaculture production is steadily increasing (Figure 1.1) (FAO, 2016).

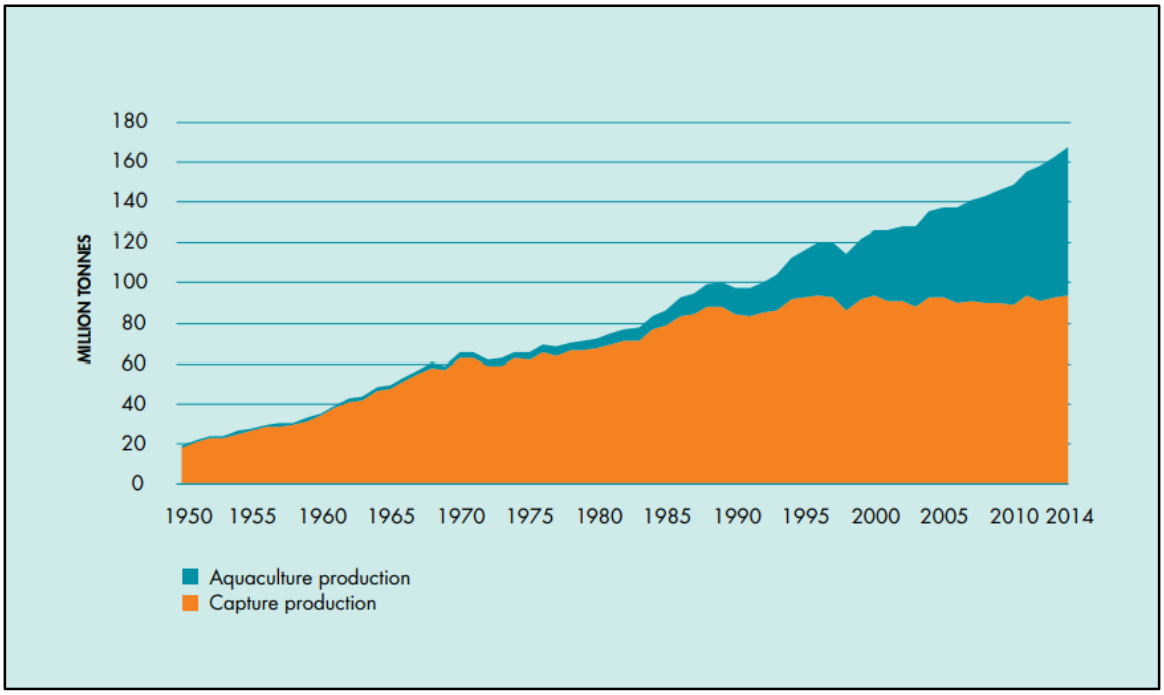


Figure 1.1 World capture fisheries and aquaculture production (FAO, 2016).

In the United States of America, the aquaculture industry has become well established in the last 35 years, and the majority of its production is in catfish (FAO, 2011). Alabama, Arkansas, and Mississippi are the major catfish producing states, using 54.2 thousand acres of surface water for catfish production in July 1 through December 31, 2017 (NASS, 2017). Sales from catfish aquaculture totaled $341 million in 2013 (USDA, 2013). However, the catfish industry suffers heavy losses due to several different diseases caused by bacteria, fungi, and parasites. A multistate group of scientists and

agencies have collaborated to address emerging diseases and provide cooperative efforts to control diseases, but there is still need for effective control methods.

1.2 Some of the diseases affecting channel catfish industry

Some of the important diseases affecting the catfish industry are discussed in detail below.

1.2.1 *Edwardsiella piscicida*

Taxonomy

Domain: Bacteria

Phylum: Proteobacteria

Class: Gammaproteobacteria

Order: Enterobacteriales

Family: Enterobacteriaceae

Genus: *Edwardsiella*

Species: *E. piscicida*

Edwardsiella piscicida is a Gram-negative facultative anaerobe causing disease in animals. It is well-known as an etiologic agent of acute to chronic edwardsiellosis in different aquatic species (Evans et al., 2011). It is motile by peritrichous flagella. The taxonomy of this species was recently clarified. Initially, in 1965, *Edwardsiella* was discovered as a new species under the family *Enterobacteriaceae* (Ewing et al., 1965). *Edwardsiella tarda* was described in 1973 (Meyer, 1973), *Edwardsiella hoshinae* in 1980 (Grimont et al., 1980), and *Edwardsiella ictaluri* in 1981 (Hawke et al., 1981). In 2013, *Edwardsiella piscicida* was reported for the first time as a new species (Abayneh et al.,

2013). Previously, *E. piscicida* strains were classified as *E. tarda*. After two years, *Edwardsiella anguillarum* (Shao et al., 2015) was defined as a separate species from *E. piscicida*. Currently, there are five *Edwardsiella* species described. These new taxonomic classifications have clarified the relatedness of different *Edwardsiella* strains.

E. piscicida is associated with gastrointestinal infections that may progress to hemorrhagic septicemia (Lv et al., 2013). *Edwardsiellosis* is one of the most important fish diseases that negatively affects aquaculture industries around the world (Park et al., 2012). *E. piscicida* infection begins with small punctate cutaneous erosions, which could progress to abscesses deep within the musculature in channel catfish (*Ictalurus punctatus*) (Evans JJ et al., 2011).

To understand the molecular mechanisms of this disease, several different *E. piscicida* and *E. piscicida*-like isolates from different host species and locations have been sequenced, including strains from grouper (Reichley et al., 2015), tilapia (Reichley, et al., 2015), channel catfish (Reichley et al., 2016), turbot (Buján et al., 2017), and olive flounder (Oguro et al., 2014). Because some of these were sequenced before the description of *E. piscicida* species, they were reported as *E. tarda*. Examples include catfish strain C07-087 (Tekedar et al., 2013), striped bass strain FL6-60 (van Soest et al., 2011), European eel isolate strain ET883 (Abayneh et al., 2012) and turbot isolate strain EIB202 (Wang et al., 2009). Recent efforts have been made to clarify confusion on the classification of *Edwardsiella* strains, and taxonomically mislabeled genomes have been updated or revised to correct taxonomical misclassification.

Repetitive sequence-mediated PCR for *E. piscicida* showed that *E. piscicida* strains are more prevalent than *E. tarda* in southeastern USA (Griffin et al., 2014). Real-

time quantitative polymerase chain reaction (qPCR) assays were developed to detect and quantify the *Edwardsiella* species in catfish pond water and tissues (Reichley et al., 2015). *E. piscicida* infection in largemouth bass causes multiorgan necrosis and granulomas (Fogelson et al., 2016).

Two-component systems are important components in virulence of pathogenic bacteria. For example, two component system EsrA-EsrB in *Edwardsiella tarda* has a role in controlling the type III secretion system (T3SS) (Zheng et al., 2005). EsrA-EsrB contribute to virulence regulation in *E. tarda* along with PhoP-PhoQ (Lv et al., 2012). In *E. piscicida*, EsrB is responsible for global regulation of physiological adaptation and activation of T3SS and its effector components during infection in turbot (Y. Liu et al., 2017). EsrB also has a role in the regulation of reactive oxygen species (ROS) (Yin et al., 2017).

Bacterial secretion systems have important roles in transporting proteins from cytoplasm to periplasm, interaction with other organism, promoting bacterial virulence, enhancing attachment, scavenging resources from the environment, and disrupting target cells (Green et al., 2016). In *E. piscicida*, EseH is a T3SS effector protein that targets MAPK signaling during infection (Hou et al., 2017). Transposon insertion site sequencing in *E. piscicida* strain EIB202 revealed that its plasmid (pEIB202) encodes a T4SS, and this system is a major genetic mechanism for transferring multi-drug resistance genes (Y. Liu et al., 2017).

Interactome analysis of *E. piscicida* and grouper (*Epinephelus drummondhayi*) liver proteins showed that bacterial proteins regulate the expression of host innate immune-related proteins (H. Li et al., 2017). These findings clearly indicate the

importance of secretome studies, which could contribute to design new vaccine candidates. In *E. piscicida* strain CK216, a mutant strain defective in cyclic AMP receptor protein (Δ*crp*) was proposed as a live attenuated vaccine candidate (Choe et al., 2017).

Due to the importance of *E. piscicida* on the aquaculture industry, research is being conducted to elucidate molecular pathogenesis of this disease. Comparative genomics of this pathogen against all the other *Edwardsiella* species has not been described before and will be discussed in this dissertation.

1.2.2 *Aeromonas hydrophila*

Taxonomy

Domain: Bacteria

Phylum: Proteobacteria

Class: Gammaproteobacteria

Order: Aeromonadales

Family: Aeromonadaceae

Genus: *Aeromonas*

Species: *A. hydrophila*

Aeromonas hydrophila is a Gram-negative motile, polar flagellated, short, rod shaped, mesophilic species that is ubiquitous in aquatic environments and capable of causing severe infection in several host species including humans, fish, invertebrates, amphibians, and reptiles (John et al., 2011; Tekedar et al., 2013). In the genus *Aeromonas*, the motile aeromonads are loosely associated species that have similar serology, biochemistry, and genetics (Cipriano, 2001). Common clinical motile

Aeromonas species and their differentiation are listed in Figure 1.2 (Woo et al., 2011). According to this list, *A. hydrophila* is incapable of producing ornithine decarboxylase.

Characteristic[a]	*A. hydrophila*	*A. veronii* bv. *sobria*	*A. veronii* bv. *veronii*	*A. caviae*	*A. schubertii*	*A. jandaei*	*A. trota*
Aesculin hydrolysis	+	–	+	+	–	–	–
Voges–Proskauer reaction	+	+	+	–	V	+	–
Pyrazinamidase activity	+	–	–	+	–	–	–
CAMP-like factor (aerobic only)	+	+	+	–	–	V	–
Arabinose fermentation	V	–	–	+	–	–	–
Mannitol fermentation	+	+	+	+	–	+	+
Sucrose fermentation	+	+	+	+	–	–	–
Ampicillin susceptibility	R	R	R	R	R	R	S
Carbenicillin susceptibility	R	R	R	R	R	R	S
Cephalothin susceptibility	R	S	S	R	S	R	R
Colistin susceptibility[b]	V	S	S	S	S	R	S
Lysine decarboxylase	+	+	+	–	+	+	+
Ornithine decarboxylase	–	–	+	–	–	–	–
Arbutin hydrolysis	+	–	+	+	–	–	V
Indole production	+	+		+	–	+	+
H_2S production	+	+	+	–	–	+	+
Gas from glucose	+	+	+	–	–	+	+
Haemolysis (TSA with 5% sheep erythrocytes)	+	+	+	V	+	+	V

[a]+, positive for >70% of isolates; –, negative, i.e. positive for <30% of isolates; V, variable; R, resistant; S, susceptible.
[b]MIC (single dilution), 4 µg/ml.

Figure 1.2 Common clinical motile *Aeromonas* species and their differentiation (Woo et al., 2011).

The disease caused by *A. hydrophila* and other motile aeromonads is referred to as Motile Aeromonad Septicemia (MAS), which is characterized by destructive systemic infection (Grizzle et al., 1993). Typical signs of MAS include red sores, necrosis,

ulceration, swelling of tissues, and hemorrhagic septicemia (Karunasagar et al., 1989). In addition to their capability of causing fish disease, aeromonads can be responsible for opportunistic disease in humans (Janda et al., 1998).

Feeding infected fish with antibiotic-medicated feed is a common approach for controlling *A. hydrophila* infection, but often this is not effective (DePaola et al., 1995). Currently, there are two FDA approved feed additive, oxytetracycline (Terramycin®) and florfenicol (Aquaflor®) available in the USA for motile aeromonad septicemia (Plumb et al., 1995). Use of antibiotic-medicated feed has the disadvantage of potential for developing antimicrobial resistance. Therefore, antibiotic treatment is not sufficient as a long-term strategy to control this disease. Another option is vaccination, but currently, there are no commercially available vaccines to protect farm-raised catfish against MAS. Passage with rifampicin and/or novobiocin was used to attenuate virulent *A. hydrophila* for potential development of a live vaccine to protect channel catfish (Pridgeon et al., 2011). Probiotics have also been suggested for disease control in aquaculture because this approach can reduce pathogen attachment to fish intestinal mucus (Vine et al., 2004).

As of September 2017, 62 *Aeromonas hydrophila* genomes from different sources have been sequenced, 14 of which are completely closed. The rest are in draft version (NCBI, 2017). Our group contributed to this numbers by providing two complete genomes (Tekedar et al., 2015; Tekedar et al., 2013) and eight draft genomes (Tekedar et al., 2017; Tekedar et al., 2016a, 2016b).

Since 2009, *A. hydrophila* epidemics have been affecting catfish producers in the U.S. The catfish industry has lost more than 5.5 million pounds of market-size fish. By sequencing a panel of *A. hydrophila* isolates, it was determined that the outbreaks were

caused by an emergent clonal group of *A. hydrophila* isolates with evidence of lateral genes transfer that contributed to their virulent phenotype (Hossain et al., 2013). This clonal group was named virulent *A. hydrophila* (vAh). Further investigation revealed that vAh from channel catfish in the U.S. is highly similar to an Asian clonal virulent group of *A. hydrophila*, suggesting a possible Asian origin for vAh in the U.S. through imported carp (Hossain et al., 2014). Genomic comparison of the Asian and U.S. vAh confirmed a large number of conserved genomic islands with a few notable differences (Pang et al., 2015). Our group recently reported histopathological and ultrastructural changes in channel catfish during vAh infection (Abdelhamed et al., 2017).

Recombinant vAh outer membrane proteins and fimbrial proteins show potential as protective vaccine antigens (Abdelhamed et al., 2017; Abdelhamed et al., 2016; Dash et al., 2014; Yadav et al., 2014). However, effective control of vAh epidemics has still not been successfully achieved due to lack of essential information. Therefore, there is still an urgent need to understand the pathogenic mechanisms of vAh.

1.2.3 *Aeromonas veronii*

Taxonomy

Domain: Bacteria

Phylum: Proteobacteria

Class: Gammaproteobacteria

Order: Aeromonadales

Family: Aeromonadaceae

Genus: *Aeromonas*

Species: *A. veronii*

Aeromonas veronii is a Gram-negative, rod-shaped, mesophilic, motile species that also causes motile aeromonad septicemia (MAS) (Janda et al., 2010). It can affect several different species but is mostly encountered in immunocompromised hosts with skin lesions causing wound infection with soft tissue destruction. Lesions can progress to ulcers and hemorrhagic septicemia (Austin et al., 2012; McGarey et al., 1991). *A. veronii* infection has been reported from different ornamental and aquatic animals including tilapia (Hassan et al., 2017), rainbow trout (Zepeda-Velazquez et al., 2015), sea bass (Uzun et al., 2015), channel catfish (D. Liu et al., 2016), and oscar cichlid (Sreedharan et al., 2011). Compared to *A. hydrophila*, *A. veronii* encode ornithine decarboxylase as a unique characteristic, whereas they lack pyrazinamidase activity and arabinose fermentation capabilities (Figure 1.2).

A. veronii is not typically considered a primary pathogen for aquaculture species. However, it was recently isolated from outbreaks of fish disease in different regions including China (D. Liu et al., 2016) and Saudi Arabia (Hassan et al., 2017). *A. veronii* is also the causative agent of the ulcerative syndrome in catfish (Cai et al., 2012).

Metatranscriptomics (Bomar et al., 2011) and metagenomics (Maltz et al., 2014) studies indicated that *A. veronii* is present in the leech gut as the second most prevalent symbiotic bacteria. *A. veronii* strains were also isolated with other *Aeromonas* species such as *A. sobria* from turtle (soft shell disease) (C. Chen et al., 2013) and *A. jandai* from Nile tilapia (Dong et al., 2017). Although the role of *A veronii* in each of these systems is not defined, it is interesting to note that it is often isolated with other bacteria either in a pathogenic or symbiotic relationship with a host.

Aeromonas species demonstrate different antimicrobial susceptibilities. For example, in one case *Aeromonas caviae* was more susceptible to ticarcillin than other evaluated *A. veronii* and *A. hydrophila* species (Igbinosa et al., 2012). Another study evaluated 47 *A. veronii* isolates from two different hosts, *Anguilla japonica* and *Cyprinus carpio*, and found that they tend to carry insertion elements and antibiotic resistance genes (Chung et al., 2017).

1.3 Some of the virulence factors in Gram-negative bacteria

1.3.1 Sialidase, Collagenase, Enterotoxin, and Hfq genes

To colonize and spread in host tissues, bacteria often produce extracellular enzymes that degrade tissues and proteins to serve as an energy source and to facilitate spreading and dissemination (Duarte et al., 2016; Janda et al., 2010). *A. hydrophila* is capable of producing a large number of extracellular enzymes and toxins that contribute to host inflammation and promote spreading in the tissues. Major extracellular proteins (ECPs) of *A. hydrophila* have been characterized (Yu et al., 2007); some are considered to be essential virulence factors and play an important role in the invasion of host cells and tissues. In particular, protease (Wu et al., 2012), exotoxin (Grim et al., 2013; Ponnusamy et al., 2016), hemolysin, and enterotoxin (Grim et al., 2013) contribute to *Aeromonas* disease. A recent study showed that hemolysin, aerolysin, and elastase are found in virulent isolates; also, hemolysin, protease, and nuclease activity were higher in virulent *A. hydrophila* (vAh) isolates (Pridgeon et al., 2013).

Vaccination using extracellular products of *A. hydrophila* provided protection against infection (Zhang et al., 2014). Recombinant aerolysin and hemolysin show potential as general immunogens for future development of recombinant vaccines (Zhang

et al., 2015). Several *A. hydrophila* and *Aeromonas* extracellular proteins and toxins have been studied for detection of the pathogen as well as for phenotypic and molecular characterization (Furmanek-Blaszk, 2014).

In Chapter VI of this dissertation, the role of sialidase as a virulence factor for *A. hydrophila* was investigated. Sialic acid-based pattern recognition contributes to microbial virulence, which strongly supports sialidase inhibition for dampening inflammation caused by infection (G. Y. Chen et al., 2011). Sialidase also plays an important role in the pathogenesis of *E. tarda* and *Porphyromonas gingivalis*, and it has potential as a vaccine candidate for *E. tarda* (Aruni et al., 2011; Chigwechokha et al., 2015; Jin et al., 2012). Moreover, sialidase of *Streptococcus intermedius* was identified as a virulence factor that modifies sugar chains (Takao et al., 2010). Sialidase is responsible for reducing biofilm formation and capsule biosynthesis of *Porphyromonas gingivalis* (C. Li et al., 2012). Sialidase can also modulate the host immune response (Carlson et al., 2010; Kuroiwa et al., 2009).

Bacterial collagenases are known for mediating degradation of extracellular matrices in the host due to their capability to digest native collagen (Duarte et al., 2016). Collagenolytic activity in *Aeromonas* strains is an example (Duarte et al., 2015; Han et al., 2008). In *Aeromonas veronii*, collagenase activity is not a main virulence factor, but it has a role in initial penetration (Han et al., 2008). On the other hand, in *Leptospira* collagenase is a critical virulence factor for transmission and invasiveness (Kassegne et al., 2014).

A. hydrophila strains produce a large number of exotoxins. Deletion of enterotoxin in a human isolate caused attenuation, and it was suggested that it could be

used as a potential vaccine candidate (J. Sha et al., 2002). It should be noted that many of the human *A. hydrophila* isolates have been reclassified into a separate species, *A. dhakensis*. Therefore, the majority of the enterotoxin studies reported in *A. hydrophila* were actually conducted with *A. dhakensis*. However, at least one study confirmed that T3SS and cytotoxic enterotoxin contribute to virulence of *A. hydrophila* (J. Sha et al., 2005). Enterotoxin is one of four virulence factors (enterotoxin, exotoxin A, and T3SS and T6SS effectors AexU and Hcp) found in the genomes of virulent isolates from a panel of *Aeromonas* (Grim et al., 2014). *A. hydrophila* also produces RtxA toxins that are responsible for host cell rounding and apoptotic death. Redundancy in toxin functions of *A. hydrophila* may be necessary for a successful infection (Suarez et al., 2012).

Hfq is an essential component for the activity and stability of most small RNAs (sRNAs) in bacteria. It is considered to be a master regulator of gene expression in bacteria. It has a role in virulence and fitness of several pathogens, and it contributes to the regulation of virulence factors, secretion systems, cellular processes, and iron metabolism (Feliciano et al., 2016). Hfq contributes to virulence of several bacterial pathogens organisms including *Vibrio cholerae* (Ding et al., 2004), *Listeria monocytogenes* (Christiansen et al., 2004; Nielsen et al., 2010), *Edwardsiella tarda* (Y. H. Hu et al., 2014), *Borrelia burgdorferi* (Lybecker et al., 2010), *Agrobacterium tumefaciens* (Wilms et al., 2012), *Yersinia pestis* (Geng et al., 2009), and *Brucella abortus* (Robertson et al., 1999). However, in *Legionella pneumophila* (McNealy et al., 2005) and *Neisseria gonorroeae* (Dietrich et al., 2009), deletion of *hfq* genes did not affect virulence. In *Vibrio alginolyticus*, Hfq is responsible for the oxidative envelope stress response, nutrient utilization, and colony morphology (Y. Deng et al., 2016).

1.3.2 Secretion systems in Gram-negative bacteria

Bacterial secretion systems translocate enzymes, toxins, and effector proteins that facilitate bacterial pathogenicity (Costa et al., 2015; Russell et al., 2014). Secretion systems also play an important role in defending against competitor organisms or chemical agents. In addition, they are sometimes responsible for nutrient acquisition. Our current understanding of their structure and composition is illustrated in Figure 1.3.

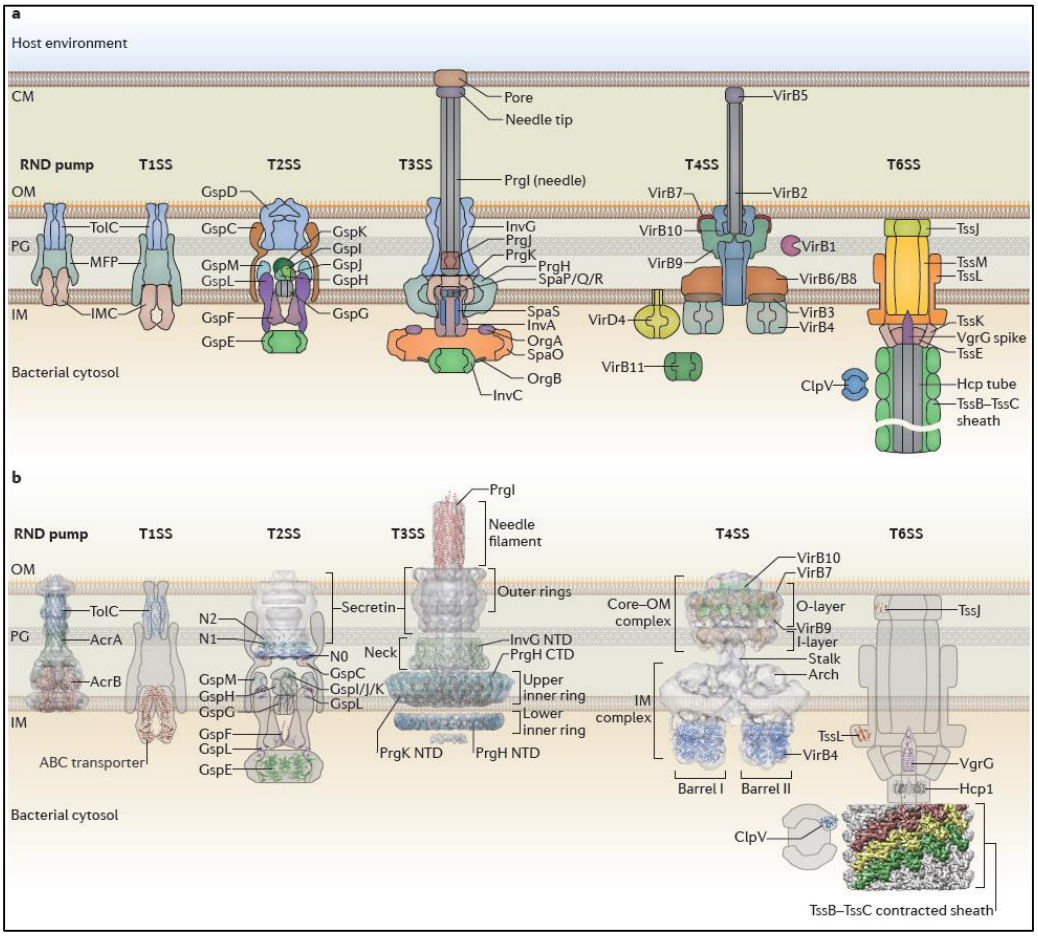


Figure 1.3 Schematic representation of the structure of the secretion systems in Gram-negative bacteria.

(a) Putative locations of each protein in the specific secretion systems are shown. (b) Structural models of each secretion system are represented (Costa et al., 2015).

Type I secretion system (T1SS) mediates one-step transportation of unfolded substrates, bypassing the periplasm and directly releasing the protein into extracellular space (Kanonenberg et al., 2013). T1SS facilitates secretion of some destructive enzymes and exotoxins with repeats-in-toxins (RTX) motifs.

Type II secretion system (T2SS) system is located in the outer membrane; it transports folded proteins from periplasm into the extracellular environment. It exports some hydrolytic enzymes, virulence factors, and toxins (including lipases, proteases, and several other proteins that process carbohydrates and substrates). (Green et al., 2016; Korotkov et al., 2012). T2SS is important in *A. hydrophila* (Ast et al., 2002; Cianciotto, 2005).

Type III secretion system (T3SS) is a protein transport nanomachine, and it is considered one of the essential virulence factors for many pathogens. T3SS secreted effectors are secreted directly from bacterial cytoplasm to host cell cytoplasm, and they modulate host cellular processes, host immune responses, and vesicle transport (Buttner, 2012; W. Deng et al., 2017). Many extracellular pathogens use T3SS to regulate the host cytoskeleton and immune response.

Type IV secretion system (T4SS) is a trans-envelope-spanning nanomachine found in many bacteria. It is responsible for transporting diverse substrates ranging from DNA to effector proteins (Darbari et al., 2015). T4SS mediates the exchange of mobile genetic elements, which facilitates adaptation to environmental changes and spread of antimicrobial resistance among bacteria (Wallden et al., 2010).

Type VI secretion system (T6SS) delivers secreted proteins directly into competitor bacteria or host cells (Zoued et al., 2014). It mediates interaction with the

environment and is a versatile defense mechanism against host and competitor organisms (Cianfanelli et al., 2016). T6SS augments bacterial adaptation to specific niches (Cianfanelli et al., 2016) and is used as a strategy to control interactions between bacterial species (Basler et al., 2013). Two T6SS proteins, in particular, Hcp and VgrG, have dual roles as effector proteins and as structural components (Pukatzki et al., 2009; Silverman et al., 2012). In *A. dhakensis* strain SSU (formerly *A. hydrophila*), deletion of these genes affected some virulence-associated features (Jian Sha et al., 2013).

Lastly, type IX secretion system (T9SS) is discovered in the Bacteroides group. It has two functions: gliding motility and protein secretion, which can contribute to virulence (Lasica et al., 2017).

1.3.3 Mobilome elements

The mobilome is defined as a set of dynamic genes located within mobile genetic elements responsible for bacterial evolution and adaptation, allowing interchangeable shuffling of genes in between close or distantly related organisms (Jorgensen et al., 2014; Piotrowska et al., 2015). For example, integrons, insertion elements, genomic islands, plasmids, phage elements, and antibiotic resistance genes are considered mobilome elements.

Integrons are well-known for facilitating spread of antibiotic resistance genes, and a recently developed tool helps identify integron elements in bacterial genomes (Cury et al., 2016). Typical integron structure is shown in Figure 1.4, consisting of an integrase gene, an *attI* recombination site, and gene cassettes (Cury et al., 2016). There are three different classes of integrons (Hall et al., 1999), and *Aeromonas* species tend to have more class 1 integrons on the oxytetracycline resistance plasmids (Schmidt, Bruun,

Dalsgaard et al., 2001). However, there have been some conflicting results on the correlation between presence of oxytetracycline resistance in *Aeromonas* and class I integrons, with one study not finding a correlation (Jacobs et al., 2007) and another finding a positive correlation between carrying an integron and *tet* genes (Moura et al., 2012).

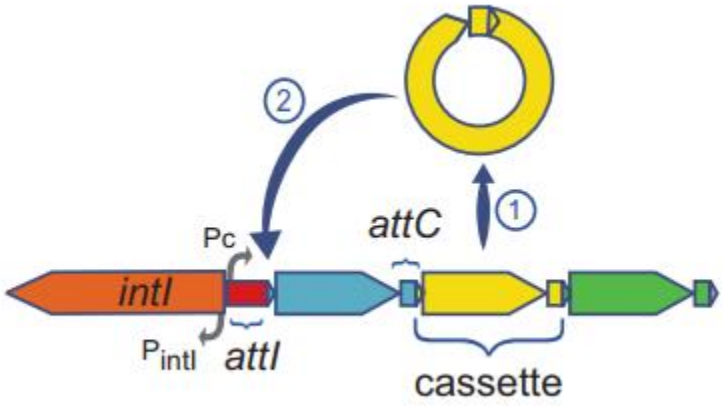


Figure 1.4 Representation of an integron element identified by IntegronFinder.

Integrons mainly consist of an integron integrase gene (intI, orange) including its own promoter (PintI), an attI recombination site (red), gene cassettes (blue, yellow, and green) along with one constitutive promoter (Pc) (Cury et al., 2016).

Genomic islands are one of the main elements in the evolution of microbial genomes. Depending on the gene functions they carry, they can be referred to as pathogenicity islands, symbiosis islands, metabolic islands, resistance islands, or fitness islands (Dobrindt et al., 2004; Juhas et al., 2009) (Figure 1.5.). Pathogenicity islands can contribute significantly to bacterial virulence or reduced virulence. In *Aeromonas caviae*,

a small genomic island is involved in flagellin glycosylation and lipopolysaccharide (LPS) O-antigen biosynthesis (Tabei et al., 2009).

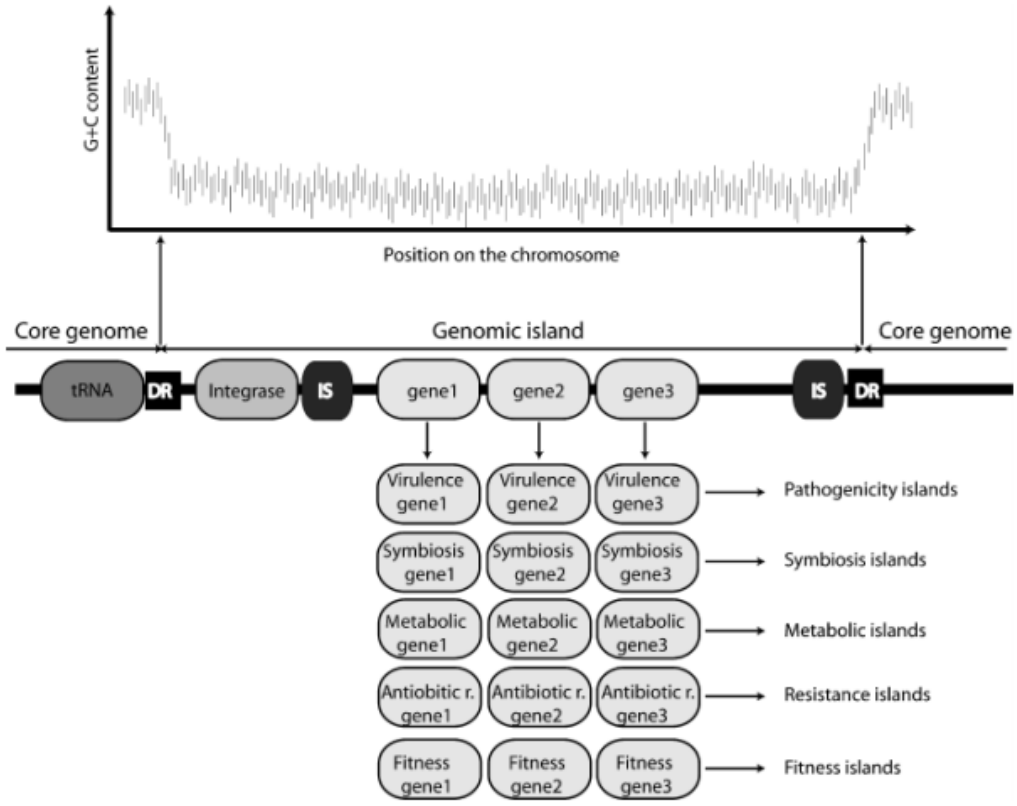


Figure 1.5 General features of genomic islands.

They typically harbor genes encoding proteins involved in pathogenicity, symbiosis, metabolism, antimicrobial resistance, or fitness for a specific environment (Juhas et al., 2009).

Bacteriophages have specific mechanisms that affect bacteria, including attachment, invasion, and survival (Boyd et al., 2012). Bacteriophages can change the genomic structure of bacteria to affect their virulence. Bacteriophages have been investigated as a treatment against bacterial pathogens. Since 1981, bacteriophages have been investigated as treatment for several bacteria (Figure 1.6) (Gon Choudhury et al.,

2017). Lytic bacteriophage (VTCCBPA6) isolated from the sewage of an organized equine breeding farm has the potential for use against *A. hydrophila* (Anand et al., 2016).

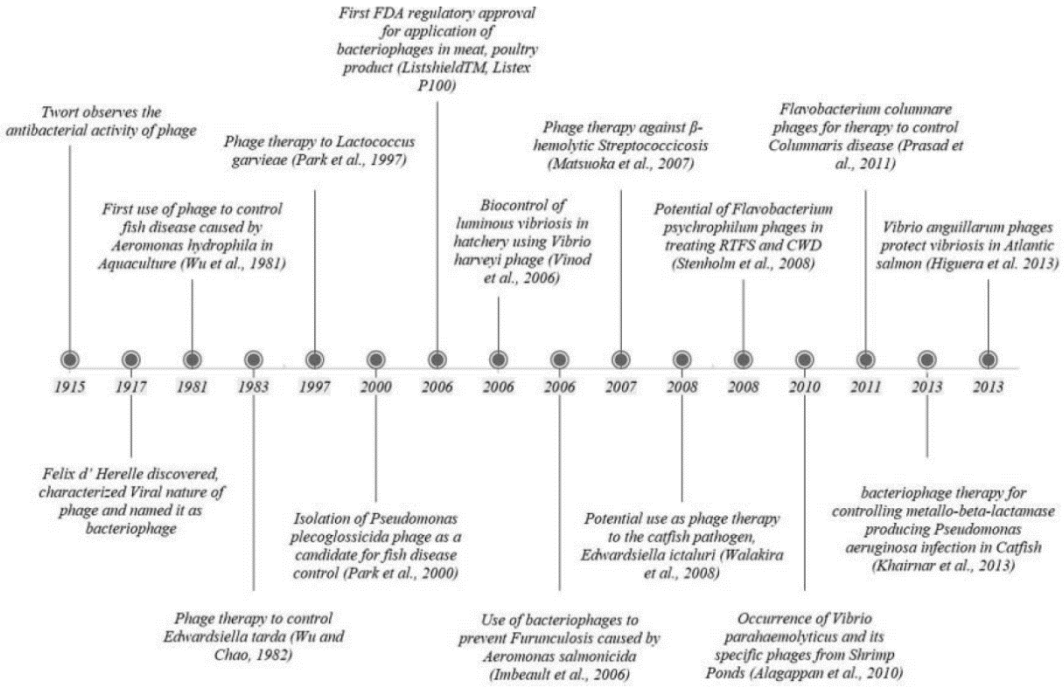


Figure 1.6 Major contributions to phage research in fish and shellfish research (Gon Choudhury et al., 2017).

Insertion elements are also important in genome evolution; they can mediate gene inactivation, genome structure rearrangement, or genome reduction. They can also incorporate genes that contribute to pathogenicity (Siguier et al., 2014; Touchon et al., 2007). In *Aeromonas salmonocida*, temperature directly affects insertion elements, which contribute to survival in harsh conditions and maintains genome structure (Vincent et al., 2016). An insertion sequence-dependent plasmid rearrangement in *A. salmonicida* caused loss of the T3SS (Tanaka et al., 2012). *A. salmonicida* insertion sequence AS5 (ISAS5) is

another active mobile genetic element involved in genomic plasticity of *A. salmonicida* (Trudel et al., 2013).

CRISPR-Cas systems are used by bacteria as an adaptable defense mechanism. They are used by many bacteria to combat predation and exposure to plasmids and viruses (Horvath et al., 2010; Sander et al., 2014). The system consists of two main components: CRISPR array and CRISPR associated genes (Cas), which are separated by spacers (Grissa et al., 2007). Plasmids are another mobilome element, the majority of which enable an advantage in a specific environment. Often they carry genes such as antimicrobial resistance, metal resistance, and virulence factors. *Aeromonas* species frequently tend to carry resistance plasmids (R-plasmids) (Piotrowska et al., 2015).

1.4 Genomics

The genome era has been a game changer to study any organism in detail. Due to rapid sequencing technology improvements and cost decreases, a quality draft genome can be generated in a couple of hours. According to NIH data, genome sequencing prices went down from $100 million per genome to $1000 in 14 years (Figure 7) (NHGRI, 2017).

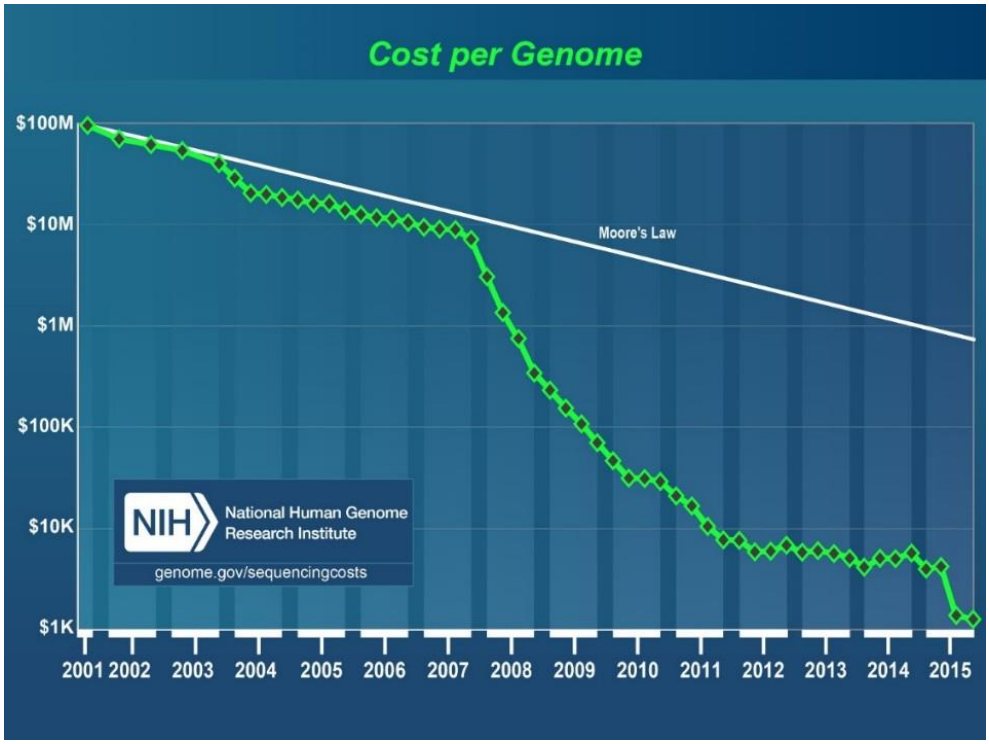


Figure 1.7 DNA Sequencing costs reduction over the years (NHGRI, 2017).

Studying pathogen genomics can decipher the origins of a pathogen; for instance, in 2011, pathogen genomics identified the origin of a Middle East respiratory syndrome (MERS) coronavirus epidemic in Europe (Luheshi et al., 2015). In 2012, two scientists from National Institutes of Health were able to track down patient-to-patient routes of transmission of highly virulent multidrug-resistant *Klebsiella pneumoniae* by applying real-time genomic sequencing (NIH, 2014). Similarly, the origin of virulent *A. hydrophila* in catfish aquaculture was determined using genomics (Hossain et al., 2014). Therefore, adoption of innovative genomics applications can be an effective and powerful strategy for solving disease outbreaks.

1.4.2 Genome sequencing

Whole Genome Sequencing (WGS) has been one of the most widely used applications of Next Generation Sequencing (NGS). It gives a comprehensive overview of genome structure and biological functions (Cirulli et al., 2010). Genome sequencing started in the mid-seventies by the development of the first generation of genome sequencing: chain termination by dideoxyribonucleotides (Sanger sequencing) (Maxam et al., 1977)(Sanger et al., 1975). Automation and improvements enabled the first genomes to be sequenced, but it was only feasible in large sequencing centers, and it was expensive.

Second generation sequencing was proposed by Ronanghi (Ronaghi et al., 1996) based on the sulfurylase-luciferase reaction (real-time sequencing) (Nyren, 1987). Roche 454 (pyrosequencing) was the first second generation sequencing method released, which was followed by Illumina Solexa (sequencing by synthesis) and later Helicos (Sequencing by Oligonucleotide Ligation and Detection or SOLiD). Ion Torrent and other alternatives were also developed. Third-generation sequencing (single-molecule sequencing) can produce longer reads than previous methods. MinION (Mikheyev et al., 2014; Quick et al., 2014, 2015) and Pacific Biosciences (PacBio) (Brown et al., 2014; Terabayashi et al., 2014) are examples. Milestones of bacterial genome sequencing are illustrated in Figure 1.8 (Loman et al., 2015).

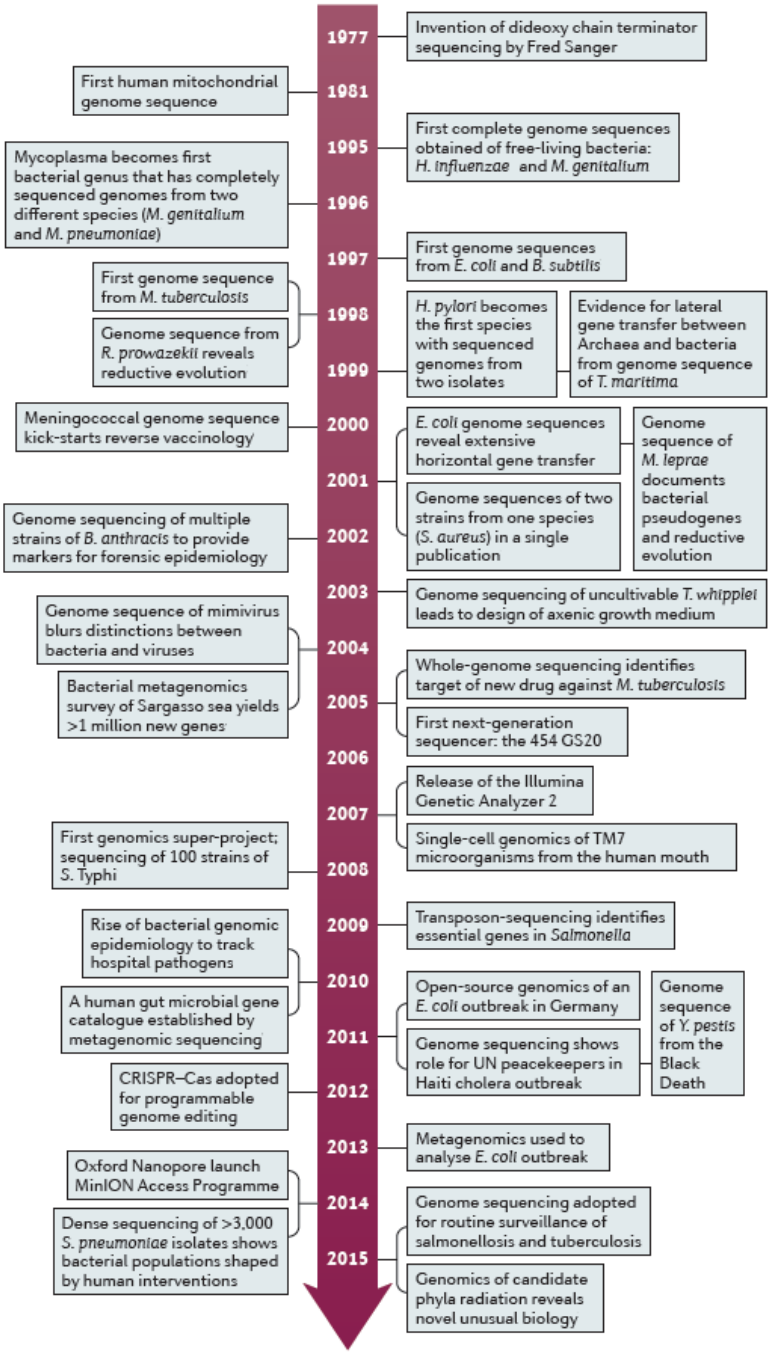


Figure 1.8 Development of bacterial genome sequencing (Loman et al., 2015).

1.4.3 Comparative genomics

Since the first generation of sequencing, understanding organisms at the DNA level has been the goal. Comparing either closely or distantly related genomes provides extensive information on intra- and inter-species diversity (Pallen et al., 2007). Evolution of pathogens occurs through horizontal acquisition of genetic information including plasmids, bacteriophages, integrons, and insertion elements (Maurelli, 2007). Random mutations, which can cause gene loss or gene variation, are another genetic mechanism affecting bacterial virulence without horizontal genetic material transfer of specific virulence factors (Sokurenko et al., 1999). Several factors contribute to bacterial genome dynamics (Figure 1.9), but briefly gene loss, gene gain, and gene change are the main reasons (Pallen et al., 2007). By applying comparative genomics methods, changes in genomic structure can be identified. Identification of lost or gained genes can also yield specific information about interactions with other organisms, including host-pathogen interactions.

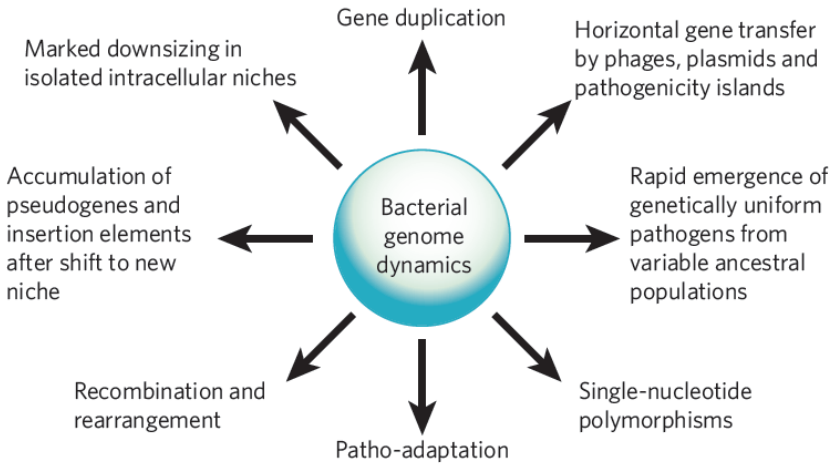


Figure 1.9 Potential effects caused by bacterial genome dynamics (Pallen et al., 2007).

Comparative genomics has yielded valuable information about several different organisms. Avian genome evolution and adaptation (G. Zhang et al., 2014) is a good example. Another example is the genomic changes underlying phenotypic convergence and adaptation of giant and red pandas to a specialized bamboo diet (Y. Hu et al., 2017). Convergence rates of evolution in ant-plant mutualisms (Rubin et al., 2016) is another good example. Comparative genomics research revealed high biological diversity and specific adaptations in the genus *Aspergillus*, which is a medically and industrially important organism (de Vries et al., 2017).

1.5 Significance of this study and objectives

Globally, aquaculture is one of the fastest growing agricultural sectors and provides almost half of the animal food for human consumption (Bostock et al., 2010). In the United States of America, the aquaculture industry has become well established in the last 35 years, and the majority of its production is in catfish (FAO, 2011). Sales from catfish aquaculture totaled $341 million in 2013 (USDA, 2013). Due to importance of catfish industry, multistate groups have collaborated to address emerging diseases and provide cooperative efforts to control the disease. In 2009 and 2010, catfish producers lost more than 5.5 million pounds of market size fish because of *A. hydrophila* epidemics. Genome sequencing revealed that these outbreaks were caused by an emergent clonal group of *A. hydrophila* strains referred to as virulent *A. hydrophila* (vAh). *A. hydrophila* produces a large number of extracellular enzymes and toxins that contribute to host inflammation and promote spreading in the tissues, and master regulator genes have not been studied extensively. Even though *E. piscicida* was recently identified and genetically characterized as a new taxon, it has been affecting U.S. aquaculture

negatively as well. Another potential problem is that non-native *A. veronii* species may have been introduced to U.S. aquaculture, which poses a potential problem.

However, to date, effective control measures for vAh have not been developed to reduce the impact of this pathogen on U.S. aquaculture. Furthermore, *E. piscicida* and *A. veronii* are emerging pathogens that threaten catfish production in the U.S. Therefore, there is a *critical gap* in our knowledge of molecular mechanisms of these pathogens and their potential virulence genes. Understanding of molecular mechanisms of these pathogens is an imperative step to fully understand their dissemination and how they induce inflammation. The *overall objectives* of the study are to sequence *E. piscicida*, *A. hydrophila*, and *A. veronii* and analyze specifically chosen invasins, toxins, and master regulator genes to understand their physiological and pathological roles in disease pathogenesis. Therefore, understanding molecular mechanism of these pathogens may allow design of novel control and/or treatment strategies. I accomplished my overall objective by pursuing the following specific aims:

1.5.1 Specific aim #1. Genome sequencing of *E. piscicida* and *A. hydrophila* isolates.

The working hypothesis was that genome sequencing of *E. piscicida* and *A. hydrophila* would accelerate comparative studies to delineate pathogenic mechanisms of these pathogens. This objective was accomplished using Illumina genome analyzer IIx and 454 GS-FLX Titanium platforms, resulting in a complete genome sequence of *E. piscicida* C07-087, one complete and four draft virulent *A. hydrophila* (vAh) genome sequences, and one complete and four draft genome sequences of other *A. hydrophila* strains.

1.5.2 Specific aim #2. Comparative genome analysis of *E. piscicida* and *E. ictaluri* genomes

The working hypothesis was that comparative genome studies would help to identify unique and shared characteristics of each pathogen, which could lead to designing a novel treatment method. This objective was accomplished applying several different comparative genomics approaches.

1.5.3 Specific aim #3. Comparative analysis of secretion systems of virulent *Aeromonas hydrophila* strains and mutation of *tssD* and *tssI* genes from T6SS.

The working hypothesis was that secretion system and effector proteins play an important role in disease pathogenesis of *A. hydrophila*. This objective was accomplished applying several different comparative genomics and along with in-frame deletion of *tssD* and *tssI* genes from T6SS.

1.5.4 Specific aim #4. Comparative genomics insights into pathogenicity of *A. veronii*

The working hypothesis was that virulent *A. veronii* strains may have been transferred from China to U.S., which would pose a risk for U.S. aquaculture. This objective was accomplished by applying several different comparative approaches to all the available *A. veronii* genomes in NCBI (as of 9/27/2017) along with our *A. veronii* genome.

1.5.5 Specific aim #5. Mutate selected *A. hydrophila* genes encoding secreted proteins (sialidase and collagenase), enterotoxin, and master regulator (Hfq) to determine their role in *A. hydrophila* virulence.

The working hypothesis was that mutation of selected *A. hydrophila* genes would result in decreased virulence of *A. hydrophila*. This objective was accomplished by in-

frame deletion mutation of selected *A. hydrophila* genes and determining virulence in catfish

1.6 References

Abayneh, T., Colquhoun, D. J., & Sorum, H. (2012). Multi-locus Sequence Analysis (MLSA) of Edwardsiella tarda isolates from fish. *Vet Microbiol, 158*(3-4), 367-375. doi:10.1016/j.vetmic.2012.03.006

Abayneh, T., Colquhoun, D. J., & Sorum, H. (2013). Edwardsiella piscicida sp. nov., a novel species pathogenic to fish. *J Appl Microbiol, 114*(3), 644-654. doi:10.1111/jam.12080

Abdelhamed, H., Ibrahim, I., Baumgartner, W., Lawrence, M. L., & Karsi, A. (2017). Characterization of Histopathological and Ultrastructural Changes in Channel Catfish Experimentally Infected with Virulent Aeromonas hydrophila. *Front Microbiol, 8*, 1519. doi:10.3389/fmicb.2017.01519

Abdelhamed, H., Ibrahim, I., Nho, S. W., Banes, M. M., Wills, R. W., Karsi, A., & Lawrence, M. L. (2017). Evaluation of three recombinant outer membrane proteins, OmpA1, Tdr, and TbpA, as potential vaccine antigens against virulent Aeromonas hydrophila infection in channel catfish (Ictalurus punctatus). *Fish Shellfish Immunol, 66*, 480-486. doi:10.1016/j.fsi.2017.05.043

Abdelhamed, H., Nho, S. W., Turaga, G., Banes, M. M., Karsi, A., & Lawrence, M. L. (2016). Protective efficacy of four recombinant fimbrial proteins of virulent Aeromonas hydrophila strain ML09-119 in channel catfish. *Vet Microbiol, 197*, 8-14. doi:10.1016/j.vetmic.2016.10.026

Anand, T., Vaid, R. K., Bera, B., Singh, J., Barua, S., Virmani, N., Rajukumar, K., Yadav, N.K., Nagar, D., Singh, R.K., Tripathi, B. N. (2016). Isolation of a lytic bacteriophage against virulent Aeromonas hydrophila from an organized equine farm. *J Basic Microbiol, 56*(4), 432-437. doi:10.1002/jobm.201500318

Aruni, W., Vanterpool, E., Osbourne, D., Roy, F., Muthiah, A., Dou, Y., & Fletcher, H. M. (2011). Sialidase and sialoglycoproteases can modulate virulence in Porphyromonas gingivalis. *Infect Immun, 79*(7), 2779-2791. doi:10.1128/IAI.00106-11

Ast, V. M., Schoenhofen, I. C., Langen, G. R., Stratilo, C. W., Chamberlain, M. D., & Howard, S. P. (2002). Expression of the ExeAB complex of Aeromonas hydrophila is required for the localization and assembly of the ExeD secretion port multimer. *Mol Microbiol, 44*(1), 217-231.

Austin, B., & Austin, A. D. (2012). *Bacterial Fish Pathogens*

Disease of Farmed and Wild Fish (5 ed.): Springer.

Basler, M., Ho, B. T., & Mekalanos, J. J. (2013). Tit-for-tat: type VI secretion system counterattack during bacterial cell-cell interactions. *Cell, 152*(4), 884-894. doi:10.1016/j.cell.2013.01.042

Beaz-Hidalgo, R., Hossain, M. J., Liles, M. R., & Figueras, M. J. (2015). Strategies to avoid wrongly labelled genomes using as example the detected wrong taxonomic affiliation for aeromonas genomes in the GenBank database. *PLoS One, 10*(1), e0115813. doi:10.1371/journal.pone.0115813

Bomar, L., Maltz, M., Colston, S., & Graf, J. (2011). Directed culturing of microorganisms using metatranscriptomics. *MBio, 2*(2), e00012-00011. doi:10.1128/mBio.00012-11

Bostock, J., McAndrew, B., Richards, R., Jauncey, K., Telfer, T., Lorenzen, K., Little, D., Ross, L., Handisyde, N., Gatward, I., Corner, R. (2010). Aquaculture: global status and trends. *Philos Trans R Soc Lond B Biol Sci, 365*(1554), 2897-2912. doi:10.1098/rstb.2010.0170

Boyd, E. F., Carpenter, M. R., & Chowdhury, N. (2012). Mobile effector proteins on phage genomes. *Bacteriophage, 2*(3), 139-148. doi:10.4161/bact.21658

Brown, S. D., Utturkar, S. M., Magnuson, T. S., Ray, A. E., Poole, F. L., Lancaster, W. A., Thorgersen, M.P., Adams, M.W., Elias, D. A. (2014). Complete Genome Sequence of Pelosinus sp. Strain UFO1 Assembled Using Single-Molecule Real-Time DNA Sequencing Technology. *Genome Announc, 2*(5). doi:10.1128/genomeA.00881-14

Buján, N., Toranzo, A., & B., M. (2017). Draft Genome Sequence of Edwardsiella piscicida Strain ACC35.1 Isolated from Diseased Turbot (Scophthalmus maximus) in Europe. *Genome Announc, 5*(7).

Buttner, D. (2012). Protein export according to schedule: architecture, assembly, and regulation of type III secretion systems from plant- and animal-pathogenic bacteria. *Microbiol Mol Biol Rev, 76*(2), 262-310. doi:10.1128/MMBR.05017-11

Cai, S. H., Wu, Z. H., Jian, J. C., Lu, Y. S., & Tang, J. F. (2012). Characterization of pathogenic Aeromonas veronii bv. veronii associated with ulcerative syndrome from chinese longsnout catfish (Leiocassis longirostris Gunther). *Braz J Microbiol, 43*(1), 382-388. doi:10.1590/S1517-838220120001000046

Carlson, C. M., Turpin, E. A., Moser, L. A., O'Brien, K. B., Cline, T. D., Jones, J. C., Tumpey, T.M., Katz, J.M., Kelley, L.A., Gauldie J., Schultz-Cherry, S. (2010). Transforming growth factor-beta: activation by neuraminidase and role in highly pathogenic H5N1 influenza pathogenesis. *PLoS Pathog, 6*(10), e1001136. doi:10.1371/journal.ppat.1001136

Chen, C., Zhu, N., Kong, L., Bei, Y., Zheng, T., Ding, X., & He, Z. (2013). First case of soft shell disease in Chinese soft-shelled turtle (Trionyx sinens) associated with Aeromonas sobria–A. veronii complex. *Aquaculture, 406-407*, 62-67.

Chen, G. Y., Chen, X., King, S., Cavassani, K. A., Cheng, J., Zheng, X., Cao, H., Yu, H., Qu, J., Fang, D., Wu, W., Bai, X.F., Liu, J.Q., Woodiga, S.A., Chen, C., Sun, L., Hogaboam, C.M., Kunkel, S.L., Zheng, P., Liu, Y. (2011). Amelioration of sepsis by inhibiting sialidase-mediated disruption of the CD24-SiglecG interaction. *Nat Biotechnol, 29*(5), 428-435. doi:10.1038/nbt.1846

Chigwechokha, P. K., Tabata, M., Shinyoshi, S., Oishi, K., Araki, K., Komatsu, M., Itakura, T., Shiozaki, K. (2015). Recombinant sialidase NanA (rNanA) cleaves alpha2-3 linked sialic acid of host cell surface N-linked glycoprotein to promote Edwardsiella tarda infection. *Fish Shellfish Immunol, 47*(1), 34-45. doi:10.1016/j.fsi.2015.08.015

Choe, Y., Park, J., Yu, J. E., Oh, J. I., Kim, S., & Kang, H. Y. (2017). Edwardsiella piscicida lacking the cyclic AMP receptor protein (Crp) is avirulent and immunogenic in fish. *Fish Shellfish Immunol, 68*, 243-250. doi:10.1016/j.fsi.2017.06.060

Christiansen, J. K., Larsen, M. H., Ingmer, H., Sogaard-Andersen, L., & Kallipolitis, B. H. (2004). The RNA-binding protein Hfq of Listeria monocytogenes: role in stress tolerance and virulence. *J Bacteriol, 186*(11), 3355-3362. doi:10.1128/JB.186.11.3355-3362.2004

Chung, T. H., Yi, S. W., & Shin, G. W. (2017). Antibiotic resistance and repetitive-element PCR fingerprinting in Aeromonas veronii isolates. *J Fish Dis, 40*(6), 821-829. doi:10.1111/jfd.12564

Cianciotto, N. P. (2005). Type II secretion: a protein secretion system for all seasons. *Trends Microbiol, 13*(12), 581-588. doi:10.1016/j.tim.2005.09.005

Cianfanelli, F. R., Monlezun, L., & Coulthurst, S. J. (2016). Aim, Load, Fire: The Type VI Secretion System, a Bacterial Nanoweapon. *Trends Microbiol, 24*(1), 51-62. doi:10.1016/j.tim.2015.10.005

Cipriano, R. C. (2001). Aeromonas hydrophila and Motile Aeromonad Septicemias of Fish. *United States Department Of The Interior-Fish and Wildlife Service Division of Fishery Research Washington*.

Cirulli, E. T., & Goldstein, D. B. (2010). Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet, 11*(6), 415-425. doi:10.1038/nrg2779

Costa, T. R., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M., & Waksman, G. (2015). Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat Rev Microbiol, 13*(6), 343-359. doi:10.1038/nrmicro3456

Cury, J., Jove, T., Touchon, M., Neron, B., & Rocha, E. P. (2016). Identification and analysis of integrons and cassette arrays in bacterial genomes. *Nucleic Acids Res, 44*(10), 4539-4550. doi:10.1093/nar/gkw319

Darbari, V. C., & Waksman, G. (2015). Structural Biology of Bacterial Type IV Secretion Systems. *Annual Review of Biochemistry, 84*(1), 603-629. doi:10.1146/annurev-biochem-062911-102821

Dash, P., Sahoo, P. K., Gupta, P. K., Garg, L. C., & Dixit, A. (2014). Immune responses and protective efficacy of recombinant outer membrane protein R (rOmpR)-based vaccine of Aeromonas hydrophila with a modified adjuvant formulation in rohu (Labeo rohita). *Fish Shellfish Immunol, 39*(2), 512-523. doi:10.1016/j.fsi.2014.06.007

de Vries, R. P., Riley, R., Wiebenga, A., Aguilar-Osorio, G., Amillis, S., Uchima, C. A., Grigoriev, I. V. et al. (2017). Comparative genomics reveals high biological diversity and specific adaptations in the industrially and medically important fungal genus Aspergillus. *Genome Biol, 18*(1), 28. doi:10.1186/s13059-017-1151-0

Deng, W., Marshall, N. C., Rowland, J. L., McCoy, J. M., Worrall, L. J., Santos, A. S., Strynadka, N.C.J., Finlay, B. B. (2017). Assembly, structure, function and regulation of type III secretion systems. *Nat Rev Microbiol, 15*(6), 323-337. doi:10.1038/nrmicro.2017.20

Deng, Y., Chen, C., Zhao, Z., Zhao, J., Jacq, A., Huang, X., & Yang, Y. (2016). The RNA Chaperone Hfq Is Involved in Colony Morphology, Nutrient Utilization and Oxidative and Envelope Stress Response in Vibrio alginolyticus. *PLoS One, 11*(9), e0163689. doi:10.1371/journal.pone.0163689

DePaola, A., Peeler, J. T., & Rodrick, G. E. (1995). Effect of oxytetracycline-medicated feed on antibiotic resistance of gram-negative bacteria in catfish ponds. *Appl Environ Microbiol, 61*(6), 2335-2340.

Dietrich, M., Munke, R., Gottschald, M., Ziska, E., Boettcher, J. P., Mollenkopf, H., & Friedrich, A. (2009). The effect of hfq on global gene expression and virulence in Neisseria gonorrhoeae. *FEBS J, 276*(19), 5507-5520. doi:10.1111/j.1742-4658.2009.07234.x

Ding, Y., Davis, B. M., & Waldor, M. K. (2004). Hfq is essential for Vibrio cholerae virulence and downregulates sigma expression. *Mol Microbiol, 53*(1), 345-354. doi:10.1111/j.1365-2958.2004.04142.x

Dobrindt, U., Hochhut, B., Hentschel, U., & Hacker, J. (2004). Genomic islands in pathogenic and environmental microorganisms. *Nat Rev Micro, 2*(5), 414-424.

Dong, H. T., Techatanakitarnan, C., Jindakittikul, P., Thaiprayoon, A., Taengphu, S., Charoensapsri, W., Khunrae, P., Rattanarojpong, T., Senapin, S. (2017). Aeromonas jandaei and Aeromonas veronii caused disease and mortality in Nile tilapia, Oreochromis niloticus (L.). *J Fish Dis*. doi:10.1111/jfd.12617

Duarte, A. S., Cavaleiro, E., Pereira, C., Merino, S., Esteves, A. C., Duarte, E. P., Tomás J.M., Correia, A. C. (2015). Aeromonas piscicola AH-3 expresses an extracellular collagenase with cytotoxic properties. *Lett Appl Microbiol, 60*(3), 288-297. doi:10.1111/lam.12373

Duarte, A. S., Correia, A., & Esteves, A. C. (2016). Bacterial collagenases - A review. *Crit Rev Microbiol, 42*(1), 106-126. doi:10.3109/1040841X.2014.904270

Evans JJ, Klesius PH, Plumb JA, & CA, S. (2011). *Fish diseases and disorders* (Woo PTK & B. DW Eds.). Oxfordshire, United Kingdom: CAB International.

Ewing, W. H., Mcwhorter, A. C., Escobar, M. R., & Lubin, A. H. (1965). Edwardsiella, a new genus of Enterobacteriaceae based on a new species, E. tarda. *International Journal of Systematic and Evolutionary Microbiology, 15*(1), 33-38. doi:doi:10.1099/00207713-15-1-33

FAO. (2016). The State of World Fisheries and Aquaculture 2016. Retrieved from http://www.fao.org/3/a-i5555e.pdf

FAO. (2017a). Aquaculture. Retrieved from http://www.fao.org/fishery/aquaculture/en

FAO. (2017b). Aquaculture development. Retrieved from http://www.fao.org/aquaculture/en/

Feliciano, J. R., Grilo, A. M., Guerreiro, S. I., Sousa, S. A., & Leitao, J. H. (2016). Hfq: a multifaceted RNA chaperone involved in virulence. *Future Microbiol, 11*(1), 137-151. doi:10.2217/fmb.15.128

Fogelson, S. B., Petty, B. D., Reichley, S. R., Ware, C., Bowser, P. R., Crim, M. J., Getchell, R.G., Sams, K.L., Marquis, H., Griffin, M. J. (2016). Histologic and molecular characterization of Edwardsiella piscicida infection in largemouth bass (Micropterus salmoides). *J Vet Diagn Invest, 28*(3), 338-344. doi:10.1177/1040638716637639

Furmanek-Blaszk, B. (2014). Phenotypic and molecular characteristics of an Aeromonas hydrophila strain isolated from the River Nile. *Microbiol Res, 169*(7-8), 547-552. doi:10.1016/j.micres.2013.11.001

Geng, J., Song, Y., Yang, L., Feng, Y., Qiu, Y., Li, G., Guo, J., Bi, Y.,Qu, Y., Wang, W., Wang, X., Guo, Z., Yang, R., Han, Y. (2009). Involvement of the post-transcriptional regulator Hfq in Yersinia pestis virulence. *PLoS One, 4*(7), e6213. doi:10.1371/journal.pone.0006213

Gon Choudhury, T., Tharabenahalli Nagaraju, V., Gita, S., Paria, A., & Parhi, J. (2017). Advances in Bacteriophage Research for Bacterial Disease Control in Aquaculture. *Reviews in Fisheries Science & Aquaculture, 25*(2), 113-125. doi:10.1080/23308249.2016.1241977

Green, E. R., & Mecsas, J. (2016). Bacterial Secretion Systems: An Overview. *Microbiol Spectr, 4*(1). doi:10.1128/microbiolspec.VMBF-0012-2015

Griffin, M. J., Ware, C., Quiniou, S. M., Steadman, J. M., Gaunt, P. S., Khoo, L. H., & Soto, E. (2014). Edwardsiella piscicida identified in the Southeastern USA by gyrB sequence, species-specific and repetitive sequence-mediated PCR. *Dis Aquat Organ, 108*(1), 23-35. doi:10.3354/dao02687

Grim, C. J., Kozlova, E. V., Ponnusamy, D., Fitts, E. C., Sha, J., Kirtley, M. L., Van Lier, C. J., Tiner, B. L., Erova, T. E., Joseph, S. J., Read, T. D., Shak, J. R., Joseph, S. W., Singletary, E., Felland, T., Baze, W. B., Horneman, A. J., Chopra, A. K. (2014). Functional genomic characterization of virulence factors from necrotizing fasciitis-causing strains of Aeromonas hydrophila. *Appl Environ Microbiol, 80*(14), 4162-4183. doi:10.1128/AEM.00486-14

Grim, C. J., Kozlova, E. V., Sha, J., Fitts, E. C., van Lier, C. J., Kirtley, M. L., Joseph, S.J., Read, T. D., Burd, E. M., Tall, B. D., Joseph, S. W., Horneman, A.J., Chopra, A. K., Shak, J. R. (2013). Characterization of Aeromonas hydrophila wound pathotypes by comparative genomic and functional analyses of virulence genes. *MBio, 4*(2), e00064-00013. doi:10.1128/mBio.00064-13

Grimont, P. A. D., Grimont, F., Richard, C., & Sakazaki, R. (1980). Edwardsiella hoshinae, a new species of enterobacteriaceae. *Current Microbiology, 4*(6), 347-351. doi:10.1007/bf02605375

Grissa, I., Vergnaud, G., & Pourcel, C. (2007). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics, 8*, 172. doi:10.1186/1471-2105-8-172

Grizzle, J. M., & Kiryu, Y. (1993). Histopathology of gill, liver, and pancreas, and serum enzyme levels of channel catfish infected with Aeromonas hydrophila complex. *Journal of Aquatic Animal Health, 5*, 36-50.

Hall, R. M., Collis, C. M., Kim, M. J., Partridge, S. R., Recchia, G. D., & Stokes, H. W. (1999). Mobile gene cassettes and integrons in evolution. *Ann N Y Acad Sci, 870*, 68-80.

Han, H. J., Taki, T., Kondo, H., Hirono, I., & Aoki, T. (2008). Pathogenic potential of a collagenase gene from Aeromonas veronii. *Can J Microbiol, 54*(1), 1-10. doi:10.1139/w07-109

Hassan, M. A., Noureldin, E. A., Mahmoud, M. A., & Fita, N. A. (2017). Molecular identification and epizootiology of Aeromonas veronii infection among farmed Oreochromis niloticus in Eastern Province, KSA. *The Egyptian Journal of Aquatic Research, 43*(2), 161-167. doi:http://dx.doi.org/10.1016/j.ejar.2017.06.001

Hawke, J. P., Mcwhorter, A. C., Steigerwalt, A. G., & Brenner, D. J. (1981). Edwardsiella ictaluri sp. nov., the Causative Agent of Enteric Septicemia of Catfish. *International Journal of Systematic and Evolutionary Microbiology, 31*(4), 396-400. doi:doi:10.1099/00207713-31-4-396

Horvath, P., & Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science, 327*(5962), 167-170. doi:10.1126/science.1179555

Hossain, M. J., Sun, D., McGarey, D. J., Wrenn, S., Alexander, L. M., Martino, M. E., Xinge, Y., Terhune, J. S., Liles, M. R. (2014). An Asian origin of virulent Aeromonas hydrophila responsible for disease epidemics in United States-farmed catfish. *MBio, 5*(3), e00848-00814. doi:10.1128/mBio.00848-14

Hossain, M. J., Waldbieser, G. C., Sun, D., Capps, N. K., Hemstreet, W. B., Carlisle, K., Griffin, M. J., Khoo, L., Goodwin, A. E., Sonstegard, T. S., Schroeder, S., Hayden, K., Newton, J. C., Terhune, J. S., Liles, M. R. (2013). Implication of lateral genetic transfer in the emergence of Aeromonas hydrophila isolates of epidemic outbreaks in channel catfish. *PLoS One, 8*(11), e80943. doi:10.1371/journal.pone.0080943

Hou, M., Chen, R., Yang, D., Nunez, G., Wang, Z., Wang, Q., Zhang, Y., Liu, Q. (2017). Identification and functional characterization of EseH, a new effector of the type III secretion system of Edwardsiella piscicida. *Cell Microbiol, 19*(1). doi:10.1111/cmi.12638

Hu, Y., Wu, Q., Ma, S., Ma, T., Shan, L., Wang, X., Niea, Y., Ningc, Z., Yana, L., Xiu, Y., Wei, F. (2017). Comparative genomics reveals convergent evolution between the bamboo-eating giant and red pandas. *Proc Natl Acad Sci U S A, 114*(5), 1081-1086. doi:10.1073/pnas.1613870114

Hu, Y. H., Li, Y. X., & Sun, L. (2014). Edwardsiella tarda Hfq: impact on host infection and global protein expression. *Vet Res, 45*, 23. doi:10.1186/1297-9716-45-23

Igbinosa, I. H., Igumbor, E. U., Aghdasi, F., Tom, M., & Okoh, A. I. (2012). Emerging Aeromonas species infections and their significance in public health. *ScientificWorldJournal, 2012*, 625023. doi:10.1100/2012/625023

Jacobs, L., & Chenia, H. Y. (2007). Characterization of integrons and tetracycline resistance determinants in Aeromonas spp. isolated from South African aquaculture systems. *Int J Food Microbiol, 114*(3), 295-306. doi:10.1016/j.ijfoodmicro.2006.09.030

Janda, J. M., & Abbott, S. L. (1998). Evolving concepts regarding the genus Aeromonas: an expanding Panorama of species, disease presentations, and unanswered questions. *Clinical Infectious Diseases*(27), 332-344.

Janda, J. M., & Abbott, S. L. (2010). The genus Aeromonas: taxonomy, pathogenicity, and infection. *Clin Microbiol Rev, 23*(1), 35-73. doi:10.1128/CMR.00039-09

Jin, R. P., Hu, Y. H., Sun, B. G., Zhang, X. H., & Sun, L. (2012). Edwardsiella tarda sialidase: pathogenicity involvement and vaccine potential. *Fish Shellfish Immunol, 33*(3), 514-521. doi:10.1016/j.fsi.2012.06.002

Jorgensen, T. S., Kiil, A. S., Hansen, M. A., Sorensen, S. J., & Hansen, L. H. (2014). Current strategies for mobilome research. *Front Microbiol, 5*, 750. doi:10.3389/fmicb.2014.00750

Juhas, M., van der Meer, J. R., Gaillard, M., Harding, R. M., Hood, D. W., & Crook, D. W. (2009). Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev, 33*(2), 376-393. doi:10.1111/j.1574-6976.2008.00136.x

Kanonenberg, K., Schwarz, C. K., & Schmitt, L. (2013). Type I secretion systems - a story of appendices. *Res Microbiol, 164*(6), 596-604. doi:10.1016/j.resmic.2013.03.011

Karunasagar, I., Rosalind, G. M., Karunasagar, I., & Gopal Rao, K. (1989). Aeromonas hydrophila septicaemia of Indian major carps in some commercial fish farms of West Godavari District, Andhra Pradesh. *Curr. Sci., 58*, 1044-1045.

Kassegne, K., Hu, W., Ojcius, D. M., Sun, D., Ge, Y., Zhao, J., Yang, X.F., Li, L., Yan, J. (2014). Identification of collagenase as a critical virulence factor for invasiveness and transmission of pathogenic Leptospira species. *J Infect Dis, 209*(7), 1105-1115. doi:10.1093/infdis/jit659

Korotkov, K. V., Sandkvist, M., & Hol, W. G. (2012). The type II secretion system: biogenesis, molecular architecture and mechanism. *Nat Rev Microbiol, 10*(5), 336-351. doi:10.1038/nrmicro2762

Kuroiwa, A., Hisatsune, A., Isohama, Y., & Katsuki, H. (2009). Bacterial neuraminidase increases IL-8 production in lung epithelial cells via NF-kappaB-dependent pathway. *Biochem Biophys Res Commun, 379*(3), 754-759. doi:10.1016/j.bbrc.2008.12.120

Lasica, A. M., Ksiazek, M., Madej, M., & Potempa, J. (2017). The Type IX Secretion System (T9SS): Highlights and Recent Insights into Its Structure and Function. *Front Cell Infect Microbiol, 7*, 215. doi:10.3389/fcimb.2017.00215

Li, C., Kurniyati, Hu, B., Bian, J., Sun, J., Zhang, W., Liu, J., Pan, Y., Li, C. (2012). Abrogation of neuraminidase reduces biofilm formation, capsule biosynthesis, and virulence of Porphyromonas gingivalis. *Infect Immun, 80*(1), 3-13. doi:10.1128/IAI.05773-11

Li, H., Zhu, Q. F., Peng, X. X., & Peng, B. (2017). Interactome of E. piscicida and grouper liver proteins reveals strategies of bacterial infection and host immune response. *Sci Rep, 7*, 39824. doi:10.1038/srep39824

Liu, D., Geng, Y., Wang, K., Chen, D., Huang, X. L., Ou, Y., He, C. L., Zhong, Z. J., Lai, W. (2016). Aeromonas veronii infection in cultured channel catfish, ictalurus punctatus, in Southwest China. *The Israeli Journal of Aquaculture, 68*.

Liu, Y., Gao, Y., Liu, X., Liu, Q., Zhang, Y., Wang, Q., & Xiao, J. (2017). Transposon insertion sequencing reveals T4SS as the major genetic trait for conjugation transfer of multi-drug resistance pEIB202 from Edwardsiella. *BMC Microbiol, 17*(1), 112. doi:10.1186/s12866-017-1013-7

Liu, Y., Zhao, L., Yang, M., Yin, K., Zhou, X., Leung, K. Y., Liu, Q., Zhang, Y., Wang, Q. (2017). Transcriptomic dissection of the horizontally acquired response regulator EsrB reveals its global regulatory roles in the physiological adaptation and activation of T3SS and the cognate effector repertoire in Edwardsiella piscicida during infection toward turbot. *Virulence*, 1-23. doi:10.1080/21505594.2017.1323157

Loman, N. J., & Pallen, M. J. (2015). Twenty years of bacterial genome sequencing. *Nat Rev Microbiol, 13*(12), 787-794. doi:10.1038/nrmicro3565

Luheshi, L. M., Raza, S., & Peacock, S. J. (2015). Moving pathogen genomics out of the lab and into the clinic: what will it take? *Genome Med, 7*(1), 132. doi:10.1186/s13073-015-0254-z

Lv, Y., Xiao, J., Liu, Q., Wu, H., Zhang, Y., & Wang, Q. (2012). Systematic mutation analysis of two-component signal transduction systems reveals EsrA-EsrB and PhoP-PhoQ as the major virulence regulators in Edwardsiella tarda. *Vet Microbiol, 157*(1-2), 190-199. doi:10.1016/j.vetmic.2011.12.018

Lv, Y., Yin, K., Shao, S., Wang, Q., & Zhang, Y. (2013). Comparative proteomic analysis reveals new components of the PhoP regulon and highlights a role for PhoP in the regulation of genes encoding the F1F0 ATP synthase in Edwardsiella tarda. *Microbiology, 159*(Pt 7), 1340-1351. doi:10.1099/mic.0.066803-0

Lybecker, M. C., Abel, C. A., Feig, A. L., & Samuels, D. S. (2010). Identification and function of the RNA chaperone Hfq in the Lyme disease spirochete Borrelia burgdorferi. *Mol Microbiol, 78*(3), 622-635. doi:10.1111/j.1365-2958.2010.07374.x

Maltz, M. A., Bomar, L., Lapierre, P., Morrison, H. G., McClure, E. A., Sogin, M. L., & Graf, J. (2014). Metagenomic analysis of the medicinal leech gut microbiota. *Front Microbiol, 5*, 151. doi:10.3389/fmicb.2014.00151

Maurelli, A. T. (2007). Black holes, antivirulence genes, and gene inactivation in the evolution of bacterial pathogens. *FEMS Microbiol Lett, 267*(1), 1-8. doi:10.1111/j.1574-6968.2006.00526.x

Maxam, A. M., & Gilbert, W. (1977). A new method for sequencing DNA. *Proc Natl Acad Sci U S A, 74*(2), 560-564.

McGarey, D. J., Milanesi, L., Foley, D. P., Reyes, B. J., Frye, L., & Lim, D. V. (1991). The role of motile aeromonads in the fish disease, ulcerative disease syndrome (UDS). *Experientia, 47*(5), 441-444.

McNealy, T. L., Forsbach-Birk, V., Shi, C., & Marre, R. (2005). The Hfq homolog in Legionella pneumophila demonstrates regulation by LetA and RpoS and interacts with the global regulator CsrA. *J Bacteriol, 187*(4), 1527-1532. doi:10.1128/JB.187.4.1527-1532.2005

Meyer, F. P., & Bullock, G. L. (1973). Edwardsiella tarda, a New Pathogen of Channel Catfish (Ictalurus punctatus). *Applied Microbiology, 25*(1), 155-156.

Mikheyev, A. S., & Tin, M. M. (2014). A first look at the Oxford Nanopore MinION sequencer. *Mol Ecol Resour, 14*(6), 1097-1102. doi:10.1111/1755-0998.12324

Moura, A., Oliveira, C., Henriques, I., Smalla, K., & Correia, A. (2012). Broad diversity of conjugative plasmids in integron-carrying bacteria from wastewater environments. *FEMS Microbiol Lett, 330*(2), 157-164. doi:10.1111/j.1574-6968.2012.02544.x

NASS. (2017). *Catfish Production*. Retrieved from

NCBI. (2017). Genome Assembly and Annotation report for *Aeromonas hydrophila* genomes. Retrieved from http://www.ncbi.nlm.nih.gov/genome/genomes/1422?

NHGRI. (2017). DNA Sequencing Costs: Data. Retrieved from https://www.genome.gov/27541954/dna-sequencing-costs-data/

Nielsen, J. S., Lei, L. K., Ebersbach, T., Olsen, A. S., Klitgaard, J. K., Valentin-Hansen, P., & Kallipolitis, B. H. (2010). Defining a role for Hfq in Gram-positive bacteria: evidence for Hfq-dependent antisense regulation in Listeria monocytogenes. *Nucleic Acids Res, 38*(3), 907-919. doi:10.1093/nar/gkp1081

NIH. (2014). Segre Group. Retrieved from https://www.genome.gov/10000354/segre--group/

Nyren, P. (1987). Enzymatic method for continuous monitoring of DNA polymerase activity. *Anal Biochem, 167*(2), 235-238.

Oguro, K., Tamura, K., Yamane, J., Shimizu, M., Yamamoto, T., Ikawa, T., Ohnishi, K., Oshima, S., Imajoh, M. (2014). Draft Genome Sequences of Two Genetic Variant Strains of Edwardsiella piscicida, JF1305 and RSB1309, Isolated from Olive Flounder (Paralichythys olivaceus) and Red Sea Bream (Pagrus major) Cultured in Japan, Respectively. *Genome Announc, 2*(3). doi:10.1128/genomeA.00546-14

Overturf, K. (2009). *Molecular Research in Aquaculture*: Wiley-Blackwell.

Pallen, M. J., & Wren, B. W. (2007). Bacterial pathogenomics. *Nature, 449*(7164), 835-842. doi:10.1038/nature06248

Pang, M., Jiang, J., Xie, X., Wu, Y., Dong, Y., Kwok, A. H., Zhang, W., Yao, H., Lu, C., Leung, F.C., Liu, Y. (2015). Novel insights into the pathogenicity of epidemic Aeromonas hydrophila ST251 clones from comparative genomics. *Sci Rep, 5*, 9833. doi:10.1038/srep09833

Park, S. B., Aoki, T., & Jung, T. S. (2012). Pathogenesis of and strategies for preventing Edwardsiella tarda infection in fish. *Vet Res, 43*, 67. doi:10.1186/1297-9716-43-67

Piotrowska, M., & Popowska, M. (2015). Insight into the mobilome of Aeromonas strains. *Frontiers in Microbiology, 6*(494). doi:10.3389/fmicb.2015.00494

Piotrowska, M., & Popowska, M. (2015). Insight into the mobilome of Aeromonas strains. *Front Microbiol, 6*, 494. doi:10.3389/fmicb.2015.00494

Plumb, J. A., & Hanson, L. A. (2011). *Health Maintenance and Principal Microbial Diseases of Cultured Fishes* (Third Edition ed.).

Plumb, J. A., Sheifinger, C. C., Shryock, T. R., & Goldsby, T. (1995). Susceptibility of Six Bacterial Pathogens of Channel Catfish to Six Antibiotics. *Journal of Aquatic Animal Health, 7*, 211-217.

Ponnusamy, D., Kozlova, E. V., Sha, J., Erova, T. E., Azar, S. R., Fitts, E. C., Kirtley, M.L., Tiner, B.L., Andersson, J.A., Grim, C.J., Isom, R.P., Hasan, N.A., Colwell, R.R., Chopra, A. K. (2016). Cross-talk among flesh-eating Aeromonas hydrophila strains in mixed infection leading to necrotizing fasciitis. *Proc Natl Acad Sci U S A, 113*(3), 722-727. doi:10.1073/pnas.1523817113

Pridgeon, J. W., & Klesius, P. H. (2011). Development and efficacy of a novobiocin-resistant Streptococcus iniae as a novel vaccine in Nile tilapia (Oreochromis niloticus). *Vaccine, 29*(35), 5986-5993. doi:10.1016/j.vaccine.2011.06.036

Pridgeon, J. W., Klesius, P. H., Song, L., Zhang, D., Kojima, K., & Mobley, J. A. (2013). Identification, virulence, and mass spectrometry of toxic ECP fractions of West Alabama isolates of Aeromonas hydrophila obtained from a 2010 disease outbreak. *Vet Microbiol, 164*(3-4), 336-343. doi:10.1016/j.vetmic.2013.02.020

Pukatzki, S., McAuley, S. B., & Miyata, S. T. (2009). The type VI secretion system: translocation of effectors and effector-domains. *Curr Opin Microbiol, 12*(1), 11-17. doi:10.1016/j.mib.2008.11.010

Quick, J., Quinlan, A. R., & Loman, N. J. (2014). A reference bacterial genome dataset generated on the MinION portable single-molecule nanopore sequencer. *Gigascience, 3*, 22. doi:10.1186/2047-217X-3-22

Quick, J., Quinlan, A. R., & Loman, N. J. (2015). Erratum: A reference bacterial genome dataset generated on the MinION(TM) portable single-molecule nanopore sequencer. *Gigascience, 4*, 6. doi:10.1186/s13742-015-0043-z

Reichley, S. R., Waldbieser, G. C., Lawrence, M. L., & Griffin, M. J. (2015). Complete Genome Sequence of an Edwardsiella piscicida-Like Species, Recovered from Tilapia in the United States. *Genome Announc, 3*(5). doi:10.1128/genomeA.01004-15

Reichley, S. R., Waldbieser, G. C., Tekedar, H. C., Lawrence, M. L., & Griffin, M. J. (2016). Complete Genome Sequence of Edwardsiella piscicida Isolate S11-285 Recovered from Channel Catfish (Ictalurus punctatus) in Mississippi, USA. *Genome Announc, 4*(6). doi:10.1128/genomeA.01259-16

Reichley, S. R., Waldbieser, G. C., Ucko, M., Colorni, A., Dubytska, L., Thune, R. L., Lawrence, M.L., Griffin, M. J. (2015). Complete Genome Sequence of an Edwardsiella piscicida-Like Species Isolated from Diseased Grouper in Israel. *Genome Announc, 3*(4). doi:10.1128/genomeA.00829-15

Reichley, S. R., Ware, C., Greenway, T. E., Wise, D. J., & Griffin, M. J. (2015). Real-time polymerase chain reaction assays for the detection and quantification of Edwardsiella tarda, Edwardsiella piscicida, and Edwardsiella piscicida-like species in catfish tissues and pond water. *J Vet Diagn Invest, 27*(2), 130-139. doi:10.1177/1040638714566672

Robertson, G. T., & Roop, R. M., Jr. (1999). The Brucella abortus host factor I (HF-I) protein contributes to stress resistance during stationary phase and is a major determinant of virulence in mice. *Mol Microbiol, 34*(4), 690-700.

Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlen, M., & Nyren, P. (1996). Real-time DNA sequencing using detection of pyrophosphate release. *Anal Biochem, 242*(1), 84-89. doi:10.1006/abio.1996.0432

Rubin, B. E., & Moreau, C. S. (2016). Comparative genomics reveals convergent rates of evolution in ant-plant mutualisms. *Nat Commun, 7*, 12679. doi:10.1038/ncomms12679

Russell, A. B., Peterson, S. B., & Mougous, J. D. (2014). Type VI secretion system effectors: poisons with a purpose. *Nat Rev Microbiol, 12*(2), 137-148. doi:10.1038/nrmicro3185

Sander, J. D., & Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol, 32*(4), 347-355. doi:10.1038/nbt.2842

Sanger, F., & Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol, 94*(3), 441-448.

Schmidt, A. S., Bruun, M. S., Dalsgaard, I., & Larsen, J. L. (2001). Incidence, distribution, and spread of tetracycline resistance determinants and integron-associated antibiotic resistance genes among motile aeromonads from a fish farming environment. *Appl Environ Microbiol, 67*(12), 5675-5682. doi:10.1128/AEM.67.12.5675-5682.2001

Sha, J., Kozlova, E. V., & Chopra, A. K. (2002). Role of various enterotoxins in Aeromonas hydrophila-induced gastroenteritis: generation of enterotoxin gene-deficient mutants and evaluation of their enterotoxic activity. *Infect Immun, 70*(4), 1924-1935.

Sha, J., Pillai, L., Fadl, A. A., Galindo, C. L., Erova, T. E., & Chopra, A. K. (2005). The type III secretion system and cytotoxic enterotoxin alter the virulence of Aeromonas hydrophila. *Infect Immun, 73*(10), 6446-6457. doi:10.1128/IAI.73.10.6446-6457.2005

Sha, J., Rosenzweig, J. A., Kozlova, E. V., Wang, S., Erova, T. E., Kirtley, M. L., Van Lier, C.J., Chopra, A. K. (2013). Evaluation of the roles played by Hcp and VgrG type 6 secretion system effectors in Aeromonas hydrophila SSU pathogenesis. *Microbiology, 159*(6), 1120-1135. doi:doi:10.1099/mic.0.063495-0

Shao, S., Lai, Q., Liu, Q., Wu, H., Xiao, J., Shao, Z., Wang, Q., Zhang, Y. (2015). Phylogenomics characterization of a highly virulent Edwardsiella strain ET080813(T) encoding two distinct T3SS and three T6SS gene clusters: Propose a novel species as Edwardsiella anguillarum sp. nov. *Syst Appl Microbiol, 38*(1), 36-47. doi:10.1016/j.syapm.2014.10.008

Siguier, P., Gourbeyre, E., & Chandler, M. (2014). Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol Rev, 38*(5), 865-891. doi:10.1111/1574-6976.12067

Silverman, J. M., Brunet, Y. R., Cascales, E., & Mougous, J. D. (2012). Structure and regulation of the type VI secretion system. *Annu Rev Microbiol, 66*, 453-472. doi:10.1146/annurev-micro-121809-151619

Sokurenko, E. V., Hasty, D. L., & Dykhuizen, D. E. (1999). Pathoadaptive mutations: gene loss and variation in bacterial pathogens. *Trends Microbiol, 7*(5), 191-195.

Sreedharan, K., Philip, R., & Singh, I. S. (2011). Isolation and characterization of virulent Aeromonas veronii from ascitic fluid of oscar Astronotus ocellatus showing signs of infectious dropsy. *Dis Aquat Organ, 94*(1), 29-39. doi:10.3354/dao02304

Suarez, G., Khajanchi, B. K., Sierra, J. C., Erova, T. E., Sha, J., & Chopra, A. K. (2012). Actin cross-linking domain of Aeromonas hydrophila repeat in toxin A (RtxA) induces host cell rounding and apoptosis. *Gene, 506*(2), 369-376. doi:10.1016/j.gene.2012.07.012

Tabei, S. M., Hitchen, P. G., Day-Williams, M. J., Merino, S., Vart, R., Pang, P. C., Horsburgh, G.J., Viches, S., Wilhelms, M., Tomás, J.M., Dell, A., Shaw, J. G. (2009). An Aeromonas caviae genomic island is required for both O-antigen lipopolysaccharide biosynthesis and flagellin glycosylation. *J Bacteriol, 191*(8), 2851-2863. doi:10.1128/JB.01406-08

Takao, A., Nagamune, H., & Maeda, N. (2010). Sialidase of Streptococcus intermedius: a putative virulence factor modifying sugar chains. *Microbiol Immunol, 54*(10), 584-595. doi:10.1111/j.1348-0421.2010.00257.x

Tanaka, K. H., Dallaire-Dufresne, S., Daher, R. K., Frenette, M., & Charette, S. J. (2012). An insertion sequence-dependent plasmid rearrangement in Aeromonas salmonicida causes the loss of the type three secretion system. *PLoS One, 7*(3), e33725. doi:10.1371/journal.pone.0033725

Tekedar, H. C., Karsi, A., Akgul, A., Kalindamar, S., Waldbieser, G. C., Sonstegard, T., Schroeder, S.G., Lawrence, M. L. (2015). Complete Genome Sequence of Fish Pathogen Aeromonas hydrophila AL06-06. *Genome Announc, 3*(2). doi:10.1128/genomeA.00368-15

Tekedar, H. C., Karsi, A., Williams, M. L., Vamenta, S., Banes, M. M., Duke, M., Scheffler, B. E., Lawrence, M. L. (2013). Complete Genome Sequence of Channel Catfish Gastrointestinal Septicemia Isolate Edwardsiella tarda C07-087. *Genome Announc, 1*(6). doi:10.1128/genomeA.00959-13

Tekedar, H. C., Kumru, S., Kalindamar, S., Karsi, A., Waldbieser, G. C., Sonstegard, T., Schroeder, S. G., Liles, M. R., Griffin, M. J., Lawrence, M. L. (2017). Draft Genome Sequences of Three Aeromonas hydrophila Isolates from Catfish and Tilapia. *Genome Announc, 5*(3). doi:10.1128/genomeA.01509-16

Tekedar, H. C., Kumru, S., Karsi, A., Waldbieser, G. C., Sonstegard, T., Schroeder, S. G., Liles, M. R., Griffin, M. J., Lawrence, M. L. (2016a). Draft Genome Sequence of Aeromonas hydrophila TN97-08. *Genome Announc, 4*(3). doi:10.1128/genomeA.00436-16

Tekedar, H. C., Kumru, S., Karsi, A., Waldbieser, G. C., Sonstegard, T., Schroeder, S. G., Liles, M. R., Griffin, M. J., Lawrence, M. L. (2016b). Draft Genome Sequences of Four Virulent Aeromonas hydrophila Strains from Catfish Aquaculture. *Genome Announc, 4*(4). doi:10.1128/genomeA.00860-16

Tekedar, H. C., Waldbieser, G. C., Karsi, A., Liles, M. R., Griffin, M. J., Vamenta, S., Sonstegard, T., Hossain, M., Schroeder, S, Khoo, L., Lawrence, M. L. (2013). Complete Genome Sequence of a Channel Catfish Epidemic Isolate, Aeromonas hydrophila Strain ML09-119. *Genome Announc, 1*(5). doi:10.1128/genomeA.00755-13

Terabayashi, Y., Juan, A., Tamotsu, H., Ashimine, N., Nakano, K., Shimoji, M., Shiroma, A., Teruya, K., Satou, K., Hirano, T. (2014). First Complete Genome Sequence of Salmonella enterica subsp. enterica Serovar Typhimurium Strain ATCC 13311 (NCTC 74), a Reference Strain of Multidrug Resistance, as Achieved by Use of PacBio Single-Molecule Real-Time Technology. *Genome Announc, 2*(5). doi:10.1128/genomeA.00986-14

Touchon, M., & Rocha, E. P. (2007). Causes of insertion sequences abundance in prokaryotic genomes. *Mol Biol Evol, 24*(4), 969-981. doi:10.1093/molbev/msm014

Trudel, M. V., Tanaka, K. H., Filion, G., Daher, R. K., Frenette, M., & Charette, S. J. (2013). Insertion sequence AS5 (IS AS5) is involved in the genomic plasticity of Aeromonas salmonicida. *Mob Genet Elements, 3*(3), e25640. doi:10.4161/mge.25640

Uzun, E., & Ogut, H. (2015). The isolation frequency of bacterial pathogens from sea bass (Dicentrarchus labrax) in the Southeastern Black Sea. *Aquaculture, 437*(1), 30-37.

van Soest, J. J., Stockhammer, O. W., Ordas, A., Bloemberg, G. V., Spaink, H. P., & Meijer, A. H. (2011). Comparison of static immersion and intravenous injection systems for exposure of zebrafish embryos to the natural pathogen Edwardsiella tarda. *BMC Immunol, 12*, 58. doi:10.1186/1471-2172-12-58

Vincent, A. T., Trudel, M. V., Freschi, L., Nagar, V., Gagne-Thivierge, C., Levesque, R. C., & Charette, S. J. (2016). Increasing genomic diversity and evidence of constrained lifestyle evolution due to insertion sequences in Aeromonas salmonicida. *BMC Genomics, 17*, 44. doi:10.1186/s12864-016-2381-3

Vine, N. G., Leukes, W. D., Kaiser, H., Daya, S., Baxter, J., & Hecht, T. (2004). Competition for attachment of aquaculture candidate probiotic and pathogenic bacteria on fish intestinal mucus. *J Fish Dis, 27*(6), 319-326. doi:10.1111/j.1365-2761.2004.00542.x

Wallden, K., Rivera-Calzada, A., & Waksman, G. (2010). Type IV secretion systems: versatility and diversity in function. *Cell Microbiol, 12*(9), 1203-1212. doi:10.1111/j.1462-5822.2010.01499.x

Wang, Q., Yang, M., Xiao, J., Wu, H., Wang, X., Lv, Y., Xu, L., Zheng, H., Wang, S., Zhao, G., Liu, Q., Zhang, Y. (2009). Genome sequence of the versatile fish pathogen Edwardsiella tarda provides insights into its adaptation to broad host ranges and intracellular niches. *PLoS One, 4*(10), e7646. doi:10.1371/journal.pone.0007646

Wilms, I., Moller, P., Stock, A. M., Gurski, R., Lai, E. M., & Narberhaus, F. (2012). Hfq influences multiple transport systems and virulence in the plant pathogen Agrobacterium tumefaciens. *J Bacteriol, 194*(19), 5209-5217. doi:10.1128/JB.00510-12

Woo, P. T. K., & Bruno, D. W. (2011). *Fish Diseases and Disorders* (P. T. K. Woo & D. W. Bruno Eds. 2 ed. Vol. Viral, Bacterial and Fungal Infections): CAB International.

Wu, L., Jiang, Y. N., Tang, Q., Lin, H. X., Lu, C. P., & Yao, H. C. (2012). Development of an Aeromonas hydrophila recombinant extracellular protease vaccine. *Microb Pathog, 53*(5-6), 183-188. doi:10.1016/j.micpath.2012.07.007

Yadav, S. K., Sahoo, P. K., & Dixit, A. (2014). Characterization of immune response elicited by the recombinant outer membrane protein OmpF of Aeromonas hydrophila, a potential vaccine candidate in murine model. *Mol Biol Rep, 41*(3), 1837-1848. doi:10.1007/s11033-014-3033-9

Yin, K., Wang, Q., Xiao, J., & Zhang, Y. (2017). Comparative proteomic analysis unravels a role for EsrB in the regulation of reactive oxygen species stress responses in Edwardsiella piscicida. *FEMS Microbiol Lett, 364*(1). doi:10.1093/femsle/fnw269

Yu, H. B., Kaur, R., Lim, S., Wang, X. H., & Leung, K. Y. (2007). Characterization of extracellular proteins produced by Aeromonas hydrophila AH-1. *Proteomics, 7*(3), 436-449. doi:10.1002/pmic.200600396

Zepeda-Velazquez, A. P., Vega-Sanchez, V., Salgado-Miranda, C., & Soriano-Vargas, E. (2015). Histopathological findings in farmed rainbow trout (Oncorhynchus mykiss) naturally infected with 3 different Aeromonas species. *Can J Vet Res, 79*(3), 250-254.

Zhang, D., Pridgeon, J. W., & Klesius, P. H. (2014). Vaccination of channel catfish with extracellular products of Aeromonas hydrophila provides protection against infection by the pathogen. *Fish Shellfish Immunol, 36*(1), 270-275. doi:10.1016/j.fsi.2013.11.015

Zhang, D., Xu, D., & Shoemaker, C. A. (2015). Immunization with recombinant aerolysin and haemolysin protected channel catfish against virulent Aeromonas hydrophila. *Aquaculture Research*(1-8).

Zhang, G., Li, C., Li, Q., Li, B., Larkin, D. M., Lee, C., Wang, J. and et al. (2014). Comparative genomics reveals insights into avian genome evolution and adaptation. *Science, 346*(6215), 1311-1320. doi:10.1126/science.1251385

Zheng, J., Tung, S. L., & Leung, K. Y. (2005). Regulation of a type III and a putative secretion system in Edwardsiella tarda by EsrC is under the control of a two-component system, EsrA-EsrB. *Infect Immun, 73*(7), 4127-4137. doi:10.1128/IAI.73.7.4127-4137.2005

Zoued, A., Brunet, Y. R., Durand, E., Aschtgen, M. S., Logger, L., Douzi, B., Journeta, L., Cambillaub, C., Cascales, E. (2014). Architecture and assembly of the Type VI secretion system. *Biochim Biophys Acta, 1843*(8), 1664-1673. doi:10.1016/j.bbamcr.2014.03.018

CHAPTER II

COMPLETE AND DRAFT GENOME SEQUENCES OF *EDWARDSIELLA PISCICIDA* AND *AEROMONAS HYDROPHILA* ISOLATES RECOVERED FROM OUTBREAKS

2.1 *Edwardsiella piscicida*

Edwardsiella piscicida is the etiologic agent of acute to chronic edwardsiellosis in fish and other species (Evans JJ et al., 2011). It is a Gram-negative facultative anaerobe that is motile by peritrichous flagella. Edwardsiellosis is an important fish disease that negatively impacts aquaculture industries throughout the world (Park et al., 2012). Moreover, *E. piscicida* can cause disease in humans, where it is typically associated with gastrointestinal infections that may progress to hemorrhagic septicemia (Lv et al., 2013). Historically, *E. piscicida* causes a sporadic disease in channel catfish (*Ictalurus punctatus*) called emphysematous putrefactive disease, which may begin as small punctate cutaneous erosions that progress to abscesses deep within the musculature (Evans JJ et al., 2011). However, *E. piscicida* is being increasingly isolated from large outbreaks of gastrointestinal septicemia in commercial channel catfish operations. *E. piscicida* isolate C07-087 is from an outbreak of gastrointestinal septicemia in a commercial catfish pond in Mississippi.

2.2 *Aeromonas hydrophila*

Aeromonas species are Gram-negative facultative anaerobes that are ubiquitous in aquatic environments and cause infections in several host species, including humans, invertebrates, reptiles, and amphibians (Chao et al., 2013; Huys et al., 2003; Noonin et al., 2010; Orozova et al., 2012; Saejung et al., 2011). In particular, many of the *Aeromonas* species are pathogenic to fish, causing septicemia in carp, tilapia, perch, salmon, catfish, and other species (Janda et al., 2010). In channel catfish aquaculture, *Aeromonas hydrophila* is historically considered an opportunistic pathogen. However, since 2009 a clonal group of virulent *A. hydrophila* (vAh) isolates have been causing large-scale disease outbreaks in Alabama and Mississippi (Griffin et al., 2013). Strain ML09-119 is an isolate from a disease outbreak on a commercial catfish farm, and it is representative of this clonal group. We reported the complete genome sequence of *A. hydrophila* ML09-119 (Tekedar et al., 2013). For comparative purposes, we next reported the complete genome sequence of *A. hydrophila* AL06-06 (Tekedar et al., 2015), which was isolated from a diseased goldfish in 2006 from the Auburn University Southeastern Cooperative Fish Disease Laboratory in Greensboro, Alabama. We subsequently reported the draft genome sequence of strain TN97-08, which was isolated in 1997 from diseased bluegill (*Lepomis macrochirus*) (Tekedar et al., 2016a). We have also reported four draft genomes of vAh strains isolated from farm-raised catfish in 2009 and 2010 (strains AL10-121, AL09-79, ML09-121, and ML09-122) (Tekedar et al., 2016b). Lastly, we reported the draft genomes of three additional *A. hydrophila* genomes for comparison purposes, one isolated from diseased catfish (strain Arkansas 2010) and two isolated from diseased tilapia (strains AL97-91 and MN98-04) (Tekedar et al., 2017).

Therefore, we have released one complete genome (strain ML09-119) (Tekedar, Waldbieser, et al., 2013) and four draft genomes (strains AL10-121, AL09-79, ML09-121, and ML09-122) that represent this clonal group of vAh affecting catfish (Tekedar et al., 2016b). For comparison, we have also released one complete genome (strain AL06-06, goldfish isolate) (Tekedar et al., 2015) and four draft genomes (TN97-08, bluegill isolate; Arkansas 2010, catfish isolate; and AL97-91 and MN98-04, tilapia isolates) of strains representing various *A. hydrophila* clades (Tekedar et al., 2016a). Comparative genomics of these strains will enable a better understanding of the variation in virulence genes and antigenic structures of fish-pathogenic *A. hydrophila*.

2.3 Material and methods

2.3.1 Sequencing

2.3.1.1 *Edwardsiella piscicida* C07-087

To determine the circularized *E. piscicida* C07-087 genome sequence, genomic DNA was shotgun sequenced using two different methods: Illumina genome analyzer IIx (8,144,858 reads with 76x coverage) (Illumina, Inc., San Diego, CA) and 454 GS-FLX Titanium platform (312,969 reads with 180x coverage) (Roche Applied Science, Indianapolis, IN, USA). CLC Workbench 5.0.1 (CLC bio, Cambridge, MA) and Sequencher 5.1 (Gene Codes Corporation, Ann Arbor, MI) were used to trim sequences, and reads were de novo assembled using the CLC Workbench. Unscaffolded gaps were amplified and sequenced by single-primer PCR (Karlyshev et al., 2000). rRNA gene operons were amplified and sequenced to resolve misassemblies.

2.3.1.2 *Aeromonas hydrophila* ML09-119

The genome sequence of *Aeromonas hydrophila* ML09-119 was completed using a combination of Illumina Genome Analyzer IIx next-generation sequencing (a total of 4,077,018 reads with 104x coverage) (Hossain et al., 2013) and the 454 GS-FLX titanium platform (a total of 96,601 reads with 308x coverage). Sequences from both platforms were assessed for errors and trimmed for quality using CLC workbench 5.0.1 and Sequencher 5.1 (Gene Codes Corporation). Assembly was performed by CLC workbench 5.0.1. Scaffolded gaps were closed by Sanger sequencing of PCR amplicons. Unscaffolded gaps were closed by sequencing single-primer PCR amplicons (Karlyshev et al., 2000). rRNA operons and other repeat regions were amplified and sequenced to resolve misassemblies.

2.3.1.3 *Aeromonas hydrophila* AL06-06

The genome of *A. hydrophila* AL06-06 was fully sequenced using an Illumina Genome Analyzer IIx (a total of 6,629,874 reads with 150x coverage). Trimming, error correction, contig creation, and quality control of sequence reads were conducted using CLC Workbench 6.5.1 and Sequencher 5.2.3. De novo assembly was performed by CLC Workbench 6.5.1. Scaffolded gaps were closed by Sanger sequencing of PCR amplicons. For the unscaffolded gaps, single-primer PCR was used for amplification of sequence templates (Karlyshev et al., 2000). Ribosomal operons and other repetitive regions were amplified and completely resequenced to create a reliable assembly.

2.3.1.4 *Aeromonas hydrophila* TN97-08

The *A. hydrophila* TN97-08 genome was sequenced using an Illumina Genome Analyzer IIx (9,280,112 reads with 243x coverage). Adaptor trimming, contig creation, and quality control of sequence reads were conducted using CLC Workbench 6.5.1 and Sequencher 5.3. De novo assembly was performed by CLC Workbench 6.5.1. To reduce contig numbers, some of the scaffolded gaps were closed by Sanger sequencing of PCR amplicons. Additionally, some of the unscaffolded gaps were closed by single-primer PCR (Karlyshev et al., 2000).

2.3.1.5 *Aeromonas hydrophila* AL10-121, AL09-79, ML09-121, ML09-122

A. hydrophila AL10-121, AL09-79, ML09-121, and ML09-122 genomes were sequenced using an Illumina Genome Analyzer IIx. The total number of reads and fold genome coverages were as follows: 8,107,077 reads with 224x coverage (AL10-121); 7,613,846 reads with 206x coverage (AL09-79); 7,094,460 reads with 192x coverage (ML09-121); and 6,199,052 reads with 176x coverage (ML09-122). Adaptor trimming, contig creation, and quality control of sequence reads were conducted using CLC Workbench version 6.5.1 and Sequencher version 5.4. De novo assembly was performed by CLC Workbench version 6.5.1.

2.3.1.6 *Aeromonas hydrophila* Arkansas 2010, AL97-91, MN98-04

A. hydrophila Arkansas 2010, AL97-91, and MN98-04 were sequenced using an Illumina Genome Analyzer IIx (11,143,909 reads with 301x coverage, 6,018,377 reads with 168x coverage, and 7,578,657 reads with 211x coverage, respectively). Read

trimming, error correction, and contig creation were conducted using CLC Genomics Workbench version 6.5.1 and Sequencher version 5.4.5.

2.3.2 Plasmid sequencing

Native plasmid DNA was isolated from the strains using QIAprep Spin Miniprep Kit (QIAGEN, Maryland, USA). Plasmids were separated on 1% agarose gel, and size of the plasmids was estimated with concurrently run standards (Supercoiled DNA ladder, New England Biolabs, Ipswitch, MA, USA). Isolated plasmids were sequenced commercially using the complete plasmid sequencing service of the DNA Core Facility of the Center for Computational and Integrative Biology at Massachusetts General Hospital (Boston, MA, USA), and acquired sequences from this service were assembled by MGH CCIB NGS de novo assembler UltraCycler v1.0 (Seed and Wang Unpublished).

2.3.3 Genome annotation and relatedness

The circularized and completed genome of *E. piscicida* strain C07- 087 and *A. hydrophila* strains were submitted to two different annotation pipelines. The NCBI Prokaryotic Genome Automatic Annotation Pipeline (PGAAP) (Angiuoli et al., 2008) was used for annotation and submission to GenBank, and Rapid Annotations using Subsystems Technology (RAST) (Aziz et al., 2008) with the Glimmer option was used to facilitate comparative analysis of the *E. piscicida* and *A. hydrophila* genomes. Furthermore, publicly available complete and draft *E. piscicida* genomes (Table 2.1) and *A. hydrophila* genomes (Table 2.2) were downloaded from NCBI, and average nucleotide identities (ANI) were determined by EDGAR (Figures 2.1 and 2.3) (Goris et al., 2007; Konstantinidis et al., 2006; Konstantinidis et al., 2005). BLAST Ring Image Generator

(BRIG) (Alikhan et al., 2011) was used to compare *E. piscicida* and *A. hydrophila* genomes (Figures 2.2 and 2.4). To accomplish this, nucleotide files were downloaded from NCBI, and BRIG analysis was performed using default criteria: upper identity threshold (%) 70, lower identity threshold (%) 50, and ring size 30; ring height and weight were 2250 pixels.

Table 2.1 Features of *Edwardsiella piscicida* genomes used for analysis.

These strains are classified as *E. tarda* in NCBI, but ANI results show they are *E. piscicida.*

#	Species	Strain	Location	Source	Genome close level/contig numbers	Size (Mb)	GC %	Plasmid	Accession	Reference
1	*E. tarda**	ET883	Norway	European eel	Contig/167	3.82967	59.60		PRJNA259400	(Abayneh, et al., 2012)
2	*E. tarda**	FL6-60	USA	Striped bass	Complete	3.7288	59.70	1	NC_017309	(van Soest et al., 2011)
3	*E. piscicida*	ETW41	Korea	Eel pond water	Chromosome	3.96904	59.48	1	NZ_CP019440	N/A
4	*E. piscicida*	JF1305	Japan	Olive Flounder	Contig/72	3.71682	59.80		PRJDB1727	(Oguro et al., 2014)
5	*E. tarda**	EIB202	China	Turbot	Complete	3.80417	59.67	1	NC_013508	(Wang et al., 2009)
6	*E. piscicida*	C07-087	USA	Channel catfish	Complete	3.85704	59.60		NC_020796	(Tekedar et al., 2013)
7	*E. piscicida*	ACC35.1	Europe	Turbot	Scaffold	3.84163	59.70		PRJNA353909	(Bujan et al., 2017)
8	*E. piscicida*	S11-285	USA	Channel catfish	Complete	3.92677	59.59	1	NZ_CP016044	(Reichley et al., 2016)

2.4 Results

2.4.1 *Edwardsiella piscicida*

The *E. piscicida* genome consists of one circular chromosome with 3,857,040 bp and 59.6% G+C content. PGAAP annotation predicted 3,525 genes coding for 3,405 proteins. tRNAscan-SE (Lagesen et al., 2007) and RNAmmer (Lowe et al., 1997) predicted 95 tRNAs and 8 rRNA operons, respectively. Griffin et al. (Griffin et al., 2013) showed that *E. piscicida* consists of two genetic subtypes, DNA groups I and II. Based on sequence comparison and G+C content, C07-087 is in DNA group II. Two *E. piscicida* genome sequences from DNA group II were previously reported: that of EIB202, which was isolated from diseased turbot in China (Wang et al., 2009), and FL6-60, which was isolated from catfish (Pressley et al., 2005). Plasmid preparations showed that *E. piscicida* C07-087 does not carry any plasmids. However, strain C07-087 has 31,387 bp in its chromosome that is identical to part of a 44,194-bp plasmid isolated from *E. piscicida* FL6-60. Integration of this partial plasmid sequence into the *E. piscicida* strain C07-087 chromosome was confirmed by PCR.

2.4.2 *Aeromonas hydrophila*

General features of the complete and draft *A. hydrophila* genomes (*A. hydrophila* ML09-119, *A. hydrophila* AL09-79, *A. hydrophila* Arkansas 2010, *A. hydrophila* ML09-122, *A. hydrophila* AL10-121, *A. hydrophila* ML09-121, *A. hydrophila* MN98-04, *A. hydrophila* AL97-91, *A. hydrophila* AL06-06, and *A. hydrophila* TN97-08) are summarized in Table 2. Results indicate that all the complete or draft vAh genomes are quite similar in terms of genome size, G+C content, number of coding genes, and tRNA number.

Table 2.2 Features of *Aeromonas hydrophila* genomes used in this analysis

#	Species	Strain	Location	Source	Genome close level/contig numbers	Size (Mb)	GC %	Plasmid	Accession
1	*A. hydrophila*	Arkansas 2010	USA	Catfish	NZ_LYZH00000000.1	4.97356	60.90	-	(Tekedar et al., 2017)
2	*A. hydrophila*	ML09-119	USA	Catfish	NC_021290.1	5.0245	60.80	-	(Tekedar, Waldbieser, et al., 2013)
3	*A. hydrophila*	ML09-122	USA	Catfish	NZ_LRRY00000000.1	4.96999	60.90	-	(Tekedar et al., 2016b)
4	*A. hydrophila*	ML09-121	USA	Catfish	NZ_LRRX00000000.1	4.96594	60.90	-	(Tekedar et al., 2016b)
5	*A. hydrophila*	AL10-121	USA	Catfish	NZ_LRRW00000000.1	4.96991	60.90	-	(Tekedar et al., 2016b)
6	*A. hydrophila*	TN-97-08	USA	Bluegill	NZ_LNUR00000000.1	5.08731	60.80	-	(Tekedar et al., 2016a)
7	*A. hydrophila*	AL09-79	USA	Catfish	NZ_LRRV00000000.1	4.96786	60.90	-	(Tekedar et al., 2016b)
8	*A. hydrophila*	AL06-06	USA	Goldfish	NZ_CP010947.1	4.90139	61.37	3	(Tekedar et al., 2015)
9	*A. hydrophila*	AL97-91	USA	Tilapia	NZ_CM004591.1	4.83027	61.19	1	(Tekedar et al., 2017)
10	*A. hydrophila*	MN98-04	USA	Tilapia	NZ_CM004592.1	4.88294	61.10	1	(Tekedar et al., 2017)

The final closed-circle version of the *A. hydrophila* ML09-119 genome sequence was submitted to PGAAP (Angiuoli et al., 2008) for annotation, followed by submission to GenBank. The total *A. hydrophila* genome comprises 5,024,500 bp with 60.8% GC content. It contains 4,577 predicted genes, of which 4,434 are protein-coding sequences. A total of 112 tRNAs and 10 rRNA operons were predicted by using tRNAscan-SE (Lowe et al., 1997) and RNAmmer 1.2 (Lagesen et al., 2007), respectively. The ML09-119 reads were assembled against the *A. hydrophila* ATCC 7966T genome (NC_008570.1) in CLC workbench 5.0.1 to identify contiguous regions of the ML09-119 genome that are not present in the ATCC 7966T genome. Functional analysis of

predicted open reading frames (ORFs) in these unique contigs indicated that strain ML09-119 has a complete inositol utilization pathway that is not present in ATCC 7966T. More than 20 unique prophage-linked ORFs and several transposons were identified, and several putative virulence loci appear to be linked to prophage elements. Relative to strain ATCC 7966T, ML09-119 contains a unique 33-kb O polysaccharide biosynthesis gene cluster with 29 total predicted ORFs. Twenty-four of these do not have any similarity to ATCC 7966T genes. Thus, it appears that ML09-119 has a different O antigen serotype than ATCC 7966T. In summary, the *A. hydrophila* ML09-119 genome encodes putative proteins suggesting that it has unique biochemical and serological features relative to strain ATCC 7966T.

The complete genome of *A. hydrophila* AL06-06 comprises 4,884,823 bp with 61.3% G+C content. It contains 4,453 predicted genes, of which 4,235 are protein coding. A total of 112 tRNAs and 10 rRNA operons were predicted by PGAAP. The *A. hydrophila* AL06-06 genome was also annotated by RAST (Overbeek et al., 2014) to facilitate comparison with the *A. hydrophila* ML09- 119 genome. Based on functional comparative results, *A. hydrophila* AL06-06 encodes 81 unique proteins including arsenic resistance proteins, heme and hemin uptake-utilization systems, some membrane transport proteins for type I and type V secretion systems, transposable elements, and nitrogen metabolism proteins. In particular, strain AL06-06 has a specific arsenic resistance mechanism that is missing in the *A. hydrophila* ML09-119 genome. Due to its ubiquitous distribution in the environment, *A. hydrophila* is prone to arsenic exposure (Goswami et al., 2011). The *A. hydrophila* AL06-06 genome also has three plasmids compared to strain ML09-119, which does not carry any plasmids.

The draft genome of *A. hydrophila* TN97-08 is composed of 16 contigs and 5,087,310 bp. It contains 4,602 predicted genes, of which 4,445 are protein coding genes. The annotated *A. hydrophila* TN97-08 genome was compared against the genomes of *A. hydrophila* ML09-119 and AL06-06. The results indicated that *A. hydrophila* TN97-08 has 100 and 74 unique proteins compared to strains ML09-119 and AL06-06, respectively. One of the unique features identified is a type VI secretion system, which is considered a virulence factor involved in the translocation of potential effector proteins into host cells (Bingle et al., 2008). Other unique elements include multidrug resistance efflux pumps, cobalt zinc cadmium resistance, stress response proteins, phage and prophage elements, and toxin-antitoxin replicon stabilization system proteins. In contrast to the AL06-06 genome, TN97-08 does not carry any plasmids. The estimated average nucleotide identity (ANI) (http://enve-omics.ce.gatech.edu/ani/) between the strain TN97-08 genome and the strain ML09-199 genome was 96.61%, and ANI was 96.72% between TN97-08 and AL06-06 (Goris et al., 2007).

The ANI between each of the four draft vAh genomes and the complete genome of vAh strain ML09-119 was 99.99% (Goris et al., 2007). By comparison, the ANI between *A. hydrophila* strain AL06-06 and the four vAh draft genomes was 96.6%. Functional comparison of the annotated vAh draft genomes against *A. hydrophila* ML09-119 showed that the vAh strains share the same functional elements except a few metabolic genes in glycine and serine utilization as well as cysteine, biotin, and molybdenum biosynthesis. On the other hand, the vAh strains all encode several pathways that are missing in other *A. hydrophila* strains such as AL06-06. For example, vAh strains carry an inositol catabolism pathway, phage elements, and an RTX toxin

cluster that are missing in *A. hydrophila* strain AL06-06. In contrast to the AL06-06 genome, none of the vAh strain genomes carry plasmid.

The draft genomes from Arkansas 2010, AL97-91, and MN98-04 and their plasmids were submitted to the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (Angiuoli et al., 2008) for annotation. The draft genomes were submitted to the (RAST) server for additional annotation and analyses (Aziz et al., 2008; Overbeek et al., 2014). The ANI mean against the genomes of previously sequenced strains ML09-119 and AL06-06 were calculated using EDGAR (Blom et al., 2009). Interestingly, the ANI mean between the genomes of strain Arkansas 2010 and vAh strain ML09-119 indicate that Arkansas 2010 is a member of the vAh clonal group. Strain Arkansas 2010 resulted from vAh-infected fish being transported from Alabama to Arkansas. Compared to previously published *A. hydrophila* genomes (Tekedar et al., 2015; Tekedar et al., 2013), the genomes of strains Arkansas 2010, AL97-91, and MN98-04 carry toxin-antitoxin replicon stabilization system components. Additionally, the AL97-91 and MN98-04 genomes are unique in that these strains are capable of utilizing taurine.

2.5 Discussion

In the current study, we applied genome sequencing to two different bacterial pathogens affecting U.S. aquaculture to accelerate research on their pathogenic mechanisms. *Edwardsiella* was identified in 1965 as a new genus in the *Enterobacteriaceae* (Ewing et al., 1965). Three *Edwardsiella* species have since been added to this genus: *Edwardsiella tarda* in 1973 (Meyer et al., 1973), *Edwardsiella hoshinae* in 1980 (Grimont et al., 1980), and *Edwardsiella ictaluri* in 1981 (Hawke et al., 1981). We originally reported *E. piscicida* C07-087 as *E. tarda*, but recently two new

Edwardsiella species were defined, *Edwardsiella anguillarum* (Shao et al., 2015) and *Edwardsiella piscicida* (Abayneh et al., 2013). This clarification allowed us to clearly place strain CO7-087 in the correct taxonomic designation of *E. piscicida*.

We evaluated the relatedness of all the available *Edwardsiella tarda* and *piscicida* genomes (Table 1) by applying ANI, which has become the preferred tool to estimate genome differences (Kim et al., 2014). Strains ET883, FL6-60, and EIB202, although classified as *E. tarda* in NCBI, are actually *E. piscicida* based on ANI values over 99% (strains with >95% ANI are classified as the same species); thus, we labeled these strains as *E. piscicida* in Figures 2.1 and 2.2 It is interesting that these *E. piscicida* strains share high sequence identity, yet they are from diverse host species and geographic regions. These findings suggest ANI calculation is a reliable method to classify bacterial genomes.

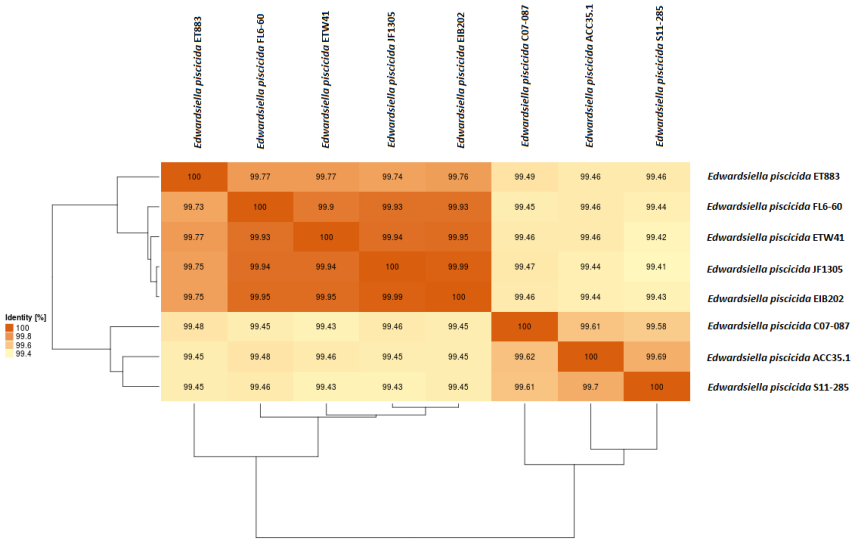


Figure 2.1 Average nucleotide identities (ANI) for *E. piscicida* genomes (analysis conducted on 9/13/2017).

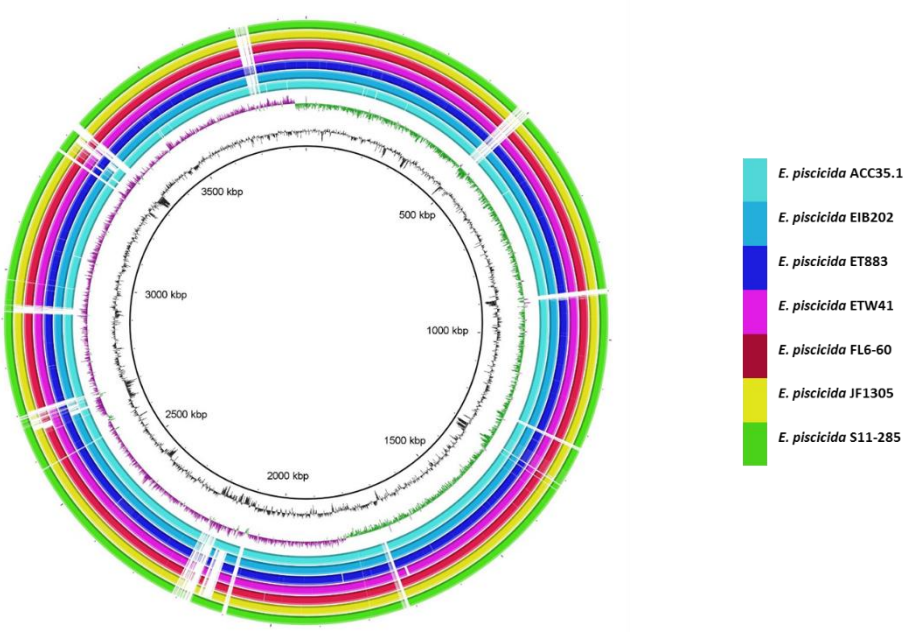


Figure 2.2 Sequence identity of *E. piscicida* genomes using *E. piscicida* C07-87 as the reference.

Image was created using the BLAST Ring Image Generator (BRIG). Each ring represents a query genome, and each genome is colored to indicate the presence of orthologous proteins to the reference genome (*Edwardsiella piscicida* C07-087). The order of rings from inner to outer is as follows: G+C content, G+C skew, *E. piscicida* ACC35.1, *E. piscicida* EIB202, *E. piscicida* ET883, *E. piscicida* ETW41, *E. piscicida* FL6-60, *E. piscicida* JF1305, and *E. piscicida* S11-285.

The fully sequenced circular *E. piscicida* C07-087 genome has unique features relative to previously sequenced *E. piscicida* genomes and will be useful for helping to delineate the pathogenesis of gastrointestinal septicemia caused by *E. piscicida* in channel catfish. Using BRIG analysis to compare it to other *E. piscicida* genomes, the majority of genome regions unique to *E. piscicida* C07-087 encode mobile elements.

Since 2009, vAh has been affecting the U.S. catfish aquaculture industry. However, the pathogenic mechanisms of this clade that cause increased virulence are largely unknown. To accelerate pathogenesis research on vAh, we conducted genome sequencing of 10 *A. hydrophila* strains (6 vAh strains and 4 *A. hydrophila* strains from other fish species). It was previously reported that many *Aeromonas* species were taxonomically labeled incorrectly in NCBI, and it was recommend to use ANI for determining taxonomic affiliations (Beaz-Hidalgo et al., 2015). Therefore, we determined ANI of our 10 genomes, and results showed that all of them are *A. hydrophila* (Figure 2.3). ANI also clearly showed that vAh isolates share very high sequence identity (Figure 2.3), while the other *A. hydrophila* genomes are more distantly related to this clonal group.

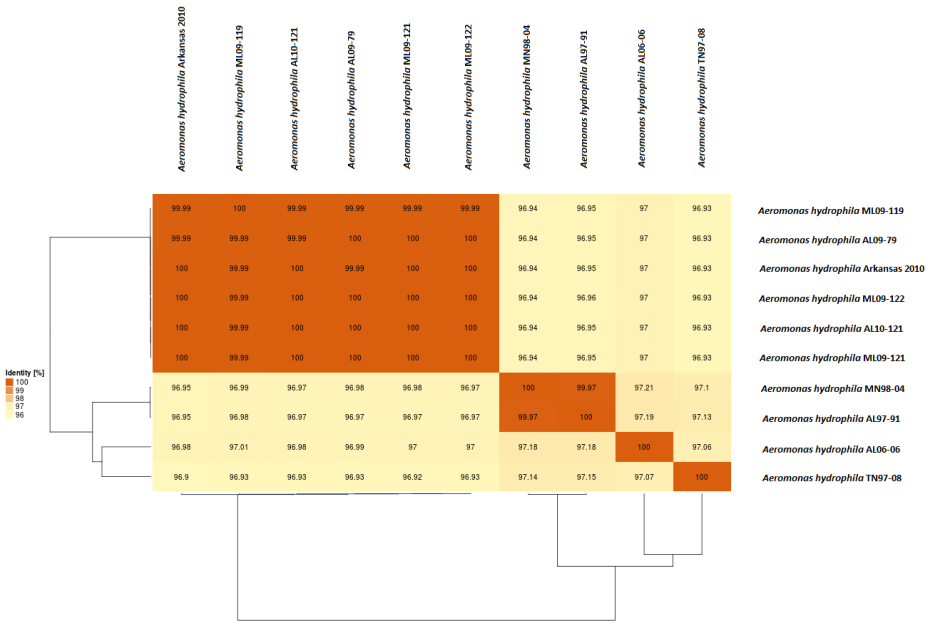


Figure 2.3 Average nucleotide identities (ANI) for *Aeromonas hydrophila* genomes (analysis conducted on 9/13/2017).

BRIG analysis of strain ML09-119 compared to the other *A. hydrophila* strains show that vAh strains have several unique genomic regions (Figure 2.4), suggesting that lateral gene transfer contributed to the virulent phenotype of vAh (Hossain et al., 2013). It has been reported that an Asian clonal group of virulent *A. hydrophila* was introduced to the U.S. catfish aquaculture industry and may be the source of vAh in the U.S. (Hossain et al., 2014).

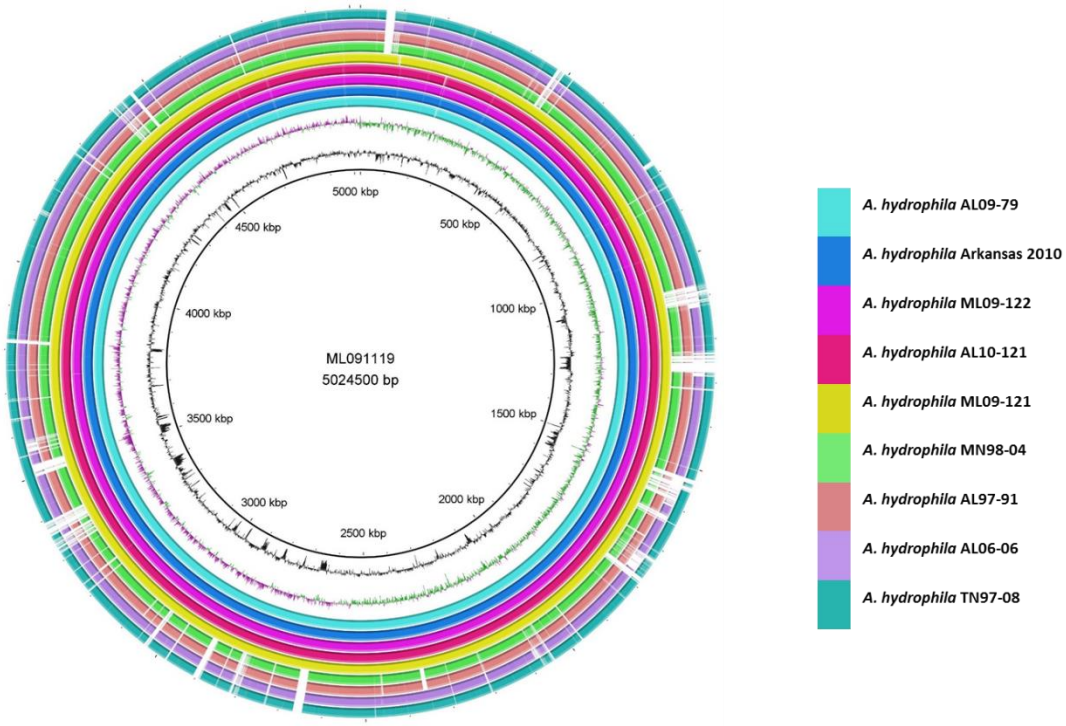


Figure 2.4 Sequence identity of *A. hydrophila* genomes using *A. hydrophila* ML09-119 as the reference.

Image was created using the BLAST Ring Image Generator (BRIG). Each ring represents a query genome, and each genome is colored to indicate the presence of orthologous proteins to the reference genome (*A. hydrophila* ML09-119). The order of rings from inner to outer is as follows: G+C content, G+C skew, *A. hydrophila* AL09-79, *A. hydrophila* Arkansas 2010, *A. hydrophila* ML09-122, *A. hydrophila* AL10-121, *A. hydrophila* ML09-121, *A. hydrophila* MN98-04, *A. hydrophila* AL97-91, *A. hydrophila* AL06-06, *A. hydrophila* TN97-08.

In summary, the complete and draft genomes of *E. piscicida*, vAh strains, and other strains of *A. hydrophila* will contribute to our knowledge of virulence and environmental adaptations in these important fish pathogens. Our findings are also useful for comparison between fish- and human-pathogenic *A. hydrophila* and *E. tarda* strains. We expect that unique characteristics uncovered in the genomes of vAh and *E. piscicida* will enable us to develop new control strategies to prevent these diseases impacting U.S. aquaculture.

2.6 References

Abayneh, T., Colquhoun, D. J., & Sorum, H. (2012). Multi-locus Sequence Analysis (MLSA) of Edwardsiella tarda isolates from fish. *Vet Microbiol, 158*(3-4), 367-375. doi:10.1016/j.vetmic.2012.03.006

Abayneh, T., Colquhoun, D. J., & Sorum, H. (2013). Edwardsiella piscicida sp. nov., a novel species pathogenic to fish. *J Appl Microbiol, 114*(3), 644-654. doi:10.1111/jam.12080

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

Angiuoli, S. V., Gussman, A., Klimke, W., Cochrane, G., Field, D., Garrity, G., Kodira, C.D., Kyrpides, N., Madupu, R., Markowitz, V., Tatusova, T., Thomson, N., White, O. (2008). Toward an online repository of Standard Operating Procedures (SOPs) for (meta)genomic annotation. *OMICS, 12*(2), 137-141. doi:10.1089/omi.2008.0017

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

Beaz-Hidalgo, R., Hossain, M. J., Liles, M. R., & Figueras, M. J. (2015). Strategies to avoid wrongly labelled genomes using as example the detected wrong taxonomic affiliation for aeromonas genomes in the GenBank database. *PLoS One, 10*(1), e0115813. doi:10.1371/journal.pone.0115813

Bingle, L. E., Bailey, C. M., & Pallen, M. J. (2008). Type VI secretion: a beginner's guide. *Curr Opin Microbiol, 11*(1), 3-8. doi:10.1016/j.mib.2008.01.006

Blom, J., Albaum, S. P., Doppmeier, D., Puhler, A., Vorholter, F. J., Zakrzewski, M., & Goesmann, A. (2009). EDGAR: a software framework for the comparative analysis of prokaryotic genomes. *BMC Bioinformatics, 10*, 154. doi:10.1186/1471-2105-10-154

Bujan, N., Toranzo, A. E., & Magarinos, B. (2017). Draft Genome Sequence of Edwardsiellapiscicida Strain ACC35.1 Isolated from Diseased Turbot (Scophthalmus maximus) in Europe. *Genome Announc, 5*(7). doi:10.1128/genomeA.01626-16

Chao, C. M., Lai, C. C., Tang, H. J., Ko, W. C., & Hsueh, P. R. (2013). Skin and soft-tissue infections caused by Aeromonas species. *Eur J Clin Microbiol Infect Dis, 32*(4), 543-547. doi:10.1007/s10096-012-1771-y

Evans JJ, Klesius PH, Plumb JA, & CA, S. (2011). *Fish diseases and disorders* (Woo PTK & B. DW Eds.). Oxfordshire, United Kingdom: CAB International.

Ewing, W. H., Mcwhorter, A. C., Escobar, M. R., & Lubin, A. H. (1965). Edwardsiella, a new genus of Enterobacteriaceae based on a new species, E. tarda. *International Journal of Systematic and Evolutionary Microbiology, 15*(1), 33-38. doi:doi:10.1099/00207713-15-1-33

Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., & Tiedje, J. M. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol, 57*(Pt 1), 81-91. doi:10.1099/ijs.0.64483-0

Goswami, R., Ghosh, D., Saha, D. R., Padhy, P. K., & Mazumder, S. (2011). Effect of acute and chronic arsenic exposure on growth, structure and virulence of Aeromonas hydrophila isolated from fish. *Microb Pathog, 50*(2), 63-69. doi:10.1016/j.micpath.2010.10.005

Griffin, M. J., Goodwin, A. E., Merry, G. E., Liles, M. R., Williams, M. A., Ware, C., & Waldbieser, G. C. (2013). Rapid quantitative detection of Aeromonas hydrophila strains associated with disease outbreaks in catfish aquaculture. *J Vet Diagn Invest, 25*(4), 473-481. doi:10.1177/1040638713494210

Griffin, M. J., Quiniou, S. M., Cody, T., Tabuchi, M., Ware, C., Cipriano, R. C., Mauel, M.J., Soto, E. (2013). Comparative analysis of Edwardsiella isolates from fish in the eastern United States identifies two distinct genetic taxa amongst organisms phenotypically classified as E. tarda. *Vet Microbiol, 165*(3-4), 358-372. doi:10.1016/j.vetmic.2013.03.027

Grimont, P. A. D., Grimont, F., Richard, C., & Sakazaki, R. (1980). Edwardsiella hoshinae, a new species of enterobacteriaceae. *Current Microbiology, 4*(6), 347-351. doi:10.1007/bf02605375

HAWKE, J. P., McWHORTER, A. C., STEIGERWALT, A. G., & BRENNER, D. J. (1981). Edwardsiella ictaluri sp. nov., the Causative Agent of Enteric Septicemia of Catfish. *International Journal of Systematic and Evolutionary Microbiology, 31*(4), 396-400. doi:doi:10.1099/00207713-31-4-396

Hossain, M. J., Sun, D., McGarey, D. J., Wrenn, S., Alexander, L. M., Martino, M. E., Xinge, Y., Terhune, J. S., Liles, M. R. (2014). An Asian origin of virulent Aeromonas hydrophila responsible for disease epidemics in United States-farmed catfish. *MBio, 5*(3), e00848-00814. doi:10.1128/mBio.00848-14

Hossain, M. J., Waldbieser, G. C., Sun, D., Capps, N. K., Hemstreet, W. B., Carlisle, K., Griffin, M.J., Khoo, L., Goodwin, A. E., Sonstegard, T. S., Schroeder, S., Hayden, K., Newton, J. C., Terhune, J. S., Liles, M. R. (2013). Implication of lateral genetic transfer in the emergence of Aeromonas hydrophila isolates of epidemic outbreaks in channel catfish. *PLoS One, 8*(11), e80943. doi:10.1371/journal.pone.0080943

Huys, G., Pearson, M., Kampfer, P., Denys, R., Cnockaert, M., Inglis, V., & Swings, J. (2003). Aeromonas hydrophila subsp. ranae subsp. nov., isolated from septicaemic farmed frogs in Thailand. *Int J Syst Evol Microbiol, 53*(Pt 3), 885-891. doi:10.1099/ijs.0.02357-0

Janda, J. M., & Abbott, S. L. (2010). The genus Aeromonas: taxonomy, pathogenicity, and infection. *Clin Microbiol Rev, 23*(1), 35-73. doi:10.1128/CMR.00039-09

Karlyshev, A. V., Pallen, M. J., & Wren, B. W. (2000). Single-primer PCR procedure for rapid identification of transposon insertion sites. *Biotechniques, 28*(6), 1078, 1080, 1082.

Kim, M., Oh, H. S., Park, S. C., & 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*(Pt 2), 346-351. doi:10.1099/ijs.0.059774-0

Konstantinidis, K. T., Ramette, A., & Tiedje, J. M. (2006). The bacterial species definition in the genomic era. *Philos Trans R Soc Lond B Biol Sci, 361*(1475), 1929-1940. doi:10.1098/rstb.2006.1920

Konstantinidis, K. T., & Tiedje, J. M. (2005). Towards a genome-based taxonomy for prokaryotes. *J Bacteriol, 187*(18), 6258-6264. doi:10.1128/JB.187.18.6258-6264.2005

Lagesen, K., Hallin, P., Rodland, E. A., Staerfeldt, H. H., Rognes, T., & Ussery, D. W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res, 35*(9), 3100-3108. doi:10.1093/nar/gkm160

Lowe, T. M., & Eddy, S. R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res, 25*(5), 955-964.

Lv, Y., Yin, K., Shao, S., Wang, Q., & Zhang, Y. (2013). Comparative proteomic analysis reveals new components of the PhoP regulon and highlights a role for PhoP in the regulation of genes encoding the F1F0 ATP synthase in Edwardsiella tarda. *Microbiology, 159*(Pt 7), 1340-1351. doi:10.1099/mic.0.066803-0

Meyer, F. P., & Bullock, G. L. (1973). Edwardsiella tarda, a New Pathogen of Channel Catfish (Ictalurus punctatus). *Applied Microbiology, 25*(1), 155-156.

Noonin, C., Jiravanichpaisal, P., Soderhall, I., Merino, S., Tomas, J. M., & Soderhall, K. (2010). Melanization and pathogenicity in the insect, Tenebrio molitor, and the crustacean, Pacifastacus leniusculus, by Aeromonas hydrophila AH-3. *PLoS One, 5*(12), e15728. doi:10.1371/journal.pone.0015728

Oguro, K., Tamura, K., Yamane, J., Shimizu, M., Yamamoto, T., Ikawa, T., Ohnishi, K., Oshima, S., Imajoh, M. (2014). Draft Genome Sequences of Two Genetic Variant Strains of Edwardsiella piscicida, JF1305 and RSB1309, Isolated from Olive Flounder (Paralichythys olivaceus) and Red Sea Bream (Pagrus major) Cultured in Japan, Respectively. *Genome Announc, 2*(3). doi:10.1128/genomeA.00546-14

Orozova, P., Sirakov, I., Petkov, I., Crumlish, M., & Austin, B. (2012). Recovery of Aeromonas hydrophila associated with bacteraemia in captive snakes. *FEMS Microbiol Lett, 334*(1), 22-26. doi:10.1111/j.1574-6968.2012.02613.x

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A. R., Xia, F., Stevens, R. (2014). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res, 42*(Database issue), D206-214. doi:10.1093/nar/gkt1226

Park, S. B., Aoki, T., & Jung, T. S. (2012). Pathogenesis of and strategies for preventing Edwardsiella tarda infection in fish. *Vet Res, 43*, 67. doi:10.1186/1297-9716-43-67

Pressley, M. E., Phelan, P. E., 3rd, Witten, P. E., Mellon, M. T., & Kim, C. H. (2005). Pathogenesis and inflammatory response to Edwardsiella tarda infection in the zebrafish. *Dev Comp Immunol, 29*(6), 501-513. doi:10.1016/j.dci.2004.10.007

Reichley, S. R., Waldbieser, G. C., Tekedar, H. C., Lawrence, M. L., & Griffin, M. J. (2016). Complete Genome Sequence of Edwardsiella piscicida Isolate S11-285 Recovered from Channel Catfish (Ictalurus punctatus) in Mississippi, USA. *Genome Announc, 4*(6). doi:10.1128/genomeA.01259-16

Saejung, C., Hatai, K., Wada, S., Kurata, O., & Sanoamuang, L. (2011). Clinical observations of black disease in fairy shrimps, Streptocephalus sirindhornae and Branchinella thailandensis, from Thailand and pathogen verification. *J Fish Dis, 34*(12), 911-920. doi:10.1111/j.1365-2761.2011.01314.x

Shao, S., Lai, Q., Liu, Q., Wu, H., Xiao, J., Shao, Z., Wang, Q., Zhang, Y. (2015). Phylogenomics characterization of a highly virulent Edwardsiella strain ET080813(T) encoding two distinct T3SS and three T6SS gene clusters: Propose a novel species as Edwardsiella anguillarum sp. nov. *Syst Appl Microbiol, 38*(1), 36-47. doi:10.1016/j.syapm.2014.10.008

Tekedar, H. C., Karsi, A., Akgul, A., Kalindamar, S., Waldbieser, G. C., Sonstegard, T., Schroeder, S.G., Lawrence, M. L. (2015). Complete Genome Sequence of Fish Pathogen Aeromonas hydrophila AL06-06. *Genome Announc, 3*(2). doi:10.1128/genomeA.00368-15

Tekedar, H. C., Karsi, A., Williams, M. L., Vamenta, S., Banes, M. M., Duke, M., Scheffler, B. E., Lawrence, M. L. (2013). Complete Genome Sequence of Channel Catfish Gastrointestinal Septicemia Isolate Edwardsiella tarda C07-087. *Genome Announc, 1*(6). doi:10.1128/genomeA.00959-13

Tekedar, H. C., Kumru, S., Kalindamar, S., Karsi, A., Waldbieser, G. C., Sonstegard, T., Schroeder, S. G., Liles, M. R., Griffin, M. J., Lawrence, M. L. (2017). Draft Genome Sequences of Three Aeromonas hydrophila Isolates from Catfish and Tilapia. *Genome Announc, 5*(3). doi:10.1128/genomeA.01509-16

Tekedar, H. C., Kumru, S., Karsi, A., Waldbieser, G. C., Sonstegard, T., Schroeder, S. G., Liles, M. R., Griffin, M. J., Lawrence, M. L. (2016a). Draft Genome Sequence of Aeromonas hydrophila TN97-08. *Genome Announc, 4*(3). doi:10.1128/genomeA.00436-16

Tekedar, H. C., Kumru, S., Karsi, A., Waldbieser, G. C., Sonstegard, T., Schroeder, S. G., Liles, M. R., Griffin, M. J., Lawrence, M. L. (2016b). Draft Genome Sequences of Four Virulent Aeromonas hydrophila Strains from Catfish Aquaculture. *Genome Announc, 4*(4). doi:10.1128/genomeA.00860-16

Tekedar, H. C., Waldbieser, G. C., Karsi, A., Liles, M. R., Griffin, M. J., Vamenta, S., Sonstegard, T., Hossain, M., Schroeder, S, Khoo, L., Lawrence, M. L. (2013). Complete Genome Sequence of a Channel Catfish Epidemic Isolate, Aeromonas hydrophila Strain ML09-119. *Genome Announc, 1*(5). doi:10.1128/genomeA.00755-13

van Soest, J. J., Stockhammer, O. W., Ordas, A., Bloemberg, G. V., Spaink, H. P., & Meijer, A. H. (2011). Comparison of static immersion and intravenous injection systems for exposure of zebrafish embryos to the natural pathogen Edwardsiella tarda. *BMC Immunol, 12*, 58. doi:10.1186/1471-2172-12-58

Wang, Q., Yang, M., Xiao, J., Wu, H., Wang, X., Lv, Y., Xu, L., Zheng, H., Wang, S., Zhao, G., Liu, Q., Zhang, Y. (2009). Genome sequence of the versatile fish pathogen Edwardsiella tarda provides insights into its adaptation to broad host ranges and intracellular niches. *PLoS One, 4*(10), e7646. doi:10.1371/journal.pone.0007646

CHAPTER III

COMPLETE GENOME SEQUENCING AND COMPARATIVE GENOMICS OF THE FISH PATHOGENS *EDWARDSIELLA ICTALURI* 93-146 AND *EDWARDSIELLA PISCICIDA* C07-087

3.1 Abstract

Edwardsiella ictaluri and *Edwardsiella piscicida* are important fish pathogens affecting cultured and wild fish worldwide. To investigate the molecular mechanisms of each species, we sequenced the *E. ictaluri* 93-146 and *E. piscicida* C07-087 genomes and compared them to other available *Edwardsiella* species. Average Nucleotide Identity and core genome phylogenetic tree analyses indicated that the five *Edwardsiella* species separated from each other clearly. Pan-Core genome analysis showed that *Edwardsiella* genus members have 9,436 genes in their pan-genome while the core genome consists of 1,426 genes. Orthology clusters analysis showed that *E. ictaluri* and *E. piscicida* genomes have the most shared clusters. However, *E. ictaluri* and *E. piscicida* also have unique features; for example, the *E. ictaluri* genome encodes urease enzymes and cytochrome o ubiquinol oxidase subunits, whereas *E. piscicida* genomes encode tetrathionate reductase operons, capsular polysaccharide synthesis enzymes, and vibrioferrin related genes. Additionally, we report for the first time that *E. ictaluri* 93-146 and three other *E. ictaluri*

genomes encode a Type IV secretion system (T4SS), whereas none of the *E. piscicida* genomes encode this system. Additionally, the *E. piscicida* C07-087 genome encodes two different Type VI secretion systems (T6SS). *E. ictaluri* genomes tend to encode more insertion elements, phage regions, and genomic islands than *E. piscicida*. We speculate that the T4SS could contribute to the increased number of mobilome elements in *E. ictaluri* compared to *E. piscicida*. Two of the *E. piscicida* genomes encode full CRISPR-Cas regions, whereas none of the *E. ictaluri* genomes encode Cas proteins. Overall, comparison of the *E. ictaluri* and *E. piscicida* genomes can provide new insights on pathogenicity and host adaptation of the two species.

3.2 Introduction

The genus *Edwardsiella* is classified in the *Enterobacteriaceae* family and contains several species that are facultative intracellular pathogens (Ewing et al., 1965). *E. ictaluri* is a primary bacterial pathogen that was originally identified as the causative agent of enteric septicemia of catfish (ESC) in channel catfish in the U.S. (Hawke et al., 1981). It is known to infect a broad range of other freshwater fish species in North America and Asia (Crumlish et al., 2002; Hawke et al., 2013; J. Y. Liu et al., 2010; Plumb et al., 1983; Sakai et al., 2008; Soto et al., 2012; Yuasa et al., 2003). *E. ictaluri* isolates from catfish, zebrafish, and tilapia have distinct genotypes and plasmid profiles (Griffin et al., 2016). Comparison of U.S. and Vietnamese catfish isolates revealed no detectable genetic difference, but they had distinct plasmid profiles (Rogge et al., 2013). *E. piscicida* has been recently identified and genetically characterized as a new species. It was previously classified as *E. tarda* based on phenotypic and biochemical tests, but genetically it is distinguishable from *E. tarda* strains (Abayneh et al., 2013). *E. piscicida*

has been isolated from diseased catfish, whitefish, sea bream, and bass in North America and Europe (Fogelson et al., 2016; Griffin et al., 2014; Katharios et al., 2015; Shafiei et al., 2016).

E. ictaluri survives in pond water and sediment under variable environmental conditions (Plumb et al., 1986). It can actively invade healthy fish through multiple routes of entry and establish either acute or chronic infection (Baldwin et al., 1993; Morrison et al., 1994; Nusbaum et al., 1996). *E. ictaluri* can evade innate immune mechanisms (Miyazaki et al., 1985; Shoemaker et al., 1997) and replicate inside professional phagocytic cells (Booth et al., 2006) despite triggering oxidative and nitrosative response (Chen et al., 1991; Yao et al., 2014; Yeh et al., 2013). Type III secretion system (T3SS) and type VI secretion system (T6SS) transfer effector proteins directly into host cells to manipulate host cell function and enable intracellular replication (Nakamura et al., 2013; Zheng et al., 2007). The pathogenesis of *E. piscicida* is not well known, but clinical signs of the disease it causes and its virulence factors are similar to *E. ictaluri*. To understand the pathogenic properties of these two aquaculture pathogens, our research group sequenced the *E. ictaluri* 93-146 (Williams et al., 2012) and *E. piscicida* C07-87 (Tekedar et al., 2013) genomes. Here we report a comparison of these two genomes to identify unique and shared features, as well as a broader analysis of these two species compared to other species in the *Edwardsiella* genus.

Based on average nucleotide identity (ANI) and core genome phylogenetic analysis, *E. ictaluri* and *E. piscicida* are more closely related to each other than *E. hoshinae* and *E. tarda*. The two species have some unique features. *E. ictaluri* genomes contain urease production operons and encode cytochrome o ubiquinol oxidase subunits,

whereas *E. piscicida* genomes encode tetrathionate reductase operons, capsular polysaccharide synthesis enzymes, and vibrioferrin related genes. *E. piscicida* C07-087 encodes two different T6SSs (T6SS-type I subtype i4b and T6SS-type I subtype i2), whereas *E. ictaluri* 93-146 genomes encode T4SS-typeG. *E. ictaluri* contains more mobilome elements than *E. piscicida*, which could be explained by the presence of a T4SS. In summary, this analysis provides valuable information about unique and shared features of two important pathogens in the *Edwardsiella* genus.

3.3 Material and Methods

3.3.1 Bacterial strains and *Edwardsiella* genomes

E. ictaluri 93-146 was isolated from a natural outbreak in Louisiana in 1993, and *E. piscicida* C07-87 was isolated from catfish with gastrointestinal septicemia in a commercial aquaculture pond in Mississippi. For comparative genomics, all the available genomes in the *Edwardsiella* genus were obtained from NCBI (as of 07/28/2017), which are listed in Table 1.

Table 3.1 *Edwardsiella* genomes used in the comparative genomic analyses.

#	Species	Strain	Location	Source	Genome close level/contig numbers	Size (Mb)	GC %	Plasmid	Accession	Reference
1	*E. ictaluri*	RUSVM-1	Western Hemisphere	Nile Tilapia	Complete	3.63064	57.40		NZ_CP020466	(Reichley et al., 2017)
2	*E. ictaluri*	LADL11-100	USA	Zebrafish	Scaffold	3.70126	57.40	1	PRJNA285663	(R. Wang et al., 2015)
3	*E. ictaluri*	LADL11-194	USA	Zebrafish	Scaffold	3.70614	57.40	2	PRJNA285852	(R. Wang et al., 2015)
4	*E. ictaluri*	ATCC 33202	USA	Channel catfish	Contig/117	3.70369	57.60		PRJNA66365	(van Soest et al., 2011)

Table 3.1 (continued)

5	*E. ictaluri*	93-146	USA	Channel catfish	Complete	3.8123	57.40		NC_012779	(Williams et al., 2012)
6	*E. piscicida*	RSB1309	Japan	Red Sea Bream	Contig/38	3.83865	59.30		PRJDB1727	(Oguro et al., 2014)
7	*E. sp.*	EA181011	Israel	White grouper	Complete	3.93417	59.10		NZ_CP011364	(Reichley et al., 2015)
8	*E. sp.*	LADL05-105	USA	Tilapia	Complete	4.14204	58.80		NZ_CP011516	(Reichley et al., 2015)
9	*E. anguillarum*	ET080813	China	Japanese eel	Complete	4.32965	58.37	2	NZ_CP006664	(Shao et al., 2015)
10	*E. anguillarum*	ET081126R	China	European eel	Draft 102 contig	4.23812	58.50		NZ_JACP01000006	(Shao et al., 2015)
11	*E. anguillarum*	ET070829	China	Marbled eel	Contig/121	4.28636	58.40		NZ_JABY01000001	(Shao et al., 2015)
12	*E. anguillarum*	ET080729	China	Marbled eel	Contig/115	4.15319	58.40		NZ_JABS01000002	(Shao et al., 2015)
13	*E. tarda*	ET883	Norway	European eel	Contig/167	3.82967	59.60		PRJNA259400	(Abayneh et al., 2012)
14	*E. tarda*	FL6-60	USA	Striped bass	Complete	3.7288	59.70	1	NC_017309	(van Soest et al., 2011)
15	*E. piscicida*	ETW41	Korea	Eel pond water	Chromosome	3.96904	59.48	1	NZ_CP019440	N/A
16	*E. piscicida*	JF1305	Japan	Olive Flounder	Contig/72	3.71682	59.80		PRJDB1727	(Oguro et al., 2014)
17	*E. tarda*	EIB202	China	Turbot	Complete	3.80417	59.67	1	NC_013508	(Q. Wang et al., 2009)
18	*E. piscicida*	C07-087	USA	Channel catfish	Complete	3.85704	59.60		NC_020796	(Tekedar et al., 2013)
19	*E. piscicida*	ACC35.1	Europe	Turbot	Scaffold	3.84163	59.70		PRJNA353909	(Bujan et al., 2017)
20	*E. piscicida*	S11-285	USA	Channel catfish	Complete	3.92677	59.59	1	NZ_CP016044	(Reichley et al., 2016)
21	*E. hoshinae*	ATCC 33379	France	Female puffin	Contig/110	3.70661	56.80		PRJDB228	(Grimont et al., 1980)
22	*E. hoshinae*	ATCC 35051	USA	Monitor lizard	Complete	3.81165	56.90		NZ_CP016043	(Reichley et al., 2017)
23	*E. tarda*	ASE201307	China	Swamp eel	Scaffold	3.68479	57.10		PRJNA329181	N/A
24	*E. tarda*	NBRC 105688	USA	Human feces	Contig/34	3.61038	57.30		PRJDB227	(Ewing et al., 1965)
25	*E. tarda*	ATCC 23685	USA	Human feces	Scaffold	3.74457	56.80		PRJNA28661	(Ewing et al., 1965)
26	*E. tarda*	ATCC 15947	USA	Human feces	Contig/247	3.69469	57.10		PRJNA39897	(Ewing et al., 1965)
27	*E. tarda*	FL95-01	USA	Channel catfish	Complete	3.6207	57.40		NZ_CP011359	(Reichley et al., 2015)
28	*E. tarda*	DT	China	Oscar fish	Contig/83	3.75958	57.00		PRJNA66369	(Yang et al., 2012)
29	*E. tarda*	NCIMB2034	USA	Unknown fish	Contig/221	3.76083	57.20		PRJNA357015	(Bujan et al., 2017)

3.3.2 Sequencing, assembly, and annotation

To determine the circularized genome sequences of *E. ictaluri* 93-146 and *E. piscicida* C07-87, high-throughput shotgun sequencing was used, which was followed by assembly and annotation as described (Tekedar et al., 2013; Williams et al., 2012). For all the evaluated genomes in this study, annotation and gene prediction were conducted by RAST (Rapid Annotation using Subsystem Technology) (Overbeek et al., 2014) and PGAP (Prokaryotic Genome Annotation Pipeline) (Tatusova et al., 2016).

3.3.3 ANI calculation

Average nucleotide identity (ANI) values (Konstantinidis et al., 2005) were calculated based on BLASTN results as described (Goris et al., 2007) using EDGAR (Blom et al., 2016) and JSpecies (Richter et al., 2009).

3.3.4 Phylogenetic tree creation

Phylogenetic distances were calculated based on the complete core genome derived from 29 *Edwardsiella* genomes with two *Escherichia* genomes as outgroups. All orthologous gene sets of the core genome were individually aligned using MUSCLE (Edgar, 2004), the resulting alignments were concatenated, and phylogenetic distances were calculated using the neighbor-joining method in PHYLIP (Felsenstein, 1989). Tree topology was validated using 500 bootstrapping iterations.

3.3.5 Pan-Core genome analysis

Core- and pan-genome analyses were performed using EDGAR 2.0. EDGAR is based on a generic orthology criterion, which in turn is based on BLAST score ratio values (SRV). In SRV, BLAST scores are normalized in relation to the best hit possible,

which is the BLAST result of a query gene against itself. Based on the distribution of SRVs in the dataset, a cutoff is estimated as described (Blom et al., 2016). In the current study, orthologs were defined as genes with a reciprocal best BLAST hit and with both single hits having a SRV above 31%.

3.3.6 Singleton analysis

Singleton genes were identified with EDGAR 2.0. In contrast to the core and pan-genome calculation, the singleton calculation is not based on reciprocal best BLAST hits, but on a stricter definition. Only genes that show no BLAST hits with a SRV >31% against any other analyzed genome are identified as singleton genes. Identified singleton elements were downloaded and searched against the microbial virulence database (MVirDB) as described below.

3.3.7 Orthologous gene clusters analysis

Genome-wide analysis of orthologous clusters was calculated for only complete *E. ictaluri* and *E. piscicida* genome sequences using OrthoVenn (Y. Wang et al., 2015). With this tool, genome wide comparisons and visualization of orthologous clusters from complete *Edwardsiella* genomes were created and visualized using the following parameters: E-value, 1e-5; and inflation value, 1.5. Protein files for each genome were downloaded from NCBI and uploaded to OrthoVenn (http://www.bioinfogenome.net/OrthoVenn/) for identification of orthologous clusters. For accuracy, only complete genomes were used in this analysis. We performed four different analyses using only complete, circular genomes. First, *E. ictaluri* 93-146 genome was compared against *E. ictaluri* RUSVM-1. Second, *E. piscicida* C07-087 and

other complete *E. piscicida* genomes (strains S11-285, EIB202, and FL6-60) were compared. Third, *E. ictaluri* 93-146 genome was compared against *E. piscicida* C07-087. Finally, 6 *Edwardsiella* genomes were compared against each other: 2 *E. ictaluri* (strains 93-146 and RUSVM-1) and 4 *E. piscicida* (strains C07-087, S11-285, EIB202, and FL6-60) genomes.

3.3.8 Genome structure variation

Homologous genes in the *E. ictaluri* (strains 93-146, LADL11-100, LADL11-194, ATCC33202, and RUSVM-1), *E. piscicida* (strains C07-087, ACC35.1 ETW41, JF1305, and S11-285), and *E. tarda* (strains EIB202, ET883, and FL6-60) genomes were identified using comparative tools in RAST (Overbeek et al., 2014) to allow for potential genome-to-genome annotation.

3.3.9 Protein Secretion Systems

To detect secretion systems and their components in the evaluated *E. ictaluri* and *E. piscicida* genomes, bacterial Type I, II, III, IV, V, and VI secretion system proteins were identified by uploading protein sequences to MacSyDB/TXSSdb (Abby et al., 2016; Abby et al., 2012; Guglielmini et al., 2014). In addition to MacSyDB/TXSSdb, SecRet6 was used to confirm T6SS proteins (Li et al., 2015). A bacterial type IV secretion system resource, SecRet4, was used to confirm identified elements from this secretion system (Bi et al., 2013).

3.3.10 Insertion elements

Issaga was used to identify individual insertion sequences in each *E. ictaluri* and *E. piscicida* genome (Varani et al., 2011). After identification, results were filtered by removing false predicted elements from the final list.

3.3.11 Genomic Islands (GI)

IslandViewer 4 was used to identify genomic islands in the *E. ictaluri* and *E. piscicida* genomes (Bertelli et al., 2017). Some of the *Edwardsiella* genus members were pre-analyzed in IslandViewer. For the unanalyzed genomes, their GBK files were downloaded from the RAST annotation server, and *E. ictaluri* 93-146 and *E. piscicida* C07-087 genomes were used as a reference. In this tool, four different island prediction methods were used, and they are shown in different colors: IslandPick (Green), SIGI-HMM (Orange), IslandPath-DIMOB (Blue), and Islander (Cyan). In addition, red color represents identified genomic islands that were predicted by at least one method. For unfinished genomes, contig boundaries are shown with zigzag marks. Also, a virulence-associated genes identification search was implemented. The identified genomic islands (based on integrated results in IslandViewer 4) were downloaded and searched against MvirDB as described below to assess the potential role of genomic islands in bacterial virulence.

3.3.12 Prophages

PHASTER (PHAge Search Tool Enhanced Release) was used to identify prophages in the *E. ictaluri* and *E. piscicida* genomes (Arndt et al., 2016). Some of the genomes were pre-computed in the PHASTER database, and some were not. For those

with pre-computed genomes, nucleotide files were concatenated to serve as input files and submitted for phage elements identification in the *E. ictaluri* and *E. piscicida* genomes. Depending on the result from PHAST, identified phage regions were placed into three categories: 1) if the identified region's score was >90, these were considered intact phage regions; if it was between 70-90, it was considered questionable; and if it was <70, it was classified as an incomplete phage region.

3.3.13 Integron identification

Presence of integron elements in the *E. ictaluri* and *E. piscicida* genomes was investigated by IntegronFinder (Cury et al., 2016) following parameters were used based on nucleotide sequences; threshold for clustering: 4000 bp; *attC* e-value:1; maximum value for *attC* size:200; minimum value for *attC* size:40.

3.3.14 CRISPR and Cas elements analysis

To determine CRISPR-Cas systems and their elements in the *Edwardsiella* genus, CRISPRfinder (Grissa et al., 2007) and MacSyFinder (Abby et al., 2014) were used. In MacSyFinder, maximal e-value was set to 1.0, independent E-value was set to 0.001, and minimal profile coverage parameters were set for Cas elements identification in the *Edwardsiella* genus. Results were sorted into three categories: *mandatory*, *accessory*, and *forbidden*. If the elements were ubiquitous and identifiable, they were considered *mandatory*. *Accessory* components could be essential but not identifiable due to rapid evolution or other reasons. If the evaluated element was partly homologous, the system could identify it as a *forbidden* element.

3.3.15 Virulence factors

Potential virulence factors of singleton elements and genomic islands in *E. ictaluri* 93-146 and *E. piscicida* C07-087 genomes were assessed by searching them against MVirDB, which was downloaded from http://mvirdb.llnl.gov/ (Zhou et al., 2007). Local BLAST was conducted with all the predicted proteins for all evaluated genomes using CLC Genomics Workbench (Version 6.5). In this analysis, BLAST results were considered significant when the E-value was $<1 \times 10^{-20}$.

3.4 Results

3.4.1 Genome Features

Genome summaries and features are listed in Table 1. *Edwardsiella* genomes G+C content ranges from 56.8 to 59.80%. *E. hoshinae* genomes have the lowest G+C content, whereas *E. ictaluri* genomes have approximately 57.4%, and *E. piscicida* genomes have the highest G+C content (59.7%). Plasmids were present in *E. hoshinae* strain ET080813, four *E. piscicida* genomes (strains FL6-60, ETW41, EIB202, and S11-285), and two *E. ictaluri* genomes (strains LADL11-100 and LADL11-194).

3.4.2 ANI and phylogenetic analysis

ANI showed that all *Edwardsiella* species (*E. ictaluri*, *E. piscicida*, *E. anguillarum*, *E. hoshinae*, and *E. tarda*) are distinctly different from each other (Figure 1). ANI results indicated that *E. anguillarum* and *E. piscicida* are the closest related species in the genus, followed by *E. ictaluri*. *E. tarda* and *E. hoshinae* are more distantly related to three other *Edwardsiella* species. Phylogenetic analysis based on the complete core genome of 29 *Edwardsiella* genomes and 2 *Escherichia coli* genomes (strains 0157-

H7 and CFT073; used as outliers) was conducted. Results showed that the five *Edwardsiella* species formed a slightly conserved branch and were clearly distinguishable from each other (Figure 2). Except for one branch in the *E. tarda* subclade [*E. tarda* ATCC 23685-*E. tarda* strain NBRC 105688-*E. tarda* strain FL95 01] with only 57% conservation, all branches showed at least 94% bootstrap support.

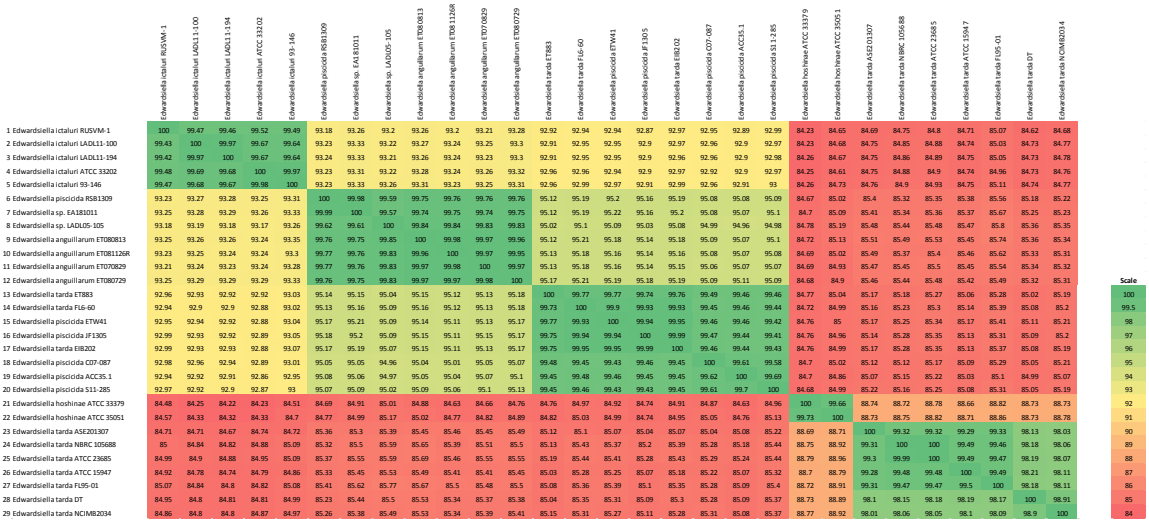


Figure 3.1 Average nucleotide identities (ANI) of *Edwardsiella* genomes.

Note that: All the strain names used in the ANI table are extracted from NCBI but should be named properly. For instance, *E. piscicida* RSB1309, *Edwardsiella sp.* EA181011 and LADL05-105 should be reclassified as *E. anguillarum* species. Also, *E. tarda* strains ET883, FL6-60, EIB202 should be reclassified under *E. piscicida* species.

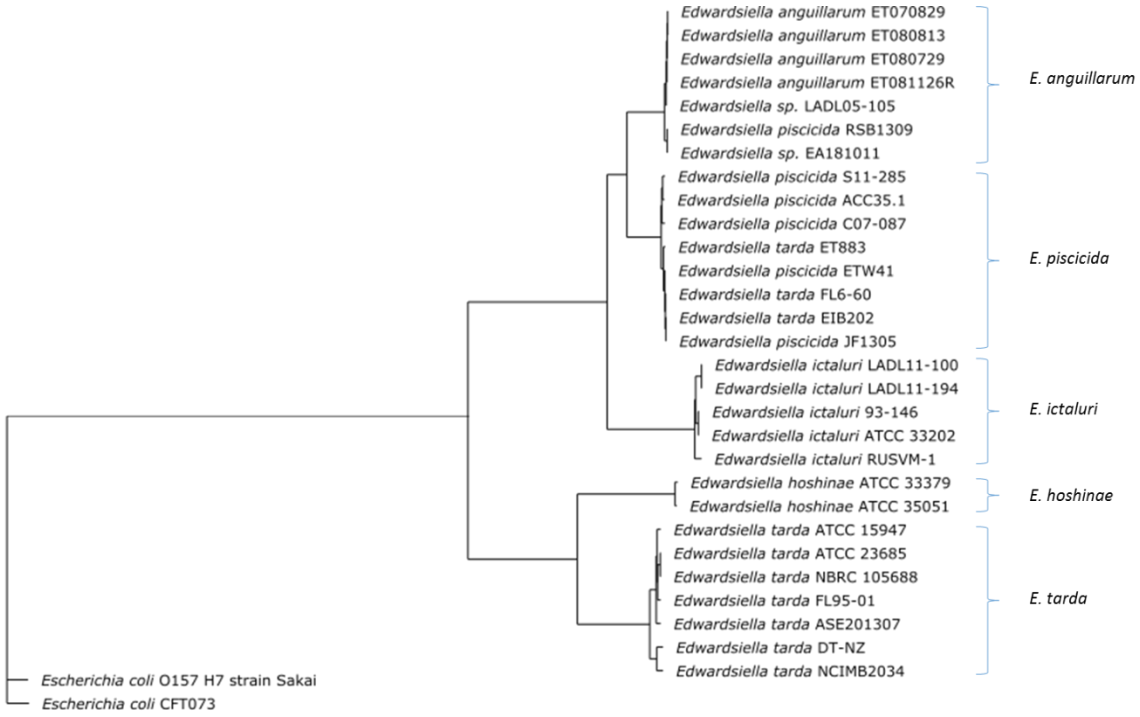


Figure 3.2 Phylogenetic tree of *Edwardsiella* genomes based on core genomes.

Note that the main *Edwardsiella* branch was manually shortened for improved visualization.

3.4.3 Pan-Core genome analysis

Pan genome analysis of the 29 *Edwardsiella* strains identified a total of 9,436 genes. On the other hand, there were 1,426 genes identified in the core genome (Figure 3).

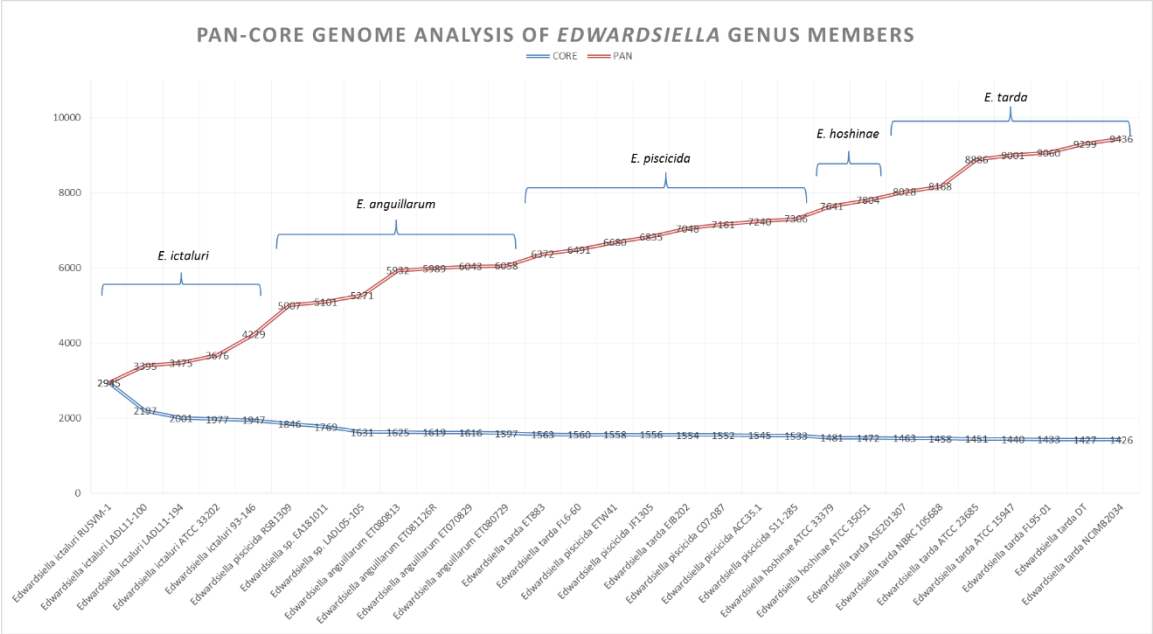


Figure 3.3 Pan vs. Core Plot analysis of *Edwardsiella* genus members.

Note that, genomes are grouped and ordered by ANI result.

3.4.4 Singleton analysis

Using all 29 *Edwardsiella* genomes, the *E. ictaluri* 93-146 genome had 268 identified singletons, whereas the *E. piscicida* C07-087 genome had 47 singleton genes. Four of these encode *E. ictaluri* 93-146 proteins (NT01EI_2881, NT01EI_2480, NT01EI_2479, and NT01EI_1391) that had significant matches in MVirDB, whereas only one of the proteins encoded by singletons in the *E. piscicida* genome (ETAC_16540) had a significant match in MVirDB.

3.4.5 Orthology analysis

In our orthology analysis, we used four different combinations. In the first one, we performed a comparison of proteins encoded by the *E. ictaluri* 93-146 and *E. ictaluri* RUSVM-1 genomes. These two strains shared 2,774 clusters, and only 13 clusters

(including insertion sequences, phage, putative glycosyltransferase, 2-dihydro-3-deoxy-D-gluconate 5-dehydrogenase, and NADH oxidase elements) were unique to *E. ictaluri* 93-146. Four clusters (including prophage CP4-57 regulatory protein AlpA, type I restriction enzyme EcoAI R, and anticodon nuclease) were unique to RUSVM-1 (Figure 4-A).

Second, we compared the *E. piscicida* strains with complete genomes and found that the *E. piscicida* S11-285 genome encodes five unique clusters (including putative lambdoid prophage e14 repressor protein C2, ornithine decarboxylase, and antitermination protein Q homolog from lambdoid prophage Qin), which is more than the *E. piscicida* C07-087, EIB202, and FL06-60 genomes. *E. piscicida* C07-087 had two unique clusters (Figure 4-B).

Third, we evaluated orthologous clusters encoded by the *E. ictaluri* 93-146 and *E. piscicida* C07-087 genomes and found that they share 2,708 orthologous clusters. Specifically, the *E. ictaluri* 93-146 genome encodes 18 unique clusters (such as insertion elements, transposons, and invasion lpaB proteins), whereas the *E. piscicida* C07-087 genome encodes 20 unique clusters (such as probable dipeptidase, platelet binding protein GspB, sn-glycerol-3-phosphate-binding periplasmic protein UgpB, transcriptional regulatory protein FixJ, and cellulose synthase catalytic subunits) (Figure 4-C).

Fourth, all the *E. ictaluri* strains (93-146 and RUSVM-1) and *E. piscicida* strains (S11-285, C07-087, EIB202 and FL06-60) that were fully sequenced were evaluated. They shared 2,495 clusters, and only two hypothetical proteins were uniquely encoded by the *E. piscicida* C07-087 genome, whereas the *E. ictaluri* genome encoded only two unique insertion element clusters (Figure 4-D).

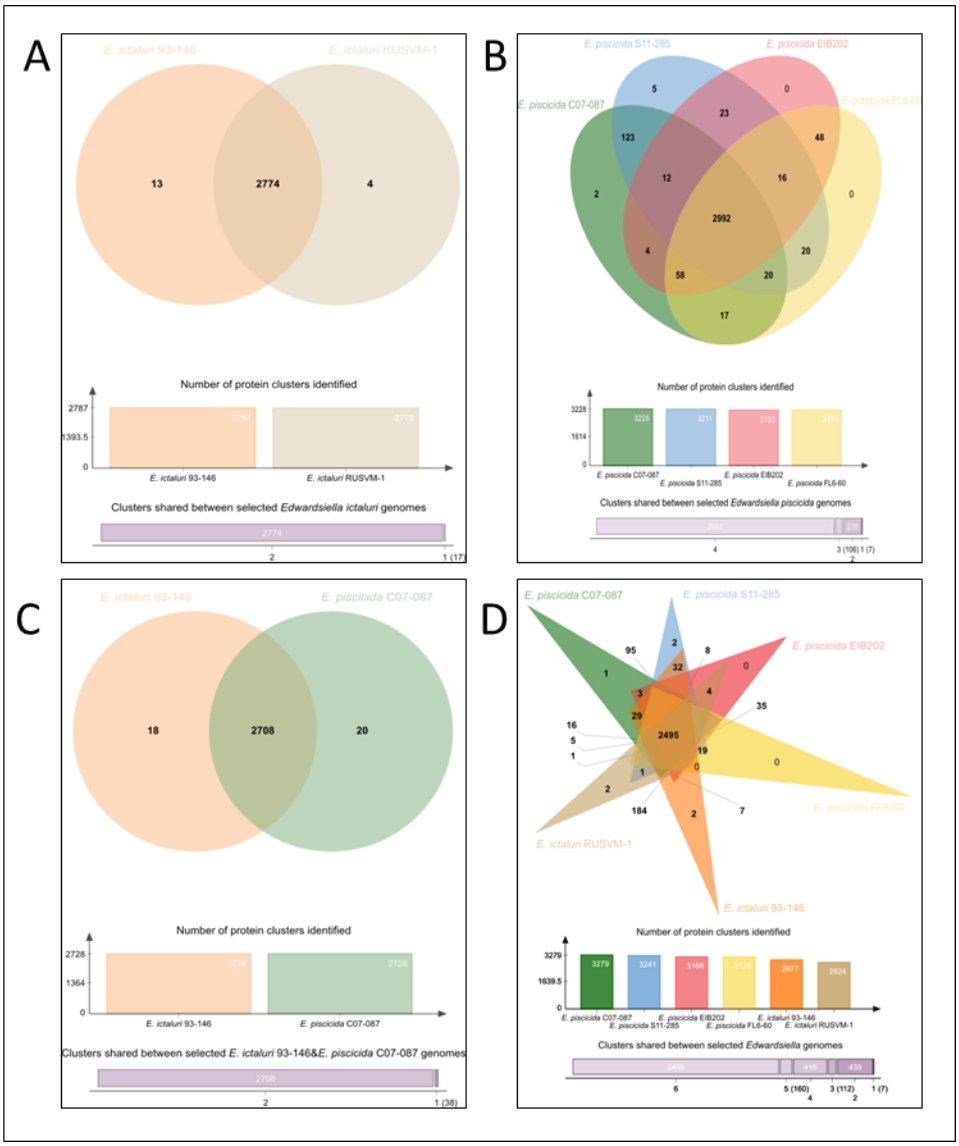


Figure 3.4 Putative functions of orthologs and unique clusters for the *E. piscicida* and *E. ictaluri* genomes

3.4.6 Gene variation in the thirteen *E. ictaluri* and *piscicida* genomes

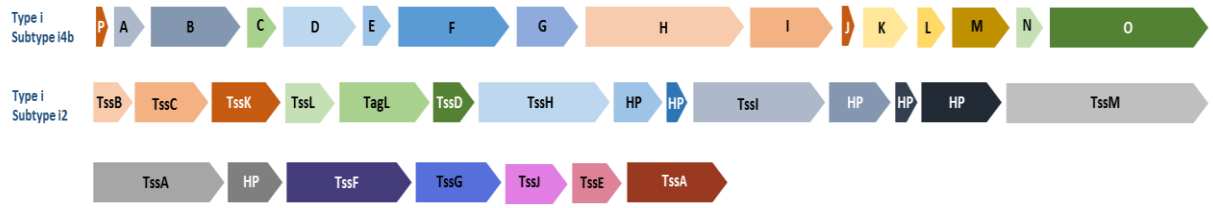
Evaluation of gene variation using all the sequenced *E. ictaluri* and *E. piscicida* genomes (complete and draft) showed that the *E. ictaluri* genomes encode urease and cytochrome o ubiquinol oxidase subunits. On the other hand, the *E. piscicida* genomes

encode tetrathionate reductase, capsular polysaccharide synthesis enzymes, and vibrioferrin related genes.

3.4.7 Secretion Systems

We evaluated all the secretion elements encoded in the *E. ictaluri* and *E. piscicida* genomes. All of the evaluated *E. ictaluri* and *E. piscicida* genomes encoded T1SS, T3SS, T5SS, and T6SS. Only the *E. ictaluri* 93-146, LADL11-100, LADL11-194, and ATCC 33202 genomes encoded T4SS-typeG (Figure 5-B; genes listed in Table 3). Intriguingly, the *E. ictaluri* RUSVM-1 genome did not encode any of the T4SS. T6SS-type i subtype i4b was encoded by all the *E. ictaluri* strains. Moreover, only the *E. piscicida* C07-087, S11-285, and ACC35.1 genomes encoded two different T6SS; one is T6SS-type i subtype i4b, and the other one is T6SS-type I subtype i2 (Figure 5-A). The other evaluated *E. piscicida* genomes (strains ET883, FL6-60, ETW41, JF1305, and EIB202) encoded only T6SS-type I subtype i4b. Finally, *E. piscicida* strain EIB202 carried a plasmid that encodes T4SS elements.

A



B

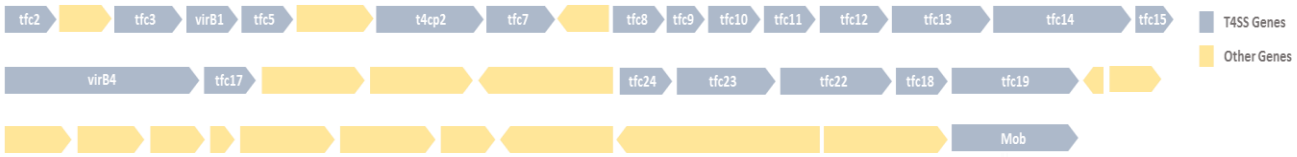


Figure 3.5 A- T6SS Type-I subtype I4B and I2 in the *E. ictaluri* and *E. piscicida* genomes B-T4SS in *E. ictaluri* 93-146.

Table 3.2 Type Four Secretion System genes

#	Protein	Profile	Gene	Protein
1	tfc2	Core	PilL	Hypothetical protein, conjugative transfer region protein
2	tfc3		HP	Hypothetical protein, integrating conjugative element protein, PFL_4693 family
3	virB1		virB1	Hypothetical protein, lytic transglycosylase
4	tfc5		HP	Hypothetical protein
5	t4cp2		TraD	Conjugative coupling factor TraD
6	tfc7		HP	Hypothetical protein, integrating conjugative element membrane protein, PFL_4697 family
7	tfc8		HP	Conserved hypothetical protein, RAQPRD motif
8	tfc9	Pilin	virB2-like	Hypothetical protein, integrating conjugative element protein
9	tfc10	Pilin	virB2-like	Integrating conjugative element membrane protein
10	tfc11	Inner Scaffold	virB3/TraL	Hypothetical protein
11	tfc12	Inner Scaffold	virB8/TraE	Integrating conjugative element protein
12	tfc13	Core	TraK	Hypothetical protein
13	tfc14	Core	TraB	Integrating conjugative element protein, Bacterial conjugation TrbI-like protein
14	tfc15	Core	TraV	Conjugative transfer region lipoprotein
15	virB4		TraC	F pilus assembly Type-IV secretion system for plasmid transfer, TraC
16	tfc17		HP	Hypothetical protein, acetyltransferase
17	tfc24			Integrating conjugative element protein
18	tfc23			Integrating conjugative element protein
19	tfc22			Integrating conjugative element protein
20	tfc18	Inner Scaffold	virB6/TraG	Hypothetical protein
21	tfc19		virB6/TraG	TraG_N
22	MobH		TraI	TraI_2_C, Conjugative transfer protein MobH, relaxase

3.4.8 Insertion sequences

Several insertion sequences were identified in the *E. ictaluri* and *E. piscicida* genomes. IS3 family members (IS407, IS51, and IS3) were the only family encoded by all the evaluated *E. ictaluri* and *E. piscicida* genomes. The *E. ictaluri* 93-146, LADL11-100, ATCC 33202, and RUSVM-1 genomes encoded eleven different IS families, whereas strain LADL11-194 lacked one insertion element from the IS*As*1 family. Thus, *E. ictaluri* genomes appeared to have relatively conserved insertion sequences. On the other hand, *E. piscicida* genomes showed more variable insertion sequence elements. Specifically, *E. ictaluri* genomes encoded more IS1, IS427, IS110, IS256, and ISL3 elements than *E. piscicida*. IS481 elements were unique to *E. ictaluri*, whereas IS50, ISH8, IS6, IS21, IS91, and Tn3 elements were unique to *E. piscicida*, but they were encoded by only some of the *E. piscicida* genomes (Table 2).

Table 3.3 Insertion elements distribution in the *E. ictaluri* and *E. piscicida* genomes.

Families	Sub-groups	Typical size-range	Mechanisms	*Edwardsiella ictaluri* RUSVM-1	*Edwardsiella ictaluri* LADL11-100	*Edwardsiella ictaluri* LADL11-194	*Edwardsiella ictaluri* ATCC 33202	*Edwardsiella ictaluri* 93146	*Edwardsiella piscicida* ET883	*Edwardsiella piscicida* FL6-60	*Edwardsiella piscicida* ETW41	*Edwardsiella piscicida* JF1305	*Edwardsiella piscicida* EIB202	*Edwardsiella piscicida* C07-087	*Edwardsiella piscicida* ACC35.1	*Edwardsiella piscicida* S11-285
IS*1*		740-1180	Copy, paste, conintegrate	37	32	39	29	39						1		
IS*3*	IS*407*	1100-1400	Copy, paste	8	8	7	10	10	4	7	5	5	5	8	3	9
	IS*51*	1000-1400		5	3	4	4	4	1	6	2	2	2	2	2	8
	IS*3*	1150-1750		7	7	6	6	6	1	2	2	2	2	1	2	2
IS*481*	-	950-1300	Copy, paste	1	2	3	2	3								
IS*4*	IS*10*	1200-1350	Cut and paste	4	3	3	2	6				2				
	IS*50*	1350-1550								1	1	1	1			
	IS*H8*	1400-1800									5					
IS*5*	IS*903*	950-1150		3	5	5	19	5	1	4	4	4	4	4		
	IS*427*	800-1000		36	33	31	31	68					1			14
IS*6*	-	700-900	Conintegrate								7	1				
IS*21*	-	1750-2600									39	2	23			
IS*91*	-	1500-2000	Rolling circle								3	2				
IS*110*	IS*1111*	1200-1500		16	7	9	16	19	2	3	3	3	3		2	3
IS*200* /IS*605* [t]		1300-2000	Peel and paste	3	8	3	3	5								
	IS*200*	600-750		1	1	1	1	1	1	1				1	1	5
IS*256*	-	1200-1500	Copy and paste	104	23	28	13	47		1	1	2	2		2	1
IS*630*	-	1000-1400	Cut and paste	27	19	25	11	22								
IS*As1*	-	1200-1500		1	1		1	1	1	1	1	1	1	1	1	1
IS*L3*	-	1300-2300		30	13	12	20	20		2	3	3	3	2	3	
Tn*3*	-	>3000	Cointegrate								3	1				

Scale: 104, 68, 47, 39, 37, 31, 27, 25, 22, 19, 15, 5, 3, 2, 1

3.4.9 Genomic islands

The evaluated *E. ictaluri* strains had more genomic islands (RUSVM-1, 35; LADL11-100, 40; LADL11-194, 38; ATCC 33202, 39; and 93-146, 36) than *E. piscicida* strains (ET883, 34; FL6-60, 23; ETW41, 34; JF1305, 26; EIB202, 30; C07-087, 29; ACC35.1, 28; and S11-285, 36) (Figure 6). *E. ictaluri* 93-146 had 431 genes in its genomic islands, and 165 of them had significant BLAST matches with MVirDB. On the other hand, the *E. piscicida* C07-087 genome had 292 genes in its genomic islands, and 88 of them had significant matches in MVirDB.

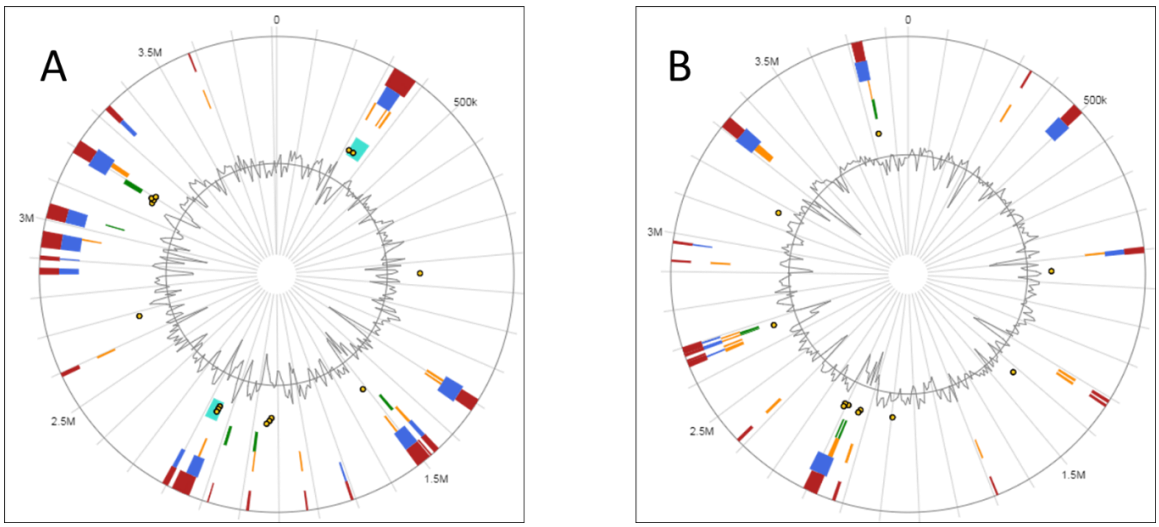


Figure 3.6 Genomic islands in *Edwardsiella* genomes

A-*Edwardsiella ictaluri* 93-146 B-*Edwardsiella piscicida* C07-87

3.4.10 Phage sequences

E. ictaluri strains tended to have more prophages than *E. piscicida* strains. *E. ictaluri* strains RUSVM-1 had 5 incomplete (In); LADL11-100 had 4 intact (I) and 2 In; LADL11-194 had 4 I, 1 In, and 2 questionable (Q); ATCC 33202 had 3 I, 1 Q, and 1 In; 93-146 had 3 I and 1 Q. *E. piscicida* strains ET883 had 3 I, 3 Q, and 1 In; FL6-60 had 1 In; pFL6-60 had 1 I; ETW41 had 2 I, 2 Q, and 2 In; JF1305 had 1 Q and 1 In; EIB202 had 2 Q and 1 I; C07-087 had 3 I; ACC35.1 had 3 I and 1 Q; and S11-285 had 1 I. G+C content of *E. ictaluri* and *E. piscicida* phage regions varied between 42.25% to 56.93 % and 46% to 59.82 %, respectively.

3.4.11 Integron identification

Integron identification analysis showed that there were no integrons in the *E. ictaluri* genomes, and two *E. piscicida* strains (ETW41 and JF1305) encoded integrons.

3.4.12 CRISPR (Clustered regularly interspaced short palindromic repeats) and Cas elements

There are three major Cas systems (Type I, Type II, and Type III) with 10 different subtypes (Type I-A to F, Type II-A and B, and Type III-A and B) (Abby et al., 2014). The *E. ictaluri* and *E. piscicida* genomes had *csm2* gene from Type III-A and *cas3* gene from Type I. Only *E. piscicida* ACC35.1 and S11-285 genomes encoded Type I-E including *cas1*, *cse2*, *cas5*, *cas3*, *cas6*, *cse1*, *cas2*, and *cas7* genes. Often CRISPR regions flank cas elements. Only *E. piscicida* ACC35.1 and S11-285 genomes had confirmed CRISPR regions, whereas the rest of the *E. piscicida* genomes encoded only questionable CRISPR regions. Furthermore, only *E. ictaluri* LADL11-100 and LADL11-194 encoded questionable CRISPR elements, and rest of the *E. ictaluri* genomes did not carry any CRISPR elements.

3.5 Discussion

The goal of our study was to compare two fish pathogens, *E. ictaluri* and *E. piscicida*, by applying several different aspects of comparative genomics. We showed how these two pathogens are genetically distinguishable as species compared to all the available *Edwardsiella* genus members by using Pan-Core genome analysis, phylogenetic analysis, and ANI. Moreover, secretion systems and mobilome elements for each of these pathogens were evaluated. Additionally, orthology analysis of complete *E. ictaluri* and *E. piscicida* genomes was conducted.

ANI is the preferred digital tool for accurately estimating genome differences (Kim et al., 2014). ANI showed that all the *Edwardsiella* genus members are clearly separated from each other. Importantly, we discovered that some of the *Edwardsiella*

genus members are not categorized properly according to ANI. Strains such as ET883, FL6-60, and EIB202 were identified as *E. tarda* but should be classified as *E. piscicida*. Therefore, ANI is an effective tool to accurately classify genomes and correct mis-classifications. To be considered in the same species, ANI should typically be 95% or more. When compared to each other, *E. anguillarum* and *E. piscicida* genomes had ANI above 95%, making them the most closely related species in the genus. However, *E. anguillarum* are clearly distinguishable based on ANI, and although they don't strictly meet the ANI cutoff as a separate species, it is useful to classify these strains as a separate species based on host fish species affected.

Phylogenetic analysis based on the complete core genomes of 29 *Edwardsiella* strains (with 2 *Escherichia coli* strains as outgroups) confirmed that the five *Edwardsiella* species are clearly distinguishable from each other. Similar to ANI, the core genome comparison indicated that *E. anguillarum* and *E. piscicida* are the most closely related, and *E. ictaluri* is more closely related to these two species than *E. hoshinae* and *E. tarda*. It is interesting that phenotypically *E. tarda*, *E. piscicida*, and *E. anguillarum* are difficult to distinguish, and *E. ictaluri* is phenotypically distinct from these three species. However, genetically *E. ictaluri* is clearly more closely related to *E. piscicida* and *E. anguillarum* than *E. tarda* is.

Pan-core genome analysis is one of the main approaches to evaluate genomic diversity. Analysis of all the available sequenced *Edwardsiella* strains (as of 7/30/2017) revealed 9,436 genes in its pan-genome and 1,426 total genes in the core genome. Thus, gene acquisition and gene loss have significantly contributed to *Edwardsiella* diversification. In particular, *E. ictaluri* 93-146 has strong evidence of gene acquisition

with 268 singleton proteins, which are proteins not having any orthologs in any other genome in the analysis (Blom et al., 2009). By comparison, *E. piscicida* C07-087 has 47 singleton proteins in our analysis. These gene acquisitions do not appear to contribute significantly to host adaptation, however; only four of the *E. ictaluri* singletons and one *E. piscicida* singleton were identified as potential virulence proteins.

Although ANI indicated that *E. piscicida* and *E. anguillarum* are the closest related *Edwardsiella* species, *E. piscicida* and *E. ictaluri* are the most functionally similar species as revealed by orthology analysis. The six completely sequenced *E. ictaluri* and *E. piscicida* strains had 2,495 orthologous clusters. Many of the cluster differences between the two species were in the mobilome, suggesting that acquisition of mobile elements has significantly contributed to species differentiation between *E. ictaluri* and *E. piscicida* while functionally the two species have retained a high degree of similarity.

E. ictaluri had several unique features compared to the other *Edwardsiella* species. In particular, the *E. ictaluri* genomes contained urease operons that are homologous in structure to those in *Yersinia enterocolitica* 8081 (NC_008800). In bacteria, the urease operon typically encodes a multimeric enzyme composed of three different polypeptides. In *E. ictaluri*, *ureA* is a hypothetical fusion gene arising from *ureA* and *ureB* genes found in other bacteria, while *E. ictaluri ureB* is analogous to *ureC* in other species. Additionally, products of accessory genes are required for urease activation, and proteins encoded by *ureF*, *ureG*, and *ureD* are involved in transporting nickel ions and incorporating them into the active center of the urease apoenzyme (Konieczna et al., 2012). Urease catalyzes the hydrolysis of urea to give ammonia and carbon dioxide, thereby providing an important nitrogen source for many bacterial

species, and the enzyme can contribute to virulence of several Gram-negative bacteria by enhancing acid resistance (Mobley et al., 1995). The three *E. ictaluri* genomes encode cytochrome o ubiquinol oxidase subunits, which is the predominant enzyme needed for regulation of respiration associated with oxygen-rich growth conditions (Cotter et al., 1990). The *E. ictaluri* strains do not encode tetrathionate reductase operons, capsular polysaccharide synthesis enzymes, and vibrioferrin related genes that are present in the *E. piscicida* genomes, the significance of which needs further work to clarify.

Secretion systems and their effector proteins contribute to the pathogenicity of *Edwardsiella* genus members (Hou et al., 2017; Y. Liu et al., 2017; Zheng et al., 2007). However, each *Edwardsiella* species has different types and subtypes of secretion systems, revealing unique adaptations to specific environments. All of the evaluated *E. piscicida* and *E. ictaluri* strains encode T1SS, T2SS, T3SS, T5SS, and T6SS.

Encoding T4SS or some of the components of this system can give bacteria a significant advantage to adapt to environmental changes, mediate horizontal gene transfer, and potentially spread antimicrobial resistance genes (Wallden et al., 2010). Even though none of the evaluated *E. piscicida* genomes encode a complete T4SS, *E. piscicida* EIB2002 carries a plasmid (pEIB202) that encodes some of the T4SS elements, contributing to transfer of multi-drug resistance genes (Y. Liu et al., 2017). Some of the currently available *E. piscicida* genomes are in draft form; therefore, they have not been experimentally evaluated for carrying plasmids, and it is possible that other *E. piscicida* strains carry plasmids with T4SS. By contrast, four of the five currently sequenced *E. ictaluri* genomes encode T4SS (strain RUSVM). T4SS is known for translocating DNA

from other species; therefore, this system may contribute to the relatively high number of mobile elements in *E. ictaluri* compared to *E. piscicida* genomes.

Three types of T6SS have been described: type i, ii, and iii. Type i T6SS are sub-grouped into six categories, i1, i2, i3, i4a, i4b and i5, based on the conservation of core components (Barret et al., 2011; Barret et al., 2013; Russell et al., 2014). All of the currently evaluated *E. ictaluri* and *E. piscicida* genomes encode T6SS type i subtype i4b, but some *E. piscicida* strains (C07-087, S11-285, and ACC35.1) also encode a second T6SS: type I subtype i2. *E. piscicida* strains ET883, FL6-60, ETW41, JF1305, and EIB202 carry only T6SS-type I subtype i4b. Interestingly, *Edwardsiella anguillarim* encodes two distinct T3SS and three T6SS (Shao et al., 2015). Thus, *Edwardsiella* species vary considerably in T4SS and T6SS. Bacterial secretion systems are potential targets for development of new anti-virulence drugs to reduce bacterial pathogenicity (Baron et al., 2007). Therefore, further investigation of T4SS and different types of T6SS in *E. ictaluri* and *E. piscicida* is warranted.

Mobile elements contribute to horizontal gene transfer, genome organization, and plasticity (Varani et al., 2011). Three types of insertion sequences tend to be present in both *E. piscicida* and *E. ictaluri* genomes: IS*407*, IS*51*, and IS*3*. IS*6*, IS*21*, and IS*9I* families are present in *E. piscicida* genomes but not *E. ictaluri*. IS*256* is present in *E. ictaluri* genomes; in *Staphylococcus aureus*, IS*256* contributes to transposon-mediated antimicrobial resistance (Schreiber et al., 2013). In the panel of strains, we evaluated, *E. ictaluri* genomes had higher number of insertion sequences than *E. piscicida* genomes (Table 2). T4SS is one possible explanation because this system is responsible for DNA and protein uptake. However, this would not explain strain RUSVM, which does not

encode a complete or partial T4SS. In fact, the *E. ictaluri* RUSVM genome carries more insertion elements than any other evaluated *E. ictaluri* genome, especially in the IS*256* and IS*3* families. In *Aeromonas salmonicida*, a large number of temperature-sensitive insertion sequences appear to contribute to genomic stability of psychrophilic strains (Vincent et al., 2016). It is interesting to speculate that the increased number of insertion sequences in *E. ictaluri*, which grows at 30 °C but not 37 °C, relative to *E. piscicida* may contribute to its genome stability.

Genomic islands contribute to acquiring virulence genes, antimicrobial resistance genes, or genes that enable adaptation to a specific environment (Juhas et al., 2009). Some *E. tarda* strains encode specific pathogenicity islands that are homologous to the genomic islands of virulent *Escherichia coli* strains (Nakamura et al., 2013). In our analysis, the evaluated *E. ictaluri* genomes encode slightly more genomic islands than the *E. piscicida* genomes. Not surprisingly, a large number of the genes in these islands encode putative virulence factors; of the 431 protein-coding genes in *E. ictaluri* 93-146 genomic islands, 165 had significant identity with known virulence genes. *E. piscicida* C07-087 is similar; 88 of 292 protein-coding genes in genomic islands had significant similarity to known virulence proteins. This included 11 hypothetical genes and 43 phage-related genes. Thus, it is likely that genomic islands in these two species contribute significantly to pathogenesis of disease, and they warrant further investigation.

Due to the limited number of evaluated genomes for *E. ictaluri* and *E. piscicida*, we may not be able to generalize, but it seems that *E. ictaluri* strains tend to carry more phage regions. There are four prophages that integrated into the chromosome of *E. ictaluri* 93-146, whereas *E. piscicida* C07-87 genome encodes 5 phage regions.

Interestingly, only the *E. ictaluri* RUSVM-1 genome encodes 5 incomplete phage regions. It is possible that none of the *E. ictaluri* prophage regions are still active and able to excise to reproduce lytically; in one study, no temperate phage was induced by mitomycin C in 11 different *E. ictaluri* strains (Walakira et al., 2008). In addition to the prophage identified in the current study, *E. ictaluri* is susceptible to lysis by at least two types of bacteriophage that are specific to *E. ictaluri*, so the species may lack strong bacteriophage resistance mechanisms.

Integrons can play a significant role in spreading antimicrobial resistance genes (Cury et al., 2016). However, none were found in any of the *E. ictaluri* strains in the current study, and only two of the *E. piscicida* genomes (strains ETW41 and JF1305) encode a complete integron structure. Thus, integrons are not a major mechanism for chromosomally-encoded antimicrobial resistance in these species.

CRISPR-Cas systems are also not prominent in *E, ictaluri* and *E. piscicida*. Only *E. piscicida* ACC35.1 and S11-285 genomes encode Type I-E Cas system and complete CRISPR regions. Some other *E. piscicida* and *E. ictaluri* strains encode questionable CRISPR systems, but most *E. ictaluri* strains do not carry any. CRISPR-Cas systems are adaptable defense mechanisms used by many bacteria to resist predation by bacteriophage and exposure to plasmids (Horvath et al., 2010; Sander et al., 2014). This could contribute to *E. ictaluri* susceptibility to insertion elements and bacteriophage.

In summary, comparison of *E. ictaluri* and *E. piscicida* genomes provided valuable information about the biology of these species. Interestingly, although *E. anguillarum* and *E. piscicida* are the most closely related species in the genus, *E. piscicida* and *E. ictaluri* are the most functionally conserved. This could be the result of

convergent acquisition of similar host adaptation mechanisms, but more likely the high number of mobile elements suggests that the mobilomes of these species have driven genome diversification with retention of functional pathways. Importantly, for the first time, we report that *E. ictaluri* genomes encode T4SS, which could play a major role in the acquisition of mobile elements and pathogenicity of this species. *E. ictaluri* and *E. piscicida* also have strain variation in their T6SS, which could contribute to virulence and host adaptation. The relatively high number of mobile elements in *E. ictaluri* may be a result of its T4SS and lack of CRISPR-Cas systems, and it is possible that the insertion sequences contribute to genome stability and temperature restriction of *E. ictaluri*. Overall, this comparative genomics evaluation has identified important questions that warrant further study to elucidate the biology and virulence of these important fish pathogens

3.6 References

Abayneh, T., Colquhoun, D. J., & Sorum, H. (2012). Multi-locus Sequence Analysis (MLSA) of Edwardsiella tarda isolates from fish. *Vet Microbiol, 158*(3-4), 367-375. doi:10.1016/j.vetmic.2012.03.006

Abayneh, T., Colquhoun, D. J., & Sorum, H. (2013). Edwardsiella piscicida sp. nov., a novel species pathogenic to fish. *J Appl Microbiol, 114*(3), 644-654. doi:10.1111/jam.12080

Abby, S. S., Cury, J., Guglielmini, J., Neron, B., Touchon, M., & Rocha, E. P. (2016). Identification of protein secretion systems in bacterial genomes. *Sci Rep, 6*, 23080. doi:10.1038/srep23080

Abby, S. S., Neron, B., Menager, H., Touchon, M., & Rocha, E. P. (2014). MacSyFinder: a program to mine genomes for molecular systems with an application to CRISPR-Cas systems. *PLoS One, 9*(10), e110726. doi:10.1371/journal.pone.0110726

Abby, S. S., & Rocha, E. P. (2012). The non-flagellar type III secretion system evolved from the bacterial flagellum and diversified into host-cell adapted systems. *PLoS Genet, 8*(9), e1002983. doi:10.1371/journal.pgen.1002983

Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., & Wishart, D. S. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res, 44*(W1), W16-21. doi:10.1093/nar/gkw387

Baldwin, T. J., & Newton, J. C. (1993). Pathogenesis of Enteric Septicemia of Channel Catfish, Caused by Edwardsiella ictaluri: Bacteriologic and Light and Electron Microscopic Findings. *J Aquat Anim Health, 5*(3), 189-198. doi:10.1577/1548-8667(1993)005<0189:POESOC>2.3.CO;2

Baron, C., & Coombes, B. (2007). Targeting bacterial secretion systems: benefits of disarmament in the microcosm. *Infect Disord Drug Targets, 7*(1), 19-27.

Barret, M., Egan, F., Fargier, E., Morrissey, J. P., & O'Gara, F. (2011). Genomic analysis of the type VI secretion systems in Pseudomonas spp.: novel clusters and putative effectors uncovered. *Microbiology, 157*(Pt 6), 1726-1739. doi:10.1099/mic.0.048645-0

Barret, M., Egan, F., & O'Gara, F. (2013). Distribution and diversity of bacterial secretion systems across metagenomic datasets. *Environ Microbiol Rep, 5*(1), 117-126. doi:10.1111/j.1758-2229.2012.00394.x

Bertelli, C., Laird, M. R., Williams, K. P., Simon Fraser University Research Computing, G., Lau, B. Y., Hoad, G., Winsor, G.L., Brinkman, F. S. (2017). IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res*. doi:10.1093/nar/gkx343

Bi, D., Liu, L., Tai, C., Deng, Z., Rajakumar, K., & Ou, H. Y. (2013). SecReT4: a web-based bacterial type IV secretion system resource. *Nucleic Acids Res, 41*(Database issue), D660-665. doi:10.1093/nar/gks1248

Blom, J., Albaum, S. P., Doppmeier, D., Puhler, A., Vorholter, F. J., Zakrzewski, M., & Goesmann, A. (2009). EDGAR: a software framework for the comparative analysis of prokaryotic genomes. *BMC Bioinformatics, 10*, 154. doi:10.1186/1471-2105-10-154

Blom, J., Kreis, J., Spanig, S., Juhre, T., Bertelli, C., Ernst, C., & Goesmann, A. (2016). EDGAR 2.0: an enhanced software platform for comparative gene content analyses. *Nucleic Acids Res, 44*(W1), W22-28. doi:10.1093/nar/gkw255

Booth, N. J., Elkamel, A., & Thune, R. L. (2006). Intracellular Replication of Edwardsiella ictaluri in Channel Catfish Macrophages. *J Aquat Anim Health, 18*(2), 101-108. doi:10.1577/H05-025.1

Bujan, N., Lasa, A., A, E. T., & Magarinos, B. (2017). Draft Genome Sequence of the Fish Strain Edwardsiella tarda NCIMB 2034. *Genome Announc, 5*(20). doi:10.1128/genomeA.00359-17

Bujan, N., Toranzo, A. E., & Magarinos, B. (2017). Draft Genome Sequence of Edwardsiellapiscicida Strain ACC35.1 Isolated from Diseased Turbot (Scophthalmus maximus) in Europe. *Genome Announc, 5*(7). doi:10.1128/genomeA.01626-16

Chen, D. X., & Ainsworth, A. J. (1991). Assessment of metabolic activation of channel catfish peripheral blood neutrophils. *Dev Comp Immunol, 15*(3), 201-208.

Cotter, P. A., Chepuri, V., Gennis, R. B., & Gunsalus, R. P. (1990). Cytochrome o (cyoABCDE) and d (cydAB) oxidase gene expression in Escherichia coli is regulated by oxygen, pH, and the fnr gene product. *J Bacteriol, 172*(11), 6333-6338.

Crumlish, M., Dung, T. T., Turnbull, J. F., Ngoc, N. T. N., & Ferguson, H. W. (2002). Identification of Edwardsiella ictaluri from diseased freshwater catfish, Pangasius hypophthalmus (Sauvage), cultured in the Mekong Delta, Vietnam. *Journal of Fish Diseases, 25*(12), 733-736. doi:10.1046/j.1365-2761.2002.00412.x

Cury, J., Jove, T., Touchon, M., Neron, B., & Rocha, E. P. (2016). Identification and analysis of integrons and cassette arrays in bacterial genomes. *Nucleic Acids Res, 44*(10), 4539-4550. doi:10.1093/nar/gkw319

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res, 32*(5), 1792-1797. doi:10.1093/nar/gkh340

Ewing, W. H., Mcwhorter, A. C., Escobar, M. R., & Lubin, A. H. (1965). Edwardsiella, a new genus of Enterobacteriaceae based on a new species, E. tarda. *International Journal of Systematic and Evolutionary Microbiology, 15*(1), 33-38. doi:doi:10.1099/00207713-15-1-33

Felsenstein, J. (1989). PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics, 5*(164-166). doi:citeulike-article-id:2344765

Fogelson, S. B., Petty, B. D., Reichley, S. R., Ware, C., Bowser, P. R., Crim, M. J., Getchell, R.G., Sams, K.L., Marquis, H., Griffin, M. J. (2016). Histologic and molecular characterization of Edwardsiella piscicida infection in largemouth bass (Micropterus salmoides). *J Vet Diagn Invest, 28*(3), 338-344. doi:10.1177/1040638716637639

Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., & Tiedje, J. M. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol, 57*(Pt 1), 81-91. doi:10.1099/ijs.0.64483-0

Griffin, M. J., Reichley, S. R., Greenway, T. E., Quiniou, S. M., Ware, C., Gao, D. X., Gaunt, P. S., Yanong, R. P., Pouder D. B., Hawke, J.P., Soto, E. (2016). Comparison of Edwardsiella ictaluri isolates from different hosts and geographic origins. *J Fish Dis, 39*(8), 947-969. doi:10.1111/jfd.12431

Griffin, M. J., Ware, C., Quiniou, S. M., Steadman, J. M., Gaunt, P. S., Khoo, L. H., & Soto, E. (2014). Edwardsiella piscicida identified in the Southeastern USA by gyrB sequence, species-specific and repetitive sequence-mediated PCR. *Dis Aquat Organ, 108*(1), 23-35. doi:10.3354/dao02687

Grimont, P. A. D., Grimont, F., Richard, C., & Sakazaki, R. (1980). Edwardsiella hoshinae, a new species of enterobacteriaceae. *Current Microbiology, 4*(6), 347-351. doi:10.1007/bf02605375

Grissa, I., Vergnaud, G., & Pourcel, C. (2007). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res, 35*(Web Server issue), W52-57. doi:10.1093/nar/gkm360

Guglielmini, J., Neron, B., Abby, S. S., Garcillan-Barcia, M. P., de la Cruz, F., & Rocha, E. P. (2014). Key components of the eight classes of type IV secretion systems involved in bacterial conjugation or protein secretion. *Nucleic Acids Res, 42*(9), 5715-5727. doi:10.1093/nar/gku194

Hawke, J. P., Kent, M., Rogge, M., Baumgartner, W., Wiles, J., Shelley, J., Savolainen L. C., Wagner, R., Murray, K., Peterson, T. S. (2013). Edwardsiellosis caused by Edwardsiella ictaluri in laboratory populations of Zebrafish Danio rerio. *J Aquat Anim Health, 25*(3), 171-183. doi:10.1080/08997659.2013.782226

Hawke, J. P., Mcwhorter, A. C., Steigerwalt, A. G., & Brenner, D. J. (1981). Edwardsiella ictaluri sp. nov., the Causative Agent of Enteric Septicemia of Catfish. *International Journal of Systematic and Evolutionary Microbiology, 31*(4), 396-400. doi:doi:10.1099/00207713-31-4-396

Horvath, P., & Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science, 327*(5962), 167-170. doi:10.1126/science.1179555

Hou, M., Chen, R., Yang, D., Nunez, G., Wang, Z., Wang, Q., Zhang, Y., Liu, Q. (2017). Identification and functional characterization of EseH, a new effector of the type III secretion system of Edwardsiella piscicida. *Cell Microbiol, 19*(1). doi:10.1111/cmi.12638

Juhas, M., van der Meer, J. R., Gaillard, M., Harding, R. M., Hood, D. W., & Crook, D. W. (2009). Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev, 33*(2), 376-393. doi:10.1111/j.1574-6976.2008.00136.x

Katharios, P., Kokkari, C., Dourala, N., & Smyrli, M. (2015). First report of Edwardsiellosis in cage-cultured sharpsnout sea bream, Diplodus puntazzo from the Mediterranean. *BMC Vet Res, 11*, 155. doi:10.1186/s12917-015-0482-x

Kim, M., Oh, H. S., Park, S. C., & 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*(Pt 2), 346-351. doi:10.1099/ijs.0.059774-0

Konieczna, I., Zarnowiec, P., Kwinkowski, M., Kolesinska, B., Fraczyk, J., Kaminski, Z., & Kaca, W. (2012). Bacterial urease and its role in long-lasting human diseases. *Curr Protein Pept Sci, 13*(8), 789-806.

Konstantinidis, K. T., & Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci U S A, 102*(7), 2567-2572. doi:10.1073/pnas.0409727102

Li, J., Yao, Y., Xu, H. H., Hao, L., Deng, Z., Rajakumar, K., & Ou, H. Y. (2015). SecReT6: a web-based resource for type VI secretion systems found in bacteria. *Environ Microbiol, 17*(7), 2196-2202. doi:10.1111/1462-2920.12794

Liu, J. Y., Li, A. H., Zhou, D. R., Wen, Z. R., & Ye, X. P. (2010). Isolation and characterization of Edwardsiella ictaluri strains as pathogens from diseased yellow catfish Pelteobagrus fulvidraco (Richardson) cultured in China. *Aquaculture Research, 41*(12), 1835-1844. doi:10.1111/j.1365-2109.2010.02571.x

Liu, Y., Gao, Y., Liu, X., Liu, Q., Zhang, Y., Wang, Q., & Xiao, J. (2017). Transposon insertion sequencing reveals T4SS as the major genetic trait for conjugation transfer of multi-drug resistance pEIB202 from Edwardsiella. *BMC Microbiol, 17*(1), 112. doi:10.1186/s12866-017-1013-7

Liu, Y., Zhao, L., Yang, M., Yin, K., Zhou, X., Leung, K. Y., Liu, Q., Zhang, Y., Wang, Q. (2017). Transcriptomic dissection of the horizontally acquired response regulator EsrB reveals its global regulatory roles in the physiological adaptation and activation of T3SS and the cognate effector repertoire in Edwardsiella piscicida during infection toward turbot. *Virulence*, 1-23. doi:10.1080/21505594.2017.1323157

Miyazaki, T., & Plumb, J. A. (1985). Histopathology of Edwardsiella ictaluri in channel catfish, Ictalurus punctatus (Rafinesque)*. *Journal of Fish Diseases, 8*(4), 389-392. doi:10.1111/j.1365-2761.1985.tb00961.x

Mobley, H. L., Island, M. D., & Hausinger, R. P. (1995). Molecular biology of microbial ureases. *Microbiol Rev, 59*(3), 451-480.

Morrison, E. E., & Plumb, J. A. (1994). Olfactory Organ of Channel Catfish as a Site of Experimental Edwardsiella ictaluri Infection. *J Aquat Anim Health, 6*(2), 101-109. doi:10.1577/1548-8667(1994)006<0101:OOOCCA>2.3.CO;2

Nakamura, Y., Takano, T., Yasuike, M., Sakai, T., Matsuyama, T., & Sano, M. (2013). Comparative genomics reveals that a fish pathogenic bacterium Edwardsiella tarda has acquired the locus of enterocyte effacement (LEE) through horizontal gene transfer. *BMC Genomics, 14*(1), 1-14. doi:10.1186/1471-2164-14-642

Nusbaum, K. E., & Morrison, E. E. (1996). Communications: Entry of 35S-Labeled Edwardsiella ictaluri into Channel Catfish. *J Aquat Anim Health, 8*(2), 146-149. doi:10.1577/1548-8667(1996)008<0146:CEOSLE>2.3.CO;2

Oguro, K., Tamura, K., Yamane, J., Shimizu, M., Yamamoto, T., Ikawa, T., Ohnishi, K., Oshima, S., Imajoh, M. (2014). Draft Genome Sequences of Two Genetic Variant Strains of Edwardsiella piscicida, JF1305 and RSB1309, Isolated from Olive Flounder (Paralichythys olivaceus) and Red Sea Bream (Pagrus major) Cultured in Japan, Respectively. *Genome Announc, 2*(3). doi:10.1128/genomeA.00546-14

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A. R., Xia, F., Stevens, R. (2014). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res, 42*(Database issue), D206-214. doi:10.1093/nar/gkt1226

Plumb, J. A., & Quinlan, E. E. (1986). Survival of Edwardsiella ictaluri in Pond Water and Bottom Mud. *The Progressive Fish-Culturist, 48*(3), 212-214. doi:10.1577/1548-8640(1986)48.212:SOEIIP.2.0.CO;2

Plumb, J. A., & Sanchez, D. J. (1983). Susceptibility of five species of fish to Edwardsiella ictaluri. *Journal of Fish Diseases, 6*(3), 261-266. doi:10.1111/j.1365-2761.1983.tb00075.x

Reichley, S. R., Waldbieser, G. C., Lawrence, M. L., & Griffin, M. J. (2015). Complete Genome Sequence of an Edwardsiella piscicida-Like Species, Recovered from Tilapia in the United States. *Genome Announc, 3*(5). doi:10.1128/genomeA.01004-15

Reichley, S. R., Waldbieser, G. C., Lawrence, M. L., & Griffin, M. J. (2017). Complete Genome Sequence of Edwardsiella hoshinae ATCC 35051. *Genome Announc, 5*(6). doi:10.1128/genomeA.01605-16

Reichley, S. R., Waldbieser, G. C., Soto, E., Lawrence, M. L., & Griffin, M. J. (2017). Complete Genome Sequence of Edwardsiella ictaluri Isolate RUSVM-1 Recovered from Nile Tilapia (Oreochromis niloticus) in the Western Hemisphere. *Genome Announc, 5*(24). doi:10.1128/genomeA.00390-17

Reichley, S. R., Waldbieser, G. C., Tekedar, H. C., Lawrence, M. L., & Griffin, M. J. (2015). Complete Genome Sequence of Edwardsiella tarda Isolate FL95-01, Recovered from Channel Catfish. *Genome Announc, 3*(3). doi:10.1128/genomeA.00682-15

Reichley, S. R., Waldbieser, G. C., Tekedar, H. C., Lawrence, M. L., & Griffin, M. J. (2016). Complete Genome Sequence of Edwardsiella piscicida Isolate S11-285 Recovered from Channel Catfish (Ictalurus punctatus) in Mississippi, USA. *Genome Announc, 4*(6). doi:10.1128/genomeA.01259-16

Reichley, S. R., Waldbieser, G. C., Ucko, M., Colorni, A., Dubytska, L., Thune, R. L., Lawrence, M.L., Griffin, M. J. (2015). Complete Genome Sequence of an Edwardsiella piscicida-Like Species Isolated from Diseased Grouper in Israel. *Genome Announc, 3*(4). doi:10.1128/genomeA.00829-15

Richter, M., & Rossello-Mora, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A, 106*(45), 19126-19131. doi:10.1073/pnas.0906412106

Rogge, M. L., Dubytska, L., Jung, T. S., Wiles, J., Elkamel, A. A., Rennhoff, A., Oanh, D.T., Thune, R. L. (2013). Comparison of Vietnamese and US isolates of Edwardsiella ictaluri. *Dis Aquat Organ, 106*(1), 17-29. doi:10.3354/dao02620

Russell, A. B., Wexler, A. G., Harding, B. N., Whitney, J. C., Bohn, A. J., Goo, Y. A., Tran, B. Q., Barry, N. A., Zheng, H., Peterson, S. B., Chou, S., Gonen, T., Goodlett, D. R., Goodman, A. L., Mougous, J. D. (2014). A type VI secretion-related pathway in Bacteroidetes mediates interbacterial antagonism. *Cell Host Microbe, 16*(2), 227-236. doi:10.1016/j.chom.2014.07.007

Sakai, T., Kamaishi, T., Sano, M., Tensha, K., Arima, T., Iida, Y., Nagai, T., Nakai, T., Iida, T. (2008). Outbreaks of Edwardsiella ictaluri Infection in Ayu Plecoglossus altivelis in Japanese Rivers. *Fish Pathology, 43*(4), 152-157. doi:10.3147/jsfp.43.152

Sander, J. D., & Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol, 32*(4), 347-355. doi:10.1038/nbt.2842

Schreiber, F., Szekat, C., Josten, M., Sahl, H. G., & Bierbaum, G. (2013). Antibiotic-induced autoactivation of IS256 in Staphylococcus aureus. *Antimicrob Agents Chemother, 57*(12), 6381-6384. doi:10.1128/AAC.01585-13

Shafiei, S., Viljamaa-Dirks, S., Sundell, K., Heinikainen, S., Abayneh, T., & Wiklund, T. (2016). Recovery of Edwardsiella piscicida from farmed whitefish, Coregonus lavaretus (L.), in Finland. *Aquaculture, 454*, 19-26. doi:http://dx.doi.org/10.1016/j.aquaculture.2015.12.011

Shao, S., Lai, Q., Liu, Q., Wu, H., Xiao, J., Shao, Z., Wang, Q., Zhang, Y. (2015). Phylogenomics characterization of a highly virulent Edwardsiella strain ET080813(T) encoding two distinct T3SS and three T6SS gene clusters: Propose a novel species as Edwardsiella anguillarum sp. nov. *Syst Appl Microbiol, 38*(1), 36-47. doi:10.1016/j.syapm.2014.10.008

Shoemaker, C. A., Klesius, P. H., & Plumb, J. A. (1997). Killing of Edwardsiella ictaluri by macrophages from channel catfish immune and susceptible to enteric septicemia of catfish. *Veterinary Immunology and Immunopathology, 58*(2), 181-190. doi:http://dx.doi.org/10.1016/S0165-2427(97)00026-3

Soto, E., Griffin, M., Arauz, M., Riofrio, A., Martinez, A., & Cabrejos, M. E. (2012). Edwardsiella ictaluri as the Causative Agent of Mortality in Cultured Nile Tilapia. *J Aquat Anim Health, 24*(2), 81-90. doi:10.1080/08997659.2012.675931

Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., Lomsadze, A., Pruitt, K. D., Borodovsky, M., Ostell, J. (2016). NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res, 44*(14), 6614-6624. doi:10.1093/nar/gkw569

Tekedar, H. C., Karsi, A., Williams, M. L., Vamenta, S., Banes, M. M., Duke, M., Scheffler, B. E., Lawrence, M. L. (2013). Complete Genome Sequence of Channel Catfish Gastrointestinal Septicemia Isolate Edwardsiella tarda C07-087. *Genome Announc, 1*(6). doi:10.1128/genomeA.00959-13

van Soest, J. J., Stockhammer, O. W., Ordas, A., Bloemberg, G. V., Spaink, H. P., & Meijer, A. H. (2011). Comparison of static immersion and intravenous injection systems for exposure of zebrafish embryos to the natural pathogen Edwardsiella tarda. *BMC Immunol, 12*, 58. doi:10.1186/1471-2172-12-58

Varani, A. M., Siguier, P., Gourbeyre, E., Charneau, V., & Chandler, M. (2011). ISsaga is an ensemble of web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. *Genome Biol, 12*(3), R30. doi:10.1186/gb-2011-12-3-r30

Vincent, A. T., Trudel, M. V., Freschi, L., Nagar, V., Gagne-Thivierge, C., Levesque, R. C., & Charette, S. J. (2016). Increasing genomic diversity and evidence of constrained lifestyle evolution due to insertion sequences in Aeromonas salmonicida. *BMC Genomics, 17*, 44. doi:10.1186/s12864-016-2381-3

Walakira, J. K., Carrias, A. A., Hossain, M. J., Jones, E., Terhune, J. S., & Liles, M. R. (2008). Identification and characterization of bacteriophages specific to the catfish pathogen, Edwardsiella ictaluri. *J Appl Microbiol, 105*(6), 2133-2142. doi:10.1111/j.1365-2672.2008.03933.x

Wallden, K., Rivera-Calzada, A., & Waksman, G. (2010). Type IV secretion systems: versatility and diversity in function. *Cell Microbiol, 12*(9), 1203-1212. doi:10.1111/j.1462-5822.2010.01499.x

Wang, Q., Yang, M., Xiao, J., Wu, H., Wang, X., Lv, Y., Xu, L., Zheng, H., Wang, S., Zhao, G., Liu, Q., Zhang, Y. (2009). Genome sequence of the versatile fish pathogen Edwardsiella tarda provides insights into its adaptation to broad host ranges and intracellular niches. *PLoS One, 4*(10), e7646. doi:10.1371/journal.pone.0007646

Wang, R., Tekedar, H. C., Lawrence, M. L., Chouljenko, V. N., Kim, J., Kim, N., Kousoulas, K. G., Hawke, J. P. (2015). Draft Genome Sequences of Edwardsiella ictaluri Strains LADL11-100 and LADL11-194 Isolated from Zebrafish Danio rerio. *Genome Announc, 3*(6). doi:10.1128/genomeA.01449-15

Wang, Y., Coleman-Derr, D., Chen, G., & Gu, Y. Q. (2015). OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res, 43*(W1), W78-84. doi:10.1093/nar/gkv487

Williams, M. L., Gillaspy, A. F., Dyer, D. W., Thune, R. L., Waldbieser, G. C., Schuster, S. C., Gipson, J., Zaitshik, J., Landry, C., Banes, M. M., Lawrence, M. L. (2012). Genome sequence of Edwardsiella ictaluri 93-146, a strain associated with a natural channel catfish outbreak of enteric septicemia of catfish. *J Bacteriol, 194*(3), 740-741. doi:10.1128/JB.06522-11

Yang, M., Lv, Y., Xiao, J., Wu, H., Zheng, H., Liu, Q., Zhang, Y., Wang, Q. (2012). Edwardsiella comparative phylogenomics reveal the new intra/inter-species taxonomic relationships, virulence evolution and niche adaptation mechanisms. *PLoS One, 7*(5), e36987. doi:10.1371/journal.pone.0036987

Yao, J., Li, C., Zhang, J., Liu, S., Feng, J., Wang, R., Li, Y., Jiang, C., Song, L., Chen, A., Liu, Z. (2014). Expression of nitric oxide synthase (NOS) genes in channel catfish is highly regulated and time dependent after bacterial challenges. *Dev Comp Immunol, 45*(1), 74-86. doi:10.1016/j.dci.2014.02.005

Yeh, H. Y., & Klesius, P. H. (2013). Changes of serum myeloperoxidase and nitric oxide in the early stage of Edwardsiella ictaluri infection in channel catfish, Ictalurus punctatus (Rafinesque). *J Fish Dis, 36*(4), 441-446. doi:10.1111/jfd.12038

Yuasa, K., Kholidin, E. B., Panigoro, N., & Hatai, K. (2003). First Isolation of Edwardsiella ictaluri from Cultured Striped Catfish Pangasius hypophthalmus in Indonesia. *Fish Pathology, 38*(4), 181-183. doi:10.3147/jsfp.38.181

Zheng, J., & Leung, K. Y. (2007). Dissection of a type VI secretion system in Edwardsiella tarda. *Mol Microbiol, 66*(5), 1192-1206. doi:10.1111/j.1365-2958.2007.05993.x

Zhou, C. E., Smith, J., Lam, M., Zemla, A., Dyer, M. D., & Slezak, T. (2007). MvirDB--a microbial database of protein toxins, virulence factors and antibiotic resistance genes for bio-defence applications. *Nucleic Acids Res, 35*(Database issue), D391-394. doi:10.1093/nar/gkl791

CHAPTER IV

COMPARATIVE ANALYSIS OF SECRETION SYSTEMS OF VIRULENT *AEROMONAS HYDROPHILA* STRAINS AND MUTATION OF *tssD* AND *tssI* GENES FROM T6SS

4.1 Abstract

Virulent *Aeromonas hydrophila* (vAh) causes severe motile *Aeromonas* septicemia (MAS) in warm water fishes. In recent years, channel catfish farming in the USA and carp farming in China have been affected by virulent *A. hydrophila*. Bacterial secretion systems are often important virulence factors. However, our knowledge on secretion systems among various *A. hydrophila* strains remains limited. Sequencing of different *A. hydrophila* genomes by our group and others have provided an opportunity for comprehensive analysis of *A. hydrophila* secretion systems. Thus, in this study, we conducted comparative secretion system analysis using 55 *A. hydrophila* genomes, including virulent *A. hydrophila* (vAh). Previous results showed that vAh strains from the southeastern U.S. and China belong to the same clonal group. Interestingly, distribution of secretion systems in this clonal group is different from other *A. hydrophila* genomes. For example, tight adherence (TaD) system present consistently in all the vAH. The majority of the U.S. isolates do not possess a complete type 6 secretion system (T6SS),

but three core elements (*tssD*, *tssH*, *tssI*) are encoded. On the other hand, Chinese isolates have a complete T6SS operon. None of the vAh isolates have a type 3 secretion system (T3SS), but they possess other secretion systems including type 1 secretion system (T1SS), type 2 secretion system (T2SS), and tight adherence (TaD). Mutation of genes encoding two T6SS proteins from vAh isolate *A. hydrophila* ML09-119 showed that deletion of the *hcpA1* and *vgrG* genes reduced vAh virulence 2.24 fold in catfish fingerlings compared to the parent strain ML09-119. We expect that these findings will improve our understanding of *A. hydrophila* secretion systems and their evolution in *A. hydrophila* genomes.

4.2 Introduction

Aeromonas hydrophila is present in freshwater environments, causing disease in fish, reptiles, amphibians, and humans (Janda and Abbott, 2010; Tomas, 2012). The U.S. and China aquaculture industries have had significant losses due to *A. hydrophila* outbreaks (Nielsen et al., 2001). In the southeastern U.S., these outbreaks impacting the catfish aquaculture industry started in 2009 and are caused by a clonal group of virulent *A. hydrophila* (vAh) strains (Hemstreet, 2010; Hossain et al., 2014).

Comparative genomics methods have helped to identify taxonomically mislabeled *A. hydrophila* genomes in Genbank (Beaz-Hidalgo et al., 2015). The same methods also revealed that the U.S. vAh clonal lineage is similar to a clonal lineage of *A. hydrophila* that is responsible for significant economic losses in the Chinese aquaculture industry (Griffin et al., 2013; Hossain et al., 2013; Zhang et al., 2014; Pang et al., 2015). Both clonal groups are categorized as sequence type ST251 (Rasmussen-Ivey et al., 2016). It

has been theorized that the U.S. vAh originated from transport of carrier fish from Asia (Hossain et al., 2014).

Comparative genomics has revealed that the vAh clonal group has unique characteristics. Some of the unique biochemical pathways include sialic acid biosynthesis, *myo*inositol utilization, and L-fucose metabolism. They also have unique O-antigen biosynthesis and characteristic mobilome elements and secretion systems (Hossain et al., 2013; Pang et al., 2015). Intriguingly, Asian vAh isolates encode all the core components of type 6 secretion system (T6SS), whereas most of the U.S. vAh isolates carry remnants of the T6SS (Rasmussen et al., 2016).

Several virulence mechanisms of *A. hydrophila* including secretion systems, motility and flagella, toxins, tissue-destructive enzymes, iron acquisition, and S-layer have been studied (Tomas, 2012). Secretion systems are used by bacteria to interact with the environment, including host adaptation and competing against other bacteria (Cianfanelli et al., 2016). A thorough investigation of secretion systems distribution in *A. hydrophila*, including vAh, other fish disease strains, and environmental strains, has not been conducted. Hence, in this study, we analyzed 55 *A. hydrophila* genomes from distinct geographical origins and hosts. We also evaluated type 4 pilus (T4P), tight adherence system (Tad), and flagellum components due to their homologous similarity to secretions systems. Potential host-pathogen interactions of the identified secretion system proteins were evaluated, and potential for having a role in virulence was evaluated by searching against the Microbial Virulence Database (MVirDB).

T6SS is known as a versatile nanomachine delivering effector proteins and toxins to host cells or competitor bacteria. T6SS is categorized into three phylogenetic subtypes:

T6SSi, T6SSii, and T6SSiii (Russell et al., 2014b). Our analysis showed that all the evaluated *A. hydrophila* genomes encode the whole operon or remnants of T6SSi. vAh isolate ML09-119, from catfish in the southeast U.S., has T6SSi genes *tssD* (also referred to as *hcpA*) and *tssI* (also referred to as *vgrG*), which are found either on a single T6SS operon or located outside of the operon. These are not only essential components of the T6SS but also play a role in activation of the system; thus, they may have a dual function as a core component and as substrates (Russell et al., 2014a). To clarify the function of these T6SSi genes in vAh, we conducted mutation of the *hcpA1* and *vgrG* genes and determined they have reduced virulence in catfish fingerlings compared to the parent strain ML09-119. Overall, the comparative genomics and mutational analyses reported here clarify the distribution of various secretion systems in *A. hydrophila* and provide functional information on the role of T6SS components in vAh.

4.3 Materials and methods

4.3.1 Genome sequencing, genomes from sequence database and annotation

The genome sequences (including complete sequences, draft assemblies, and raw reads) of 55 *A. hydrophila* strains were downloaded from the National Center for Biotechnology Information (NCBI) (Table 4.1). Raw data were assembled using CLC workbench 6.5.1 after trimming sequence reads, followed by error correction and contig creation. All unannotated genomes were annotated by RAST (Brettin et al., 2015). All selected genomes had at least 95% average nucleotide identity (ANI).

Table 4.1 *A. hydrophila* genomes used in comparative genomic analyses.

#	Species	Strain	Location	Source	Accession	Reference
1	*A. hydrophila*	Arkansas 2010	USA	Catfish	NZ_LYZH00000000.1	(Tekedar et al., 2017)
2	*A. hydrophila*	ML09-119	USA	Catfish	NC_021290.1	(Tekedar et al., 2013)
3	*A. hydrophila*	ML09-122	USA	Catfish	NZ_LRRY00000000.1	(Tekedar et al., 2016b)
4	*A. hydrophila*	ML09-121	USA	Catfish	NZ_LRRX00000000.1	(Tekedar et al., 2016b)
5	*A. hydrophila*	AL10-121	USA	Catfish	NZ_LRRW00000000.1	(Tekedar et al., 2016b)
6	*A. hydrophila*	AL09-71	USA	Catfish	NZ_CP007566.1	(Pridgeon et al., 2014b)
7	*A. hydrophila*	pc104A	USA	Soil	NZ_CP007576.1	(Pridgeon et al., 2014a)
8	*A. hydrophila*	S14-296	USA	Catfish	SAMN05292365	(Rasmussen-Ivey et al., 2016)
9	*A. hydrophila*	S14-606	USA	Catfish	SAMN05292366	(Rasmussen-Ivey et al., 2016)
10	*A. hydrophila*	S13-612	USA	Catfish	SAMN05292362	(Rasmussen-Ivey et al., 2016)
11	*A. hydrophila*	S13-700	USA	Catfish	SAMN05292363	(Rasmussen-Ivey et al., 2016)
12	*A. hydrophila*	Ahy_Idx7_1	USA	Catfish	SAMN05292361	(Rasmussen-Ivey et al., 2016)
13	*A. hydrophila*	ALG15-098	USA	Catfish	SAMN05223361	(Rasmussen-Ivey et al., 2016)
14	*A. hydrophila*	IPRS-15-28	USA	Catfish	SAMN05223362	(Rasmussen-Ivey et al., 2016)
15	*A. hydrophila*	ML10-51K	USA	Catfish	SAMN05223363	(Rasmussen-Ivey et al., 2016)
16	*A. hydrophila*	S14-458	USA	Catfish	SAMN05223364	(Rasmussen-Ivey et al., 2016)
17	*A. hydrophila*	S15-130	USA	Catfish	SAMN05223365	(Rasmussen-Ivey et al., 2016)
18	*A. hydrophila*	S15-400	USA	Catfish	SAMN05223367	(Rasmussen-Ivey et al., 2016)
19	*A. hydrophila*	ZC1	USA	Gras Carp	SAMN02404465	(Hossain et al., 2014)
20	*A. hydrophila*	AL09-79	USA	Catfish	NZ_LRRV00000000.1	(Tekedar et al., 2016b)
21	*A. hydrophila*	2JBN101	China	Crucian carp	NZ_LXME00000000.1	(Zhang et al., 2013)
22	*A. hydrophila*	D4	China	Megalobrama amblycephala	NZ_CP013965.1	(Tran et al., 2015)
23	*A. hydrophila*	JBN2301	China	Carp	NZ_CP013178.1	(Yang et al., 2016)
24	*A. hydrophila*	S15-591	USA	Catfish	SAMN05223368	(Rasmussen-Ivey et al., 2016)
25	*A. hydrophila*	J-1	China	Carp	NZ_CP006883.1	(Pang et al., 2015)
26	*A. hydrophila*	NJ-35	China	Carp	NZ_CP006870.1	(Pang et al., 2015)
27	*A. hydrophila*	GYK1	China	Siniperca chuatsi	NZ_CP016392.1	(Pan et al., 2004)
28	*A. hydrophila*	SNUFPC-A8	S. Korea	Salmon	NZ_AMQA00000000.1	(Han et al., 2013)
29	*A. hydrophila*	NF1	USA	Clinical	NZ_JDWB00000000.1	(Grim et al., 2014)
30	*A. hydrophila*	Ae34	Sri Lanka	Carp	NZ_BAXY00000000.1	(Jagoda et al., 2014)
31	*A. hydrophila*	M052	Malaysia	Waterfall	NZ_MAKI00000000.1	N/A
32	*A. hydrophila*	M053	Malaysia	Waterfall	NZ_MAKJ00000000.1	N/A
33	*A. hydrophila*	M062	Malaysia	Waterfall	NZ_JSXE00000000.1	(Chan et al., 2015)
34	*A. hydrophila*	AHNIH1	USA	Clinical	NZ_CP016380.1	(Hughes et al., 2016)
35	*A. hydrophila*	AL06-06	USA	Goldfish	NZ_CP010947.1	(Tekedar et al., 2015)

Table 4.1 (continued)

36	*A. hydrophila*	ATCC 7966	USA	Milk tin	NC_008570.1	(Seshadri et al., 2006)
37	*A. hydrophila*	AL97-91	USA	Tilapia	NZ_CM004591.1	(Tekedar et al., 2017)
38	*A. hydrophila*	MN98-04	USA	Tilapia	NZ_CM004592.1	(Tekedar et al., 2017)
39	*A. hydrophila*	AH-1	Canada	Moribund fish	NZ_LYXN00000000.1	(Forn-Cuni et al., 2016)
40	*A. hydrophila*	RB-AH	Malaysia	Soil	NZ_JPEH00000000.1	(Emond-Rheault et al., 2015)
41	*A. hydrophila*	NF2	USA	Clinical	NZ_JDWC00000000.1	(Grim et al., 2014)
42	*A. hydrophila*	S14-230	USA	Catfish	SAMN05292364	(Rasmussen-Ivey et al., 2016)
43	*A. hydrophila*	48_AHYD	USA	Clinical	NZ_JVFM00000000.1	(Roach et al., 2015)
44	*A. hydrophila*	53_AHYD	USA	Clinical	NZ_JVDL00000000.1	(Roach et al., 2015)
45	*A. hydrophila*	56_AHYD	USA	Clinical	NZ_JVCD00000000.1	(Roach et al., 2015)
46	*A. hydrophila*	52_AHYD	USA	Clinical	NZ_JVDW00000000.1	(Roach et al., 2015)
47	*A. hydrophila*	50_AHYD	USA	Clinical	NZ_JVES00000000.1	(Roach et al., 2015)
48	*A. hydrophila*	AH10	China	Grass carp	NZ_CP011100.1	(Xu et al., 2013)
49	*A. hydrophila*	TN-97-08	USA	Bluegill	NZ_LNUR00000000.1	(Tekedar et al., 2016a)
50	*A. hydrophila*	FDAARGOS_78	USA	Clinical	JTBD01000000	N/A
51	*A. hydrophila*	226	Malaysia	Clinical	NZ_JEML00000000.1	(Chan et al., 2011)
52	*A. hydrophila*	M013	Malaysia	Waterfall	NZ_JRWS00000000.1	(Tan et al., 2015a)
53	*A. hydrophila*	AD9	USA	Wetland sediment	NZ_JFJO00000000.1	(Lenneman and Barney, 2014)
54	*A. hydrophila*	M023	Malaysia	Waterfall	NZ_JSWA00000000.1	(Tan et al., 2015b)
55	*A. hydrophila*	Ranae CIP 107985	USA	Fish/Ranae	NZ_CDDC00000000.1	(Colston et al., 2014)

N/A: Not available

4.3.2 Phylogenetic tree creation

A phylogenetic tree was built from the complete core genomes of 55 *A. hydrophila* strains, which included 115,335 coding sequences (2,097/genome) with 101,851,090 amino acid residues (1,851,838/genome). The gene sets of the core genome were aligned one by one using MUSCLE (Edgar, 2004), and alignments were concatenated. This alignment was used to compute a Kimura distance matrix, which was used as input for the Neighbor-Joining algorithm as implemented in PHYLP (Felsenstein, 1989). The resulting tree was verified by bootstrapping with 250 iterations.

4.3.3 ANI and AAI calculation

ANI and average amino acid identity (AAI) values (Konstantinidis and Tiedje, 2005b; a; Konstantinidis et al., 2006) were calculated using EDGAR (Konstantinidis and Tiedje, 2005b). Briefly, the average amino acid identities were based on all gene sets of the core genome (2097 per genome). Mean identity values were extracted from BLASTP (Altschul et al., 1990) results using EDGAR. ANI was calculated as ANIb values based on BLASTN as described (Goris et al., 2007) using the same cutoffs as JSpeciesWS (Richter and Rossello-Mora, 2009).

4.3.4 Secretion systems identification

MacSyfinder was used with default features to identify secretion systems from the *A. hydrophila* genomes. The "unordered" type of dataset option was chosen because the majority of the evaluated genomes were draft genomes. The topology of the replicon was linear/circular, maximal e-value was 1.0, maximal independent E-value was 0.001, and minimal profile coverage was 0.5. Both mandatory genes and accessory genes were identified (Abby et al., 2014; Abby et al., 2016).

4.3.5 Virulence database

Secretion systems elements identified by MacyFinder were evaluated for their potential as virulence factors. To do so, Microbial Virulence Database (MVirDB) was downloaded (http://mvirdb.llnl.gov/) (Zhou et al., 2007), and a searchable database was created using CLC Workbench (version 6.5). Local BLAST was conducted, and matches were accepted when E-value < $1x10^{-10}$. Because different strains carried different secretion systems, T1SS, T2SS, T4P and Tad system proteins were from *A. hydrophila*

ML09-119 strain; T3SS and flagellum proteins were from *A. hydrophila* MN98-04, and T6SS proteins were from *A. hydrophila* TN 97-08 strain. Additionally, T6SS mutants in the *A. hydrophila* ML09-119 genome were evaluated separately to determine their potential as a virulence factor. When identified secretion systems components were encoded more than once, only one of them was included in the virulence database search analysis.

4.3.6 Bacterial strains and plasmids

Bacterial strains and plasmids used in this study are listed in Table 4.2. *A. hydrophila* strain ML09-119 represents the vAh clonal group impacting U.S. channel catfish aquaculture. The strain was grown on brain heart infusion (BHI) agar or broth (Difco, Sparks, MD, USA) and incubated at 37 °C. *Escherichia coli* strain CC118 λ*pir* was used for cloning, and strain BW19851 was used for transferring suicide plasmid pMEG-375 into *A. hydrophila* by conjugation. *E. coli* strains were cultured in Luria–Bertani (LB) agar and broth (Difco) and incubated at 37 °C. The following antibiotics and reagents (Sigma-Aldrich, Saint Louis, MN, USA) were used when needed: ampicillin (100 µg/ml), chloramphenicol (10-25 µg/ml), colistin (12.5 µg/ml), sucrose (5%), and mannitol (0.35%).

Table 4.2 Bacterial strains and plasmids used in the present study.

Strain or plasmid	Description	Source
A. hydrophila ML09-119	Isolate from a disease outbreak on a commercial catfish farm	(Griffin et al., 2013)
*vAh*Δ*hcpA1*	*A. hydrophila* ML09-119 derivative; Δ*hcpA1*	This study
*vAh*Δ*vgrG1*	*A. hydrophila* ML09-119; Δ*vgrG*	This study
E. coli		
CC118λ*pir*	Δ(*ara-leu*); *araD*; Δ*lacX74*; *galE*; *galK*; *phoA20*; *thi-1*; *rpsE*; *rpoB*; *argE*(Am); *recA1*; λ*pir*R6K	(Herrero et al., 1990)
BW19851	*RP4-2 (Km::Tn7, Tc::Mu-1), DuidA3::pir+, recA1, endA1, thi-1, hsdR17, creC510*	(Metcalf et al., 1994)
Plasmid		
pMEG-375	8,142 bp, Amp^r, Cm^r, *lacZ*, R6K *ori*, *mob incP*, *sacR sacB*	(Dozois et al., 2003)
pAhΔ*hcpA1*	10,173 bp, Δ*hcpA1*, pMEG-375	This study
pAhΔ*vgrG1*	10,160 bp, Δ*vgrG1*, pMEG-375	This study

4.3.7 In-frame deletion of *A. hydrophila* genes

Two chromosomal in-frame deletion mutants of type six secretion system (T6SS) effector genes, *hcpA1* (AHML_05970) and *vgrG1* (AHML_05975), were constructed by allelic exchange and homologous recombination using suicide plasmid pMEG-375 containing the counter-selectable marker *sacB* (Dozois et al., 2003). Recombinant DNA and mutant construction procedures were completed as described previously (Abdelhamed et al., 2013). Briefly, four primers (A, B, C, and D) were designed for each gene using Primer3 (Untergasser et al., 2012) (Table 4.3). Compatible restriction enzyme sites were embedded in A and D primers (bold line in primers A and D) for cloning, and the reverse complement of primer B was added to the 5′ end of primer C (underlined letters in primers C) to allow fusion of PCR fragments by overlap extension PCR (Horton et al., 1989). The upstream (fragment AB) and downstream (fragment CD) of each gene was amplified using two sets of primers. PCR fragments AB and CD were annealed at the overlapping regions and were amplified as a single fragment

using primers A and D. The fusion products were purified, digested, ligated into digested pMEG-375, electroporated into *E. coli* CC118λ*pir*, and spread on LB agar plus ampicillin. The resulting plasmids were transferred into *A. hydrophila* ML09-119 by conjugation using *E. coli* BW19851. Transconjugants were selected on plates containing chloramphenicol and colistin; chloramphenicol was used to select the integration of pMEG-375 in *A. hydrophila* chromosome while colistin was used as counter selection against *E. coli*. PCR analysis confirmed that the vector had integrated correctly into the chromosomal DNA. After sucrose treatment, transconjugants that were colistin resistant and chloramphenicol sensitive were selected, and the deletion was confirmed by colony PCR using A and D primers. Mutant validation was done by sequencing of AD fragments amplified from chloramphenicol sensitive mutants using hcpA1Seq and vgrG1Seq primers (Table 4.3). The *A. hydrophila* mutants were designated *vAhΔhcpA1* and *vAhΔvgrG1*.

Table 4.3 Primers used to generate and verify in-frame deletion of vAh genes.

	Primer ID		**Sequence 5-3'**	**RE**
hcpA1	hcpA1F01	A	AAA**TCTAGA**TCCTATGTGCCTGAGTGTGC	*Xba*I
	hcpA1R1000	B	AATGACACTCGACCAAACCA	
	hcpA1F1000	C	TGGTTTGGTCGAGTGTCATTGAGGCCTAACGCTCGATCT	
	hcpA1R01	D	AAA**GAGCTC**AGGTCGGTTTCCCGGTACT	*Sac*I
	HcpA1Seq		GCTGGCTCTCCATGCATATT	
vgrG1	vgrG1F01	A	AAA**TCTAGA**AAGGTAAAACCCAGGGCAAT	*Xba*I
	vgrG1R1000	B	TGTGCTGTCTGCCATGAAG	
	vgrG1F1000	C	CTTCATGGCAGACAGCACACGACTGATTGAGGTTTCCGTA	
	vgrG1R01	D	AAA**GAGCTC**CAGGCTGGTGTCTCGATTTT	*Sac*I
	vgrG1Seq		GCAAAGCACAACAGAGGCTA	

Bold letters at the 5′ end of the primer sequence represent restriction enzymes (RE) site added. AAA nucleotides were added to the end of each primer containing a RE site. Underlined bases in primer C indicate reverse complemented primer B sequence.

4.3.8 Virulence and efficacy of the vAh mutants in catfish fingerlings

All fish experiments were conducted under a protocol approved by the Institutional Animal Care and Use Committee at Mississippi State University. Virulence of *vAhΔhcpA1* and *vAhΔvgrG1* was compared to *A. hydrophila* wild-type (WT) strain ML09-119 by immersion route of exposure as described (Abdelhamed et al., 2016). Briefly, 120 six-month-old specific-pathogen-free (SPF) channel catfish fingerlings (18.10 ± 0.56 cm, 50.90± 3.76 g) were stocked into twelve 40-liter flow-through tanks (10 fish/tank) and acclimated for a week. Tanks were assigned randomly to four treatment groups: *vAhΔhcpA1*, *vAhΔvgrG1*, vAh WT (positive control for virulence), and BHI (negative control for virulence and sham control for vaccination). Each group included three replicate tanks. Water temperature was maintained at 32ºC (±2) throughout the experiments. Fish were fed twice a day with a commercial catfish feed. On the challenge day, the water levels in each tank were decreased to 10 L, and 100 mL of overnight cultures was added directly to each tank (1.02×10^{10} CFU/mL water). Negative control tanks were exposed to 100 mL of sterile BHI broth. During immersion, water was well aerated. After six hours, water flow was restored, and fish were maintained as usual. Fish mortalities were recorded daily for a total of 21 days, and percent mortalities was calculated for each group. Protection against vAh WT challenge was determined in surviving fingerlings vaccinated by the *vAhΔhcpA1* and *vAhΔvgrG1* mutants. Briefly, after 21 days post-vaccination, catfish fingerlings were re-challenged by vAh WT by immersion (2.21×10^{10} CFU/ml water), and mortalities were recorded daily for 14 days. At the end of the experiments, mean percent survival was calculated for each treatment.

4.3.9 Statistical analysis

Mean percent mortality data were arcsine transformed, and analysis of variance (ANOVA) was applied using PROC GLM in SAS for Windows v9.4 (SAS Institute, Inc., Cary, NC) to assess significance. An alpha level of 0.05 was used in all analyses.

4.3.10 Host-pathogen interaction network

Protein-protein interactions between *A. hydrophila* secretion system proteins and catfish proteins (accession: PRJNA281269) were determined using the Host-Pathogen Interaction Database (Ammari et al., 2016). For pathogen sequences, default upload options were: database search: bacterial pathogens, matrix: Blosum62, E-value: 0.00001, pathogen percent identity: 30, and query coverage filter: 50%. For host sequences, selected animal protein options were: for the database search matrix: Blosum62, E-value cutoff: 0.00001, percent identity and query coverage filter: 70% (Ammari et al., 2016).

4.4 Results

4.4.1 Genome features

The 55 genome sequences included in the current study are either vAh or non-vAh isolates from different geographical locations and hosts (Table 4.1). Of these, we sequenced vAh strains ML09-119, ML09-121, ML09-122, AL09-79, AL10-121, and Arkansas 2010. We also sequenced *A. hydrophila* strains AL06-06, AL97-91, MN98-04, and TN97-08. Additionally, we included 12 recent vAh strains' genomes (strains Ahy_Idx71, ALG15-098, IPRS15-28, ML10-51K, S13-612, S13-700, S14-296, S14-458, S14-606, S15-130, S15-400, and S15-591) and one non-vAh strain genome (S14-230) in

our analysis. Genome size of the 55 strains ranged from approximately 4.67 Mb to 5.28 Mb, and GC ratio of the genomes ranged from 60.47 to 61.60.

4.4.2 ANI and phylogenetic tree creation

A phylogenetic tree based on the complete core genome of 55 *A. hydrophila* strains shows the 27 vAh strains forming a highly conserved branch separated clearly from the other strains. The conservation of the tree topology within the cluster of 27 vAh strains was not convincing with branch conservation values down to 15%. The separation of the vAh cluster from the rest of the tree showed 100% branch conservation. These findings were confirmed by Average Nucleotide Identity (ANI) as well as Average Amino Acid Identity (AAI). ANI and AAI values within the cluster of 27 strains were above 99.88% (ANI) and 99.89% (AAI), respectively (Figure 4.1).

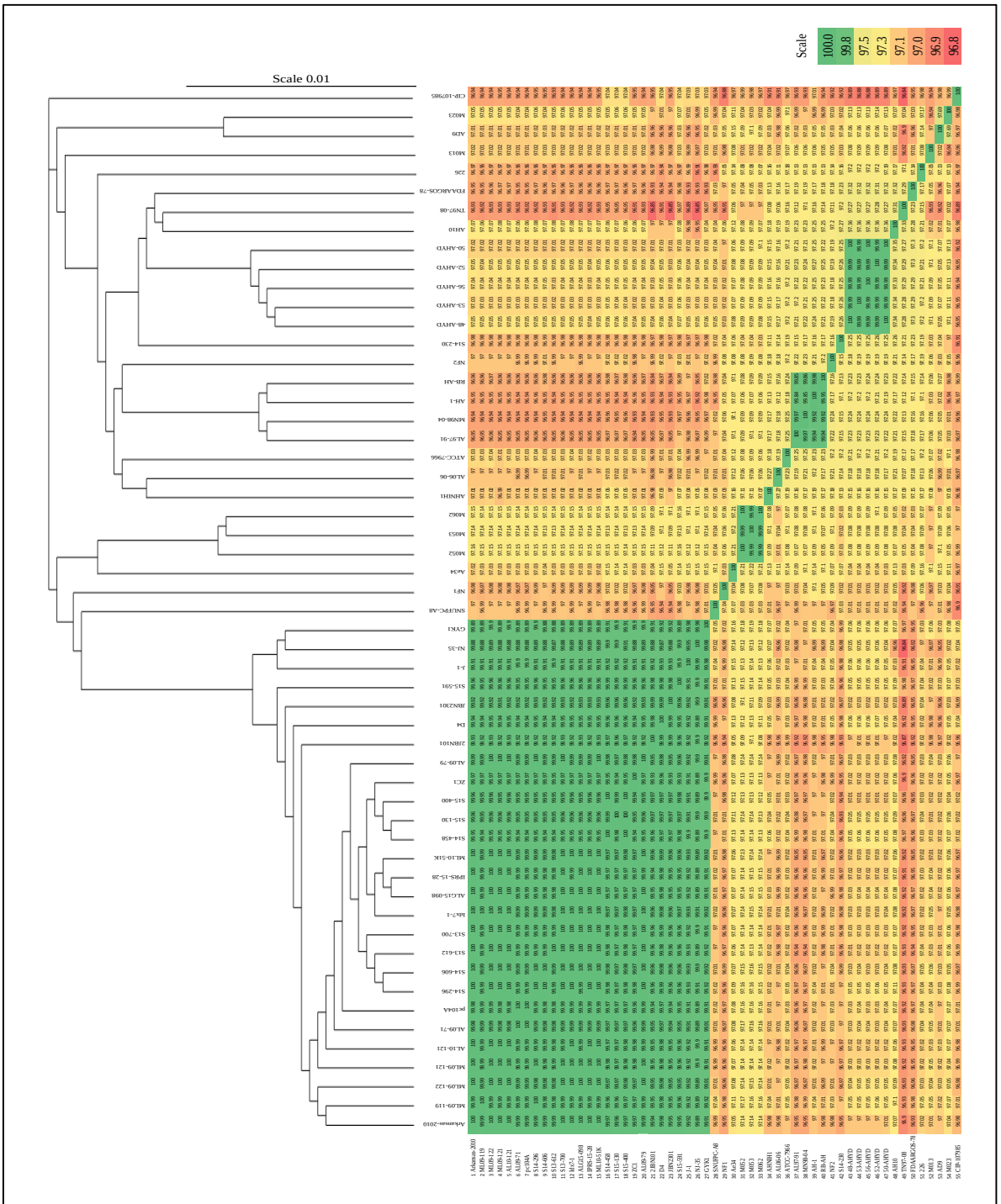


Figure 4.1 Average nucleotide identities (ANI) of *A. hydrophila* genomes and phylogenetic tree based on core genome.

Note that branch lengths of the phylogenetic tree were reduced to fit the image.

4.4.3 Secretion systems in *Aeromonas hydrophila* genomes

In our *in silico* secretion systems analysis, we identified that most of the U.S. and Chinese vAh isolates encode slightly more T1SS core components, for instance ATP-binding cassette (*abc*) and *mfp* genes, compared to environmental isolates. Additionally, the genome of human strain FDAARGOS_78 encodes more *abc* and outer membrane factor (*omf*) genes than the other 54 genomes (Figure 4.2).

All of the evaluated *A. hydrophila* genomes encode a T2SS system except one: strain S15-591. However, this strain does not have a completed genome sequence; it is a draft assembly with a large number of contigs. Therefore, it is possible that the genes may not have been detected due to the large number of gaps in the genome (Figure 4.2).

All the vAh strains in the current study encode mandatory and accessory genes of type 4 pilus (T4P). By contrast, strains from different origins and locations lacked the *pilQ* gene. One non-vAh strain (Ranae CIP 107985, which was isolated from a frog) encodes all the T4P elements (Figure 4.2).

Intriguingly, only one gene (*tadZ*) from the Tad system is present in all the evaluated *A. hydrophila* genomes. On the other hand, all the vAh strains as well as two non-vAh strains (ATCC 7966 and AHNIH1) encode this system (Figure 4.2).

The majority of vAh strains from the U.S. (except strain S14-230) and Chinese isolates do not carry all of the mandatory T3SS genes in their genomes (Figure 4.2). The mandatory T3SS gene cluster is composed of *sctU*, *sctJ*, *sctN*, *sctS*, *sctR*, *sctQ*, *sctV*, *sctU*, and *sctT*. Interestingly, only the *sctN* gene is present in all of the evaluated *A. hydrophila* genomes. Only two of the eight human isolates encode T3SS except for the *sctC* gene. Similarly, *A. hydrophila* ATCC 7966 does not encode T3SS, but it has two copies of the

sctN gene. By contrast, most of the environmental isolates carry T3SS. Interestingly, *A. hydrophila* strains that encode more T1SS components tend to have fewer or no genes encoding T3SS (Figure 4.2).

Some T3SS genes and are similar to flagella genes. Therefore, we used MacSyfinder to discriminate between T3SS and flagella genes. Of the 55 *A. hydrophila* genomes we evaluated, all carry the mandatory flagella genes (Figure 4.2).

All of the evaluated *A. hydrophila* genomes encode either the entire operon or remnants of the T6SSi. Most of the U.S. vAh isolates have only three T6SSi genes: *tssD* (*hcpA*), *tssH*, and *tssI* (*vgrG*). By contrast, almost all the China isolates encode the entire T6SSi. The exception was strain GYK1 from China, which has the same three T6SSi genes as the U.S. vAh isolates. Additionally, fish isolate Ae34 from Sri Lanka, four non-vAh isolates from the U.S. (AL06-06, MN98-04, AL97-91, and Ranae CIP 107985), and one fish isolate (AH-1) from Canada do not encode the entire T6SSi elements but have the same three genes as the U.S. vAh strains. Three of the Malaysian isolates (M023, RB-AH, and 226) encode the same three elements from T6SSi, whereas four Malaysian isolates (M013, M052, M053, and M054) encode the entire T6SSi.

Only one gene encoding a T9SS-like protein (s*prA*) was identified in the *A. hydrophila* genomes. This gene is encoded by all the evaluated *A. hydrophila* genomes.

Figure 4.2 Core and accessory proteins of secretion systems, T4P, Tad, and flagella encoded in *A. hydrophila* genomes.

Numbers and color represent the number of copies of each listed gene.

4.4.4 Virulence database comparison

Using E-value cutoff of $1*10^{-10}$, core components of 18 of 19 T1SS proteins, 12 of 14 T2SS proteins, 12 of 18 T4P proteins, 6 of 10 Tad proteins, and 16 of 21 T6SS proteins had significant protein matches in MVirDB. All of the core components of T3SS (11 proteins) and flagella (24 proteins) also had significant matches in MVirDB. T6SS effector VgrG (*tssI*) had a significant match in MVirDB using the established cutoff; however, two other T6SS proteins (Hcp (*tssD*)) had matches in MVirDB but with E-values below the established cutoff (Table 4.4).

Table 4.4 Identified secretion systems elements and related components searched against MvirDB

S. S. S. E.: Secretion system and structural elements, N/A: Not available. T1SS, T2SS, T4P and Tad system proteins were chosen from *A. hydrophila* ML09-119 strain; T3SS and flagellum proteins were chosen from *A. hydrophila* MN98-04, and T6SS proteins were chosen from *A. hydrophila* TN 97-08 strain.

S.S.S.E.	Accession numbers	Gene	Annotation description in NCBI	MVIRDB database		
				Lowest E-value	Description of E-value in MVIRDB	Matching organisms in MVIRDB
T1SS	WP_016350042.1	T1SS_mfp	HlyD family type I secretion periplasmic adaptor subunit	0	RTX toxin transporter	N/A
	WP_016351252.1	T1SS_abc	Lipid A export permease/ATP-binding protein MsbA	0	Lipid transporter ATP-binding/permease protein	Salmonella enterica subsp. enterica serovar Typhimurium str. LT2
	WP_016351504.1	T1SS_abc	Thiol reductant ABC exporter subunit CydD	0	Cysteine/glutathione ABC transporter membrane/ATP-binding component	Salmonella enterica subsp. enterica serovar Typhimurium str. LT2
	WP_048814156.1	T1SS_abc	Peptidase C39	0	Type I secretion channel, This type I complex secretes the rtx repeat containing RtxA	N/A
	WP_016351165.1	T1SS_mfp	HlyD family type I secretion periplasmic adaptor subunit	2.49E-167	Secretion protein, HlyD family	N/A
	WP_043159115.1	T1SS_abc	Peptidase C39	5.04E-161	Hemolysin B	Escherichia coli CFT073
	WP_016351505.1	T1SS_abc	Thiol reductant ABC exporter subunit CydC	1.26E-144	Cysteine/glutathione ABC transporter membrane/ATP-binding component	Salmonella enterica subsp. enterica serovar Typhimurium str. LT2
	WP_026080288.1	T1SS_omf	Outer membrane channel protein TolC	7.36E-136	Outer membrane channel protein	Enterobacter sakazakii ATCC BAA-894
	WP_016351470.1	T1SS_mfp	Macrolide transporter subunit MacA	4.62E-102	Putative drug-efflux protein	N/A
	WP_016351166.1	T1SS_abc	Type I secretion system permease/ATPase	6.53E-88	Cydolysin secretion ATP-binding protein	Bordetella pertussis Tohama I
	WP_016352402.1	T1SS_abc	Multidrug ABC transporter permease/ATP-binding protein	2.06E-80	Pyoverdine biosynthesis protein PvdE	Pseudomonas aeruginosa PAO1
	WP_016350580.1	T1SS_abc	Drug efflux ABC transporter ATP-binding and permease subunits	1.35E-73	Multidrug resistance ABC transporter ATP-binding and permease protein	N/A
	WP_043119643.1	T1SS_abc	Multidrug ABC transporter ATP-binding protein	6.71E-60	Lipid transporter ATP-binding/permease protein	Salmonella enterica subsp. enterica serovar Typhimurium str. LT2
	WP_016350338.1	T1SS_omf	RND transporter NodT	6.45E-54	Outer membrane protein OprN precursor	Pseudomonas aeruginosa PA7
	WP_016351842.1	T1SS_abc	ABC transporter ATP-binding protein	1.19E-41	ATP-binding cassette protein	Cryptosporidium parvum
	WP_043119233.1	T1SS_omf	Transporter	8.93E-36	Belongs to the outer membrane factor OMF TC 1.B.17 family.	N/A
	WP_016351164.1	T1SS_omf	Agglutination protein	5.79E-31	No definition line	N/A
	WP_016350337.1	T1SS_mfp	Secretion protein HlyD	3.36E-30	Putative membrane protein	Shigella flexneri 5 str. 8401
	WP_016351868.1	T1SS_omf	Outer membrane efflux protein	9.38E-05	No definition line	N/A
T2SS	WP_005308483.1	T2SS_gspG	Type II secretion system protein GspG	1.25E-101	General secretion pathway protein G	N/A
	WP_016349156.1	T2SS_gspD	Pilus (MSHA type) biogenesis protein MshL	3.16E-20	Putative type II secretion protein	Shigella dysenteriae Sd197
	WP_016349160.1	T2SS_gspF	GspF family protein	5.80E-59	Type IV pilus assembly protein PilC	Legionella pneumophila subsp. pneumophila str. Philadelphia 1
	WP_016349165.1	T2SS_gspJ	Msha pilin protein mshd	0.77	Respiratory nitrate reductase subunit alpha NarG	Mycobacterium tuberculosis H37Rv
	WP_016349335.1	T2SS_gspD	Type II secretion system protein GspD	0	General secretion pathway protein D	N/A
	WP_016349336.1	T2SS_gspE	Type II secretion system protein GspE	0	Putative type II secretion protein	Shigella dysenteriae Sd197
	WP_016349337.1	T2SS_gspF	Type II secretion system protein GspF	0	General secretion pathway protein F	N/A
	WP_016349339.1	T2SS_gspH	Type II secretion system protein GspH	1.90E-22	General secretion pathway protein H	N/A
	WP_016349340.1	T2SS_gspI	Type II secretion system protein GspI	3.40E-17	General secretion pathway protein I	Pseudomonas aeruginosa PAO1
	WP_016349341.1	T2SS_gspJ	Type II secretion system protein GspJ	7.93E-31	Putative type II secretion protein	Shigella dysenteriae Sd197
	WP_016349344.1	T2SS_gspM	General secretion pathway protein GspM	3.81E-15	Putative secretion pathway protein	Shigella dysenteriae Sd197
	WP_043118378.1	T2SS_gspL	Type II secretion system protein GspL	1.12E-16	GspL-like protein	Shigella dysenteriae Sd197
	WP_043118380.1	T2SS_gspN	Type II secretion protein N	N/A	N/A	N/A
	WP_043159151.1	T2SS_gspC	Type II secretion system protein GspC	5.73E-172	General secretion pathway protein C	N/A
	WP_043159236.1	T2SS_gspK	Type II secretion protein K	4.13E-38	Putative type II secretion protein	Shigella dysenteriae Sd197
T4P	WP_005305690.1	T4P_pilM	Cell division protein FtsA	0.15	Rod shape-determining protein MreB	N/A
	WP_011706550.1	T4P_pilT_pilU	Twitching motility protein PilT	1.81E-139	Twitching motility protein PilU	Pseudomonas aeruginosa PAO1
	WP_011707389.1	T4P_pilT_pilU	Twitching motility protein PilT	0	Twitching motility protein PilT	Pseudomonas aeruginosa PAO1
	WP_016349163.1	T4P_pilAE	MSHA pilin protein MshA	N/A	N/A	N/A
	WP_016349433.1	T4P_pilAE	Type IV pilin	4.22E-05	Pilin PilE	Neisseria meningitidis MC58
	WP_016349434.1	T4P_pilAE	Type IV pilin	1.23E-03	Type 4 fimbrial biogenesis protein PilE	Pseudomonas aeruginosa PAO1
	WP_016349435.1	T4P_pilAE	Type IV pilin	2.18	Trans-sialidase	Trypanosoma cruzi strain CL Brener
	WP_016349438.1	T4P_pilN_pilV	Type IV pilin	2.13	Hypothetical protein SMc03288	Sinorhizobium meliloti 1021
	WP_016351630.1	T4P_pilP	Type IV pilus biogenesis protein PilP	5.76E-22	Type 4 fimbrial biogenesis protein PilP	Pseudomonas aeruginosa PAO1
	WP_016351631.1	T4P_pilO	Type IV pilus biogenesis protein PilO	1.64E-36	Type 4 fimbrial biogenesis protein PilO	Pseudomonas aeruginosa PAO1
	WP_016351632.1	T4P_pilN	Pilus assembly protein PilN	3.61E-28	Type 4 fimbrial biogenesis protein PilN	Pseudomonas aeruginosa PAO1
	WP_016351633.1	T4P_pilM	Pilus assembly protein PilM	4.20E-55	Type 4 fimbrial biogenesis protein PilM	Pseudomonas aeruginosa PAO1
	WP_016351972.1	T4P_pilT_pilU	Twitching motility protein PilU	1.24E-156	Twitching motility protein PilU	Pseudomonas aeruginosa PAO1
	WP_016352168.1	T4P_pilB	Type IV-A pilus assembly ATPase PilB	0	Type 4 fimbrial biogenesis protein PilB	Pseudomonas aeruginosa PAO1
	WP_016352169.1	T4P_pilC	Type 4 fimbrial assembly protein PilC	6.74E-146	Still frameshift type 4 fimbrial biogenesis protein PilC	Pseudomonas aeruginosa PAO1
	WP_016352170.1	T4P_pilD	Prepilin peptidase	6.31E-125	Type 4 prepilin-like proteins leader peptide-processing enzyme	N/A
	WP_043118275.1	T4P_pilB	MSHA biogenesis protein MshE	1.93E-99	Type II protein secretion ATPase LspE	Legionella pneumophila subsp. pneumophila str. Philadelphia 1
	WP_043119325.1	T4P_pilQ	Pilus assembly protein PilQ	2.05E-132	Type 4 fimbrial biogenesis outer membrane protein PilQ precursor	Pseudomonas aeruginosa PAO1
Tad	WP_010634619.1	Tad_tadZ	Chromosome partitioning protein ParA	2.28E-102	Flagellar synthesis regulator FleN	Pseudomonas aeruginosa PAO1
	WP_011705350.1	Tad_flp	Flp family type IVb pilin	1.15	Regulatory protein	Streptomyces coelicolor A3 2
	WP_011705992.1	Tad_tadZ	Septum site-determining protein MinD	1.26E-93	Possible pump for toxic or radical ions	N/A
	WP_011707916.1	Tad_tadZ	Cobyric acid synthase CobQ	4.83E-09	Possible pump for toxic or radical ions	N/A
	WP_016350151.1	Tad_tadZ	Flp pilus assembly protein FlpE	1.69E-08	Pilus assembly, ATPase	N/A
	WP_016350152.1	Tad_tadA	Flp pilus assembly protein FlpF	1.09E-131	Pilus assembly, secretory protein kinase	N/A
	WP_016350153.1	Tad_tadB	Flp pilus assembly protein TadB	5.71E-22	Pilus assembly	N/A
	WP_043118796.1	Tad_tadV	Peptidase	3.1	Prepilin peptidase	N/A
	WP_043118798.1	Tad_rcpA	Secretin	4.49E-63	Pilus assembly, secretin	N/A
	WP_043118801.1	Tad_tadC	Pilus assembly protein TadC	2.04E-24	Pilus assembly, adherence	N/A

T3SS	WP_005307308.1	T3SS_sctN	Transcription termination factor Rho	3.34E-18	Type III secretion system ATPase	Salmonella typhimurium LT2
	WP_011707910.1	T3SS_sctN	ATP synthase subunit alpha	9.81E-41	Type III secretion	N/A
	WP_017408012.1	T3SS_sctS	EscS/YscS/HrcS family type III secretion system export apparatus protein	1.43E-43	Needle complex export protein	Yersinia pestis CO92
	WP_017408017.1	T3SS_sctN	EscN/YscN/HrcN family type III secretion system ATPase	0	Type III secretion system ATPase	Yersinia pestis CO92
	WP_029300154.1	T3SS_sctU	EscU/YscU/HrcU family type III secretion system export apparatus switch protein	4.73E-177	Needle complex export protein	Yersinia pestis CO92
	WP_043123793.1	T3SS_sctJ	EscJ/YscJ/HrcJ family type III secretion inner membrane ring protein	3.88E-120	Type III export protein PscJ	Pseudomonas aeruginosa PAO1
	WP_043123801.1	T3SS_sctC	EscC/YscC/HrcC family type III secretion system outer membrane ring protein	0	AscC	N/A
	WP_043123815.1	T3SS_sctV	EscV/YscV/HrcV family type III secretion system export apparatus protein	0	Low calcium response protein D	Yersinia pestis CO92
	WP_043123827.1	T3SS_sctQ	Type III secretion system protein	1.62E-79	YscQ	N/A
	WP_043123830.1	T3SS_sctR	EscR/YscR/HrcR family type III secretion system export apparatus protein	1.70E-119	Type III secretion system protein	Pseudomonas aeruginosa PAO1
	WP_043123832.1	T3SS_sctT	EscT/YscT/HrcT family type III secretion system export apparatus protein	5.42E-78	Needle complex export protein	Yersinia pestis CO92
Flagellum	WP_005305833.1	Flg_fliE	Flagellar hook-basal body complex protein FliE	7.41E-17	Flagellar hook-basal body complex protein FliE	N/A
	WP_005307215.1	Flg_sctN_FLG	ATP synthase subunit beta	1.87E-38	Type III secretion system ATPase	Yersinia pestis CO92
	WP_011705271.1	Flg_fliE	Flagellar hook-basal body complex protein FliE	1.06E-24	Flagellar hook-basal body protein FliE	Legionella pneumophila subsp. pneumophila str. Philadelphia 1
	WP_011705275.1	Flg_sctN_FLG	Flagellum-specific ATP synthase	9.78E-178	Flagellum-specific ATP synthase	Pseudomonas aeruginosa PAO1
	WP_011705280.1	Flg_sctQ_FLG	Flagellar motor switch protein FliN	1.23E-47	Flagellar motor switch protein FliN	N/A
	WP_011705282.1	Flg_sctS_FLG	Flagellar biosynthetic protein FliQ	2.56E-26	Flagellar biosynthesis protein FliQ [Pseudomonas aeruginosa PAO1]	N/A
	WP_011705283.1	Flg_sctT_FLG	Flagellar biosynthesis protein FliR	1.80E-63	Flagellar biosynthesis protein FliR	Pseudomonas aeruginosa PAO1
	WP_011706644.1	Flg_flgB	Flagellar basal body rod protein FlgB	1.72E-32	Flagellar basal body rod protein FlgB	Legionella pneumophila subsp. pneumophila str. Philadelphia 1
	WP_016350050.1	Flg_sctJ_FLG	Flagellar M-ring protein FliF	3.19E-80	Flagellar M-ring protein	N/A
	WP_016350056.1	Flg_sctV_FLG	Flagellar biosynthesis protein FlhA	0	Flagellar biosynthesis protein FlhA	Pseudomonas aeruginosa PAO1
	WP_017409386.1	Flg_sctQ_FLG	Flagellar motor switch protein FliN	1.28E-21	Flagellar motor switch protein FliN	N/A
	WP_017409388.1	Flg_sctS_FLG	Flagellar export apparatus protein FliQ	1.72E-15	Flagellar biosynthesis protein FliQ	Pseudomonas aeruginosa PAO1
	WP_017409389.1	Flg_sctT_FLG	Flagellar biosynthetic protein FliR	3.78E-47	Flagellar biosynthesis protein FliR	Pseudomonas aeruginosa PAO1
	WP_017409402.1	Flg_flgC	Flagellar basal body rod protein FlgC	1.31E-41	Flagellar basal body rod protein FlgC	Pseudomonas aeruginosa PAO1
	WP_017409406.1	Flg_flgC	Flagellar basal body rod protein FlgG	2.03E-96	Flagellar basal body rod protein FlgG	Pseudomonas aeruginosa PAO1
	WP_043121750.1	Flg_flgC	Flagellar basal body rod protein FlgC	3.26E-54	Flagellar basal body rod protein FlgC	Pseudomonas aeruginosa PAO1
	WP_043124302.1	Flg_sctU_FLG	Flagellar biosynthesis protein FlhB	2.42E-118	Polar flagellar assembly protein FlhB	N/A
	WP_043124305.1	Flg_sctR_FLG	Flagellar biosynthetic protein FliP	4.87E-90	Flagellar biosynthesis protein FliP	Pseudomonas aeruginosa PAO1
	WP_043126821.1	Flg_sctR_FLG	Flagellar biosynthetic protein FliP	1.56E-72	Flagellar biosynthesis protein FliP	Pseudomonas aeruginosa PAO1
	WP_043126824.1	Flg_sctU_FLG	Flagellar biosynthesis protein FlhB	2.81E-88	Flagellar biosynthesis protein FlhB	Pseudomonas aeruginosa PAO1
	WP_043126828.1	Flg_sctV_FLG	Flagellar biosynthesis protein FlhA	0	Flagellar biosynthesis protein FlhA	Pseudomonas aeruginosa PAO1
	WP_043126836.1	Flg_sctJ_FLG	Flagellar M-ring protein FliF	1.02E-79	Flagellar M-ring protein	N/A
	WP_043126844.1	Flg_sctN_FLG	Flagellar protein export ATPase FliI	1.23E-171	Flagellum-specific ATP synthase	Pseudomonas aeruginosa PAO1
	WP_043126856.1	Flg_flgB	Flagellar basal-body rod protein FlgB	2.69E-27	Flagellar basal body rod protein FlgB	Legionella pneumophila subsp. pneumophila str. Philadelphia 1
T6SS	WP_010634228.1	T6SSi_tssL	Type IV / VI secretion system protein, DotU family	2.42E-73	Hypothetical protein VCA0115	Vibrio cholerae O1 biovar eltor str. N16961
	WP_010634822.1	T6SSi_tssD	Hcp1 family type VI secretion system effector	3.79	RNA-directed RNA polymerase L	N/A
	WP_011705044.1	T6SSi_tssD	Hcp1 family type VI secretion system effector	3.79	RNA-directed RNA polymerase L	N/A
	WP_011705712.1	T6SSi_tssB	Type VI secretion protein	1.97E-57	Hypothetical protein VCA0107	Vibrio cholerae O1 biovar eltor str. N16961
	WP_011705715.1	T6SSi_tssF	Type VI secretion protein	0	Hypothetical protein VCA0110	Vibrio cholerae O1 biovar eltor str. N16961
	WP_011705718.1	T6SSi_tssJ	Type VI secretion system-associated lipoprotein	2.05E-37	Hypothetical protein VCA0113	Vibrio cholerae O1 biovar eltor str. N16961
	WP_017411057.1	T6SSI_evpI	Membrane protein	6.67	Adhesion and penetration protein autotransporter	N/A
	WP_029945187.1	T6SSi_tssH	ATP-dependent chaperone ClpB	0	ClpB protein	N/A
	WP_042019066.1	T6SSi_tssD	Hcp1 family type VI secretion system effector	3.72	Ferrichrome transport ATP-binding protein	Staphylococcus aureus subsp. aureus MW2
	WP_043163044.1	T6SSi_tssC	Type VI secretion protein	0	Hypothetical protein VCA0108	Vibrio cholerae O1 biovar eltor str. N16961
	WP_060388738.1	T6SSI_evpI	Hypothetical protein	6.87	Adhesion and penetration protein autotransporter	N/A
	WP_060389076.1	T6SSi_tssI	Type IV secretion protein Rhs	0	VgrG protein	Vibrio cholerae O1 biovar eltor str. N16961
	WP_060389654.1	T6SSi_tssI	Type IV secretion protein Rhs	0	VgrG protein	Vibrio cholerae O1 biovar eltor str. N16961
	WP_060389655.1	T6SSi_tssA	Hypothetical protein	3.34E-10	Hypothetical protein VCA0121	Vibrio cholerae O1 biovar eltor str. N16961
	WP_060389656.1	T6SSi_tssM	Type VI secretion protein IcmF	0	IcmF-related protein	Vibrio cholerae O1 biovar eltor str. N16961
	WP_060389657.1	T6SSi_tssA	Type VI secretion-associated protein	1.62E-47	Hypothetical protein VCA0119	Vibrio cholerae O1 biovar eltor str. N16961
	WP_060389659.1	T6SSi_tssH	ClpV1 family T6SS ATPase	0	ClpB protein	Vibrio cholerae O1 biovar eltor str. N16961
	WP_060389660.1	T6SSi_tssK	Type VI secretion protein	0	Hypothetical protein VCA0114	Vibrio cholerae O1 biovar eltor str. N16961
	WP_060389662.1	T6SSi_tssG	Type VI secretion protein	1.66E-77	Hypothetical protein VCA0111	Vibrio cholerae O1 biovar eltor str. N16961
	WP_060389663.1	T6SSi_tssE	Lysozyme	4.62E-26	Hypothetical protein VCA0109	Vibrio cholerae O1 biovar eltor str. N16961
	WP_060389666.1	T6SSi_tssI	Type IV secretion protein Rhs	0	VgrG protein	Vibrio cholerae O1 biovar eltor str. N16961
Mutants	WP_010634822.1	T6SSi_tssD-Hcp	Hcp1 family type VI secretion system effector	3.79	RNA-directed RNA polymerase L	N/A
	WP_011705044.1	T6SSi_tssD-Hcp	Hcp1 family type VI secretion system effector	3.79	RNA-directed RNA polymerase L	N/A
	WP_016350448.1	T6SSi_tssI-Vgr-Rhs	Rhs element Vgr family protein	0	VgrG protein	Vibrio cholerae O1 biovar eltor str. N16961

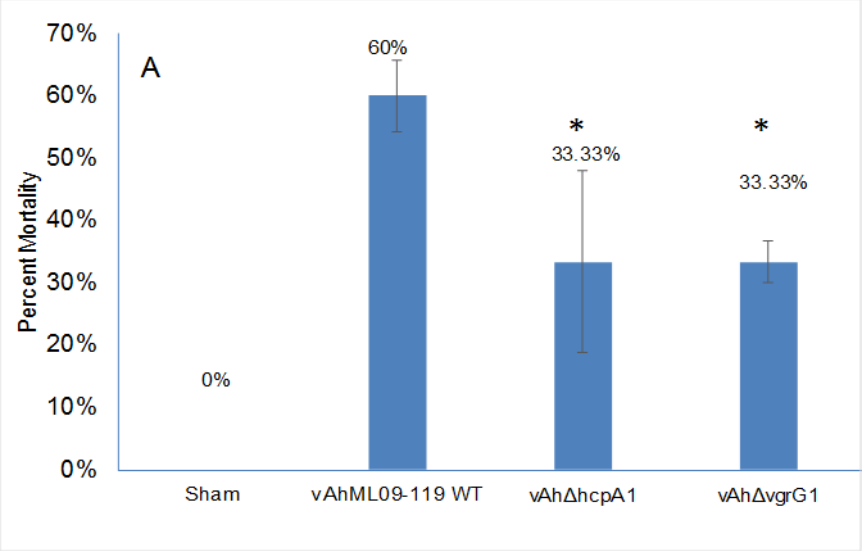
4.4.5 Construction and virulence of T6SS effector mutant strains

In the present study, we successfully introduced in-frame deletions in two genes encoding T6SS effectors: *hcpA1* and *vgrG1* (Table 4.5). The Δ*hcpA1* mutation has a deletion of 537 bp out of 564 bp (95.21%) and Δ*vgrG* has a deletion of 2,739 bp out of 1,992 bp (94.43%).

Results of the immersion challenge in catfish fingerlings indicated that the mortality rate was significantly lower ($p < 0.05$) in *vAh*Δ*hcpA1* and *vAh*Δ*vgrG1* compared with vAh WT strain (33.33% and 33.33% mortality vs. 60% mortality) (Figure 4.3A). Fingerlings immunized with *vAh*Δ*hcpA1* and *vAh*Δ*vgrG1* had 91.67% and 100.00% percent survival, respectively, compared to 60.00% survival in the sham-vaccinated control group (Figure 4.3B).

Table 4.5 The sizes of upstream (USF), downstream (DSF), and in-frame fused fragments (FF), deleted region (DR), and undeleted region (UD) by base pair (bp).

Gene name	**Gene symbols**	**Locus tag**	**USF (AB)**	**DSF (CD)**	**FF (AD)**	**DR**	**UD**
HcpA-like protein	*hcpAI*	AHML_05970	1074	957	2031	537	27
Rhs element Vgr protein	*vgrG1*	AHML_05975	990	1038	2028	2739	42



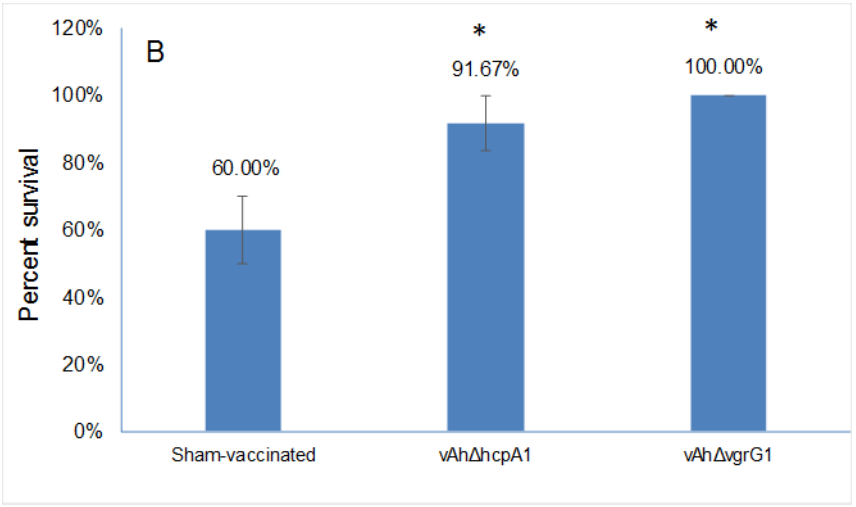


Figure 4.3 Virulence of vAh T6SS mutants in channel catfish fingerlings.

(A) Percent mortalities in catfish fingerlings experimentally infected with vAh mutants and vAh wild type (WT) strain ML09-119. (B) Percent survival in vaccinated catfish fingerlings following experimental infection with vAh WT. Data are the mean ± SE of three replicate tanks. Significant differences between challenged and non-vaccinated treatments are indicated with asterisks ($p < 0.05$).

4.4.6 Host-pathogen interaction

Using HPIDB, we predicted the interaction of identified *A. hydrophila* secretion system components with host channel catfish (*Ictalurus punctatus*) proteins. We identified that 333 catfish proteins interact with 30 different components of the *A. hydrophila* secretion systems.

4.5 Discussion

In this study, our goal was to compare secretion systems distribution in *A. hydrophila* genomes using comparative genomics. We found that some of the secretion systems commonly involved in pathogenesis of Gram-negative bacterial infections are not consistently present in U.S. vAh isolates. However, there are three secretion systems present in all *A. hydrophila* strains we analyzed, and one system that is present almost specifically in vAh strains. To clarify the role of three T6SS genes found in all *A. hydrophila*, we constructed deletion mutants of *hcpA1 (tssD)* and *vgrG (tssI)* genes in vAh strain ML09-119. The two mutants had significantly reduced virulence in catfish fingerlings compared to parent strain ML09-119. Therefore, these genes still contribute to vAh virulence despite the absence of a complete T6SS. Overall, these comparative genomics and mutational analyses clarify the distribution of secretion systems in *A. hydrophila* and partially characterize their roles in virulence.

Our phylogenetic tree based on the complete core genome of 55 selected *A. hydrophila* strains shows the 27 vAh strains formed a very highly conserved branch that is clearly separated from the rest of the *A. hydrophila* strains. Also, ANI analysis of the 55 *A. hydrophila* genomes showed that U.S. vAh isolates and Chinese epidemic isolates were derived from the same monophyletic clade. None of the genomes in our analysis

had below 96% ANI with all of the other genomes in our study. ANI and AAI values within the cluster of 27 vAh strains were above 99.88% (ANI) and 99.89% (AAI) (Figure 4.1).

A. hydrophila secretes a wide range of extracellular enzymes and toxins. Type I secretion systems are capable of secreting exotoxins and enzymes by a one-step process from the cytoplasm to outer membrane. T1SS consists of three main components: ATP-binding cassette (ABC) transporters, membrane fusion protein (MFP), and outer membrane factor (OMF) (Green and Mecsas, 2016). All of the evaluated *A. hydrophila* genomes carry core components of the T1SS. However, most of the vAh isolates and some of the clinical isolates encode more core components of the T1SS. In *Vibrio*, RTX toxin is secreted by T1SS (Boardman and Satchell, 2004). Moreover, presence of T1SS increases virulence of *Vibrio cholera* (Dolores et al., 2015) and *Serratia marcescens* (Letoffe et al., 1996).

Not surprisingly, all the evaluated *A. hydrophila* genomes possess the T2SS. This system is capable of secreting enzymes such as proteases, phosphatases, and lipases (Korotkov et al., 2012; Green and Mecsas, 2016); in *A. hydrophila*, it is also well known for exporting cytotoxic enterotoxin (Act), which has hemolytic and cytotoxic activities (Chopra et al., 2000; Galindo et al., 2004; Korotkov et al., 2012). T2SS is a large, trans-envelope apparatus encoded by a set of 12-16 core genes. It is located in the outer membrane, and it transports folded proteins from periplasm into the extracellular environment. T2SS differs from T1SS, which releases proteins to the outer medium, and T3SS, T4SS, and T6SS, which are contact-dependent (Hayes et al., 2010). T2SS secretes specific toxins, effectors, and large proteins that could not be secreted to the host or

competitor bacteria otherwise (Rondelet and Condemine, 2013; Rosenzweig and Chopra, 2013). T2SS is evolutionarily related to the type 4 pilus (T4P) system, which is responsible for motility, signaling, and adhesion (Nivaskumar et al., 2014). T4P has not been studied extensively in *A. hydrophila*. T4P and T2SS show a high degree of similarity in their components, and one of the genes encoding a T2SS component, *gspO*, is located in the T4P-encoding locus (Nivaskumar et al., 2014). Our secretion analysis assigned the *A. hydrophila gspO* gene as *pilD*, which is one of the accessory genes of T4P. Although *A. hydrophila gspC* gene is listed as missing a mandatory gene in Figure 4.2, it is present as an accessory gene (Figure 4.2). In *E. coli*, T2SS and T4P are important for persistent infection (Kulkarni et al., 2009).

The tight adherence (Tad) system contributes to biofilm formation, colonization, and virulence of several pathogens (Tomich et al., 2007). The Tad system is similar to T2SS systems (Peabody et al., 2003; Tomich et al., 2007). Intriguingly, our results showed that a complete Tad secretion system is available in vAh strains, whereas the majority of the other evaluated *A. hydrophila* genomes do not encode a Tad system except for two strains, one of which is human clinical isolate (strain AHNIH1), and the other is from milk (strain ATCC 7966). Interestingly, only one gene (*tadZ*) from this system is available in all the evaluated *A. hydrophila* genomes. TadZ is encoded by one of the mandatory genes of the Tad system and plays a major role in mediating polar localization of the Tad secretion system (Perez-Cheeks et al., 2012).

Many Gram-negative pathogens use type III secretion system, which delivers effector proteins directly into host cells. Many components of this system are homologous to flagellum proteins. T3SS is an important contributor to pathogenesis of

some *A. hydrophila* strains (Vilches et al., 2004; Yu et al., 2004); however, our comparative genomics analysis showed that 27 vAh strains lack genes encoding T3SS (except for the *sctN* gene, which encodes a highly conserved ATPase that contributes to energy metabolism and provides recognition capability for T3SS effectors and other virulence factors (Zarivach et al., 2007)). Most of the non-vAh isolates in our study encode T3SS, but the majority of these are environmental isolates from outside the U.S. On the other hand, U.S. environmental isolates (soil and wetland sediment) do not encode a T3SS, and they also lack Tad systems. Our results showing the absence of T3SS in vAh strains are consistent with a previous relatively small-scale comparative genomics study (Pang et al., 2015). Therefore, the presence of genes encoding a T3SS is not a good indicator of virulence potential for *A. hydrophila* strains.

Interestingly, there is an inverse relationship between the presence of a Tad system and a T3SS in many of the *A. hydrophila* genomes we analyzed. The Tad system is encoded in the vAh isolates, but they do not encode T3SS. On the other hand, almost all of the non-vAh *A. hydrophila* strains do not encode a Tad system, but many of these genomes encode T3SS (Figure 4.2). Ten strains have neither Tad nor T3SS sytems, and only one strain (human isolate AHNIH1) encodes both systems. Therefore, we can hypothesize thatthe three secretion systems consistently encoded in vAh strains are T2SS, T4P, and Tad.

Flagella are important in motility and often in attachment to the host. They are linked with biofilm formation, which contributes to persistent infection (Tomas, 2012). In eels, an *A. hydrophila* polar flagellum mutant had decreased survival and adherence to eel macrophages (Qin et al., 2014). Because flagella proteins are similar to T3SS proteins

(Nguyen et al., 2000; Gophna et al., 2003), we included them in our comparative genomics analysis. All the evaluated *A. hydrophila* genomes encode mandatory genes. In some bacteria, T3SS components play a role in flagellar rotation, but in *A. hydrophila*, there is only one T3SS gene (*sctN*) shared by all the evaluated *A. hydrophila* genomes. An *A. hydrophila* master regulator of T3SS (ExsA) negatively affects the lateral flagella (Zhao and Shaw, 2016), so it is possible that T3SS and flagella proteins interact in *A. hydrophila* strains encoding both systems.

T6SS is used widely distributed in Gram-negative bacteria, and it contributes to bacterial fitness in specific niches (Cianfanelli et al., 2016). In particular, it delivers secreted proteins into competitor bacteria or host cells (Zoued et al., 2014). T6SS is categorized into three phylogenetic subtypes (T6SSi, T6SSii, T6SSiii) (Russell et al., 2014b). All of the *A. hydrophila* genomes we evaluated encode the entire T6SSi operon or remnants of the T6SSi. Some of the U.S. vAh strains have only three mandatory genes (*tssD*, *tssH*, *tssI*) of T6SS, while other vAh strains from U.S. and China encode all the mandatory genes of T6SS. We extended our research to understand the role of these remnants in the pathogenicity of *A. hydrophila*.

The *tssD* and *tssI* genes in *A. hydrophila* ML09-119 are also known as *hcpA* and *vgrG*, respectively. Strain ML09-119 encodes two hemolysin co-regulated proteins (Hcp) (AHML_05970 and AHML_10025) and two valine-glycine repeat G (VgrG) proteins (AHML_05975 and AHML_10030). The *hcpA* genes are linked to the v*grG* genes in strain ML09-119; the *hcpA1* gene is adjacent to *vgrG1* gene, and *hcpA2* gene is adjacent to the *vgrG2* gene. Multiple copies of *hcp* and v*grG* genes are commonly seen in several bacterial species that possess a T6SS, including *Vibrio cholerae, Pseudomonas*

aeruginosa, *A. hydrophila* SSU, and *A. hydrophila* ATCC 7966T (Mougous et al., 2006; Podladchikova et al., 2011; Sha et al., 2013). Hcp and VgrG are effector proteins of T6SS (Cascales, 2008). However, structural analysis of Hcp and VgrG from *P. aeruginosa* and *V. cholerae* showed that these proteins independently formed a transportation channel between the inner and outer membranes through which other effector molecules can be transported to the host cell (Leiman et al., 2009; Pell et al., 2009). Thus, Hcp and VgrG could also be part of the secretion apparatus.

Deletion of the *hcpA1* and *vgrG* genes in vAh strain ML09-119 affected virulence significantly (about a 2.24-fold reduction in virulence in catfish fingerlings compared to the parent strain). Several studies emphasized the importance of Hcp and VgrG in pathogenesis. Examples include *Escherichia coli* (Dudley et al., 2006), *Pseudomonas aeruginosa* (Hood, 2010), *Edwardsiella tarda* (Rao et al., 2004), and *Aeromonas* (Sha et al., 2013). In *V. cholera*, an *hcp1*/*hcp2* mutant is avirulent, whereas individual *hcp1* or *hcp2* mutants retain virulence. Therefore, at least one Hcp protein is required and sufficient for virulence (Pukatzki et al., 2006). In an intraperitoneal murine model of infection, all Hcp and VgrG paralogues were required for optimal *A. hydrophila* SSU virulence and dissemination to mouse peripheral organs (Sha et al., 2013).

T9SS is typically only found in some species in the Bacteroidetes phylum, so it is not surprising that only one gene (s*prA*) encoding a protein similar to T9SS is present in all the evaluated *A. hydrophila* genomes. T9SS functions as a secretion system but also enables gliding motility (McBride and Zhu, 2013; Sato et al., 2013)(McBride and Nakane, 2015). In *Flavobacterium johnsoniae*, SprA is responsible (along with SprE and SprT) for secretion of SprB (Shrivastava et al., 2013).

Due to their role in secreting proteins involved in pathogenesis of multiple bacterial species, it is not surprising that almost all the *A. hydrophila* proteins identified from T1SS, T2SS, T6SS, Tad, and T4P have significant matches with known bacterial virulence proteins in MVirDB. T3SS and flagella are well-known virulence factors, and all of the *A. hydrophila* proteins from these systems have significant matches in MVirDB. Additionally, 30 of the *A. hydrophila* secretion system proteins have predicted interactions with channel catfish proteins. We chose channel catfish as the host species for this analysis because of its importance as an aquaculture species in the U.S. and due to the impacts and known virulence of vAh strains on this species. These results confirm the multiple interactions between *A. hydrophila* secretion systems and channel catfish, adding additional evidence to their potential roles in *A. hydrophila* virulence.

In summary, our results analysis indicate that vAh strains do not encode two of the contact-dependent secretion systems commonly involved in virulence of many Gram-negative pathogens, T3SS and T6SS. In fact, the T3SS is missing in all vAh strains and many other *A. hydrophila* strains. This suggests that vAh utilizes other systems to secrete effectors, toxins, and large secreted proteins. T1SS, T2SS, and T4P systems are encoded in all the *A. hydrophila* strains we sequenced, and these systems likely secrete several virulence-related proteins. Interestingly, the Tad system is present in all the vAh strains we sequenced, but it is only present in two of the non-vAh strains we analyzed. It is possible that the Tad system is one of the vAh-specific adaptations that make this clade of *A. hydrophila* more virulent.

Although only some *A. hydrophila* have a complete T6SS, all of the strains in our analysis encode three T6SS proteins. To determine their potential role in virulence, we

deleted these two genes and determined that the resulting Δ*hcpA1* and Δ*vgrG* mutant strains are significantly attenuated in channel catfish. Further investigation of the role of these T6SS genes in *A. hydrophila* is warranted, including the effects of deleting all the *hcpA* and *vrgG* alleles on *A. hydrophila* virulence.

4.6 References

Abby, S.S., Cury, J., Guglielmini, J., Neron, B., Touchon, M., and Rocha, E.P. (2016). Identification of protein secretion systems in bacterial genomes. *Sci Rep* 6**,** 23080. doi: 10.1038/srep23080.

Abby, S.S., Neron, B., Menager, H., Touchon, M., and Rocha, E.P. (2014). MacSyFinder: a program to mine genomes for molecular systems with an application to CRISPR-Cas systems. *PLoS One* 9(10)**,** e110726. doi: 10.1371/journal.pone.0110726.

Abdelhamed, H., Lu, J., Shaheen, A., Abbass, A., Lawrence, M.L., and Karsi, A. (2013). Construction and evaluation of an Edwardsiella ictaluri fhuC mutant. *Vet Microbiol* 162(2-4)**,** 858-865. doi: 10.1016/j.vetmic.2012.11.006.

Abdelhamed, H., Nho, S.W., Turaga, G., Banes, M.M., Karsi, A., and Lawrence, M.L. (2016). Protective efficacy of four recombinant fimbrial proteins of virulent Aeromonas hydrophila strain ML09-119 in channel catfish. *Vet Microbiol* 197**,** 8-14. doi: 10.1016/j.vetmic.2016.10.026.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. *J Mol Biol* 215(3)**,** 403-410. doi: 10.1016/S0022-2836(05)80360-2.

Ammari, M.G., Gresham, C.R., McCarthy, F.M., and Nanduri, B. (2016). HPIDB 2.0: a curated database for host-pathogen interactions. *Database (Oxford)* 2016. doi: 10.1093/database/baw103.

Beaz-Hidalgo, R., Hossain, M.J., Liles, M.R., and Figueras, M.J. (2015). Strategies to avoid wrongly labelled genomes using as example the detected wrong taxonomic affiliation for aeromonas genomes in the GenBank database. *PLoS One* 10(1)**,** e0115813. doi: 10.1371/journal.pone.0115813.

Boardman, B.K., and Satchell, K.J. (2004). Vibrio cholerae strains with mutations in an atypical type I secretion system accumulate RTX toxin intracellularly. *J Bacteriol* 186(23)**,** 8137-8143. doi: 10.1128/JB.186.23.8137-8143.2004.

Brettin, T., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Olsen, G.J., et al. (2015). RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* 5**,** 8365. doi: 10.1038/srep08365.

Cascales, E. (2008). The type VI secretion toolkit. *EMBO Rep* 9(8)**,** 735-741. doi: 10.1038/embor.2008.131.

Chan, K.G., Puthucheary, S.D., Chan, X.Y., Yin, W.F., Wong, C.S., Too, W.S., et al. (2011). Quorum sensing in Aeromonas species isolated from patients in Malaysia. *Curr Microbiol* 62(1)**,** 167-172. doi: 10.1007/s00284-010-9689-z.

Chan, K.G., Tan, W.S., Chang, C.Y., Yin, W.F., and Mumahad Yunos, N.Y. (2015). Genome Sequence Analysis Reveals Evidence of Quorum-Sensing Genes Present in Aeromonas hydrophila Strain M062, Isolated from Freshwater. *Genome Announc* 3(2). doi: 10.1128/genomeA.00100-15.

Chopra, A.K., Xu, X., Ribardo, D., Gonzalez, M., Kuhl, K., Peterson, J.W., et al. (2000). The cytotoxic enterotoxin of Aeromonas hydrophila induces proinflammatory cytokine production and activates arachidonic acid metabolism in macrophages. *Infect Immun* 68(5)**,** 2808-2818.

Cianfanelli, F.R., Monlezun, L., and Coulthurst, S.J. (2016). Aim, Load, Fire: The Type VI Secretion System, a Bacterial Nanoweapon. *Trends Microbiol* 24(1)**,** 51-62. doi: 10.1016/j.tim.2015.10.005.

Colston, S.M., Fullmer, M.S., Beka, L., Lamy, B., Gogarten, J.P., and Graf, J. (2014). Bioinformatic genome comparisons for taxonomic and phylogenetic assignments using Aeromonas as a test case. *MBio* 5(6)**,** e02136. doi: 10.1128/mBio.02136-14.

Dolores, J.S., Agarwal, S., Egerer, M., and Satchell, K.J. (2015). Vibrio cholerae MARTX toxin heterologous translocation of beta-lactamase and roles of individual effector domains on cytoskeleton dynamics. *Mol Microbiol* 95(4)**,** 590-604. doi: 10.1111/mmi.12879.

Dozois, C.M., Daigle, F., and Curtiss, R., 3rd (2003). Identification of pathogen-specific and conserved genes expressed in vivo by an avian pathogenic *Escherichia coli* strain. *Proc. Natl. Acad. Sci. U. S. A.* 100(1)**,** 247-252. doi: 10.1073/pnas.232686799.

Dudley, E.G., Thomson, N.R., Parkhill, J., Morin, N.P., and Nataro, J.P. (2006). Proteomic and microarray characterization of the AggR regulon identifies a pheU pathogenicity island in enteroaggregative Escherichia coli. *Mol Microbiol* 61(5)**,** 1267-1282. doi: 10.1111/j.1365-2958.2006.05281.x.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5)**,** 1792-1797. doi: 10.1093/nar/gkh340.

Emond-Rheault, J.G., Vincent, A.T., Trudel, M.V., Brochu, F., Boyle, B., Tanaka, K.H., et al. (2015). Variants of a genomic island in Aeromonas salmonicida subsp. salmonicida link isolates with their geographical origins. *Vet Microbiol* 175(1)**,** 68-76. doi: 10.1016/j.vetmic.2014.11.014.

Felsenstein, J. (1989). PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics* 5**,** 164-166. doi: citeulike-article-id:2344765.

Forn-Cuni, G., Tomas, J.M., and Merino, S. (2016). Whole-Genome Sequence of Aeromonas hydrophila Strain AH-1 (Serotype O11). *Genome Announc* 4(5). doi: 10.1128/genomeA.00920-16.

Galindo, C.L., Fadl, A.A., Sha, J., Gutierrez, C., Jr., Popov, V.L., Boldogh, I., et al. (2004). Aeromonas hydrophila cytotoxic enterotoxin activates mitogen-activated protein kinases and induces apoptosis in murine macrophages and human intestinal epithelial cells. *J Biol Chem* 279(36)**,** 37597-37612. doi: 10.1074/jbc.M404641200.

Gophna, U., Ron, E.Z., and Graur, D. (2003). Bacterial type III secretion systems are ancient and evolved by multiple horizontal-transfer events. *Gene* 312**,** 151-163.

Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P., and Tiedje, J.M. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 57(Pt 1)**,** 81-91. doi: 10.1099/ijs.0.64483-0.

Green, E.R., and Mecsas, J. (2016). Bacterial Secretion Systems: An Overview. *Microbiol Spectr* 4(1). doi: 10.1128/microbiolspec.VMBF-0012-2015.

Griffin, M.J., Goodwin, A.E., Merry, G.E., Liles, M.R., Williams, M.A., Ware, C., et al. (2013). Rapid quantitative detection of Aeromonas hydrophila strains associated with disease outbreaks in catfish aquaculture. *J Vet Diagn Invest* 25(4)**,** 473-481. doi: 10.1177/1040638713494210.

Grim, C.J., Kozlova, E.V., Ponnusamy, D., Fitts, E.C., Sha, J., Kirtley, M.L., et al. (2014). Functional genomic characterization of virulence factors from necrotizing fasciitis-causing strains of Aeromonas hydrophila. *Appl Environ Microbiol* 80(14)**,** 4162-4183. doi: 10.1128/AEM.00486-14.

Han, J.E., Kim, J.H., Choresca, C., Shin, S.P., Jun, J.W., and Park, S.C. (2013). Draft Genome Sequence of a Clinical Isolate, Aeromonas hydrophila SNUFPC-A8, from a Moribund Cherry Salmon (Oncorhynchus masou masou). *Genome Announc* 1(1). doi: 10.1128/genomeA.00133-12.

Hayes, C.S., Aoki, S.K., and Low, D.A. (2010). Bacterial contact-dependent delivery systems. *Annu Rev Genet* 44**,** 71-90. doi: 10.1146/annurev.genet.42.110807.091449.

Hemstreet, W.B. (2010). An update on Aeromonas hydrophila from a fish health specialist for summer 2010. *Catfish J.* 24(4).

Herrero, M., de Lorenzo, V., and Timmis, K.N. (1990). Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria. *J. Bacteriol.* 172(11)**,** 6557-6567.

Hood, R.D.S., P. Hsu, F. Guvener, T. Carl, M. A. Trinidad, R. R. Silverman, J. M. Ohlson, B. B. Hicks, K. G. Plemel, R. L. Li, M. Schwarz, S. Wang, W. Y. Merz, A. J. Goodlett, D. R. Mougous, J. D. (2010). A type VI secretion system of Pseudomonas aeruginosa targets a toxin to bacteria. *Cell Host Microbe* 7(1)**,** 25-37. doi: 10.1016/j.chom.2009.12.007.

Horton, R.M., Hunt, H.D., Ho, S.N., Pullen, J.K., and Pease, L.R. (1989). Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 77(1)**,** 61-68.

Hossain, M.J., Sun, D., McGarey, D.J., Wrenn, S., Alexander, L.M., Martino, M.E., et al. (2014). An Asian origin of virulent Aeromonas hydrophila responsible for disease epidemics in United States-farmed catfish. *MBio* 5(3)**,** e00848-00814. doi: 10.1128/mBio.00848-14.

Hossain, M.J., Waldbieser, G.C., Sun, D., Capps, N.K., Hemstreet, W.B., Carlisle, K., et al. (2013). Implication of lateral genetic transfer in the emergence of Aeromonas hydrophila isolates of epidemic outbreaks in channel catfish. *PLoS One* 8(11)**,** e80943. doi: 10.1371/journal.pone.0080943.

Hughes, H.Y., Conlan, S.P., Lau, A.F., Dekker, J.P., Michelin, A.V., Youn, J.H., et al. (2016). Detection and Whole-Genome Sequencing of Carbapenemase-Producing Aeromonas hydrophila Isolates from Routine Perirectal Surveillance Culture. *J Clin Microbiol* 54(4)**,** 1167-1170. doi: 10.1128/JCM.03229-15.

Jagoda, S.S., Tan, E., Arulkanthan, A., Kinoshita, S., Watabe, S., and Asakawa, S. (2014). Draft Genome Sequence of Aeromonas hydrophila Strain Ae34, Isolated from a Septicemic and Moribund Koi Carp (Cyprinus carpio koi), a Freshwater Aquarium Fish. *Genome Announc* 2(3). doi: 10.1128/genomeA.00572-14.

Janda, J.M., and Abbott, S.L. (2010). The genus Aeromonas: taxonomy, pathogenicity, and infection. *Clin Microbiol Rev* 23(1)**,** 35-73. doi: 10.1128/CMR.00039-09.

Konstantinidis, K.T., Ramette, A., and Tiedje, J.M. (2006). The bacterial species definition in the genomic era. *Philos Trans R Soc Lond B Biol Sci* 361(1475)**,** 1929-1940. doi: 10.1098/rstb.2006.1920.

Konstantinidis, K.T., and Tiedje, J.M. (2005a). Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci U S A* 102(7)**,** 2567-2572. doi: 10.1073/pnas.0409727102.

Konstantinidis, K.T., and Tiedje, J.M. (2005b). Towards a genome-based taxonomy for prokaryotes. *J Bacteriol* 187(18)**,** 6258-6264. doi: 10.1128/JB.187.18.6258-6264.2005.

Korotkov, K.V., Sandkvist, M., and Hol, W.G. (2012). The type II secretion system: biogenesis, molecular architecture and mechanism. *Nat Rev Microbiol* 10(5)**,** 336-351. doi: 10.1038/nrmicro2762.

Kulkarni, R., Dhakal, B.K., Slechta, E.S., Kurtz, Z., Mulvey, M.A., and Thanassi, D.G. (2009). Roles of putative type II secretion and type IV pilus systems in the virulence of uropathogenic Escherichia coli. *PLoS One* 4(3)**,** e4752. doi: 10.1371/journal.pone.0004752.

Leiman, P.G., Basler, M., Ramagopal, U.A., Bonanno, J.B., Sauder, J.M., Pukatzki, S., et al. (2009). Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *Proc Natl Acad Sci U S A* 106(11)**,** 4154-4159. doi: 10.1073/pnas.0813360106.

Lenneman, E.M., and Barney, B.M. (2014). Draft Genome Sequences of the Alga-Degrading Bacteria Aeromonas hydrophila Strain AD9 and Pseudomonas pseudoalcaligenes Strain AD6. *Genome Announc* 2(4). doi: 10.1128/genomeA.00709-14.

Letoffe, S., Delepelaire, P., and Wandersman, C. (1996). Protein secretion in gram-negative bacteria: assembly of the three components of ABC protein-mediated exporters is ordered and promoted by substrate binding. *EMBO J* 15(21)**,** 5804-5811.

McBride, M.J., and Nakane, D. (2015). Flavobacterium gliding motility and the type IX secretion system. *Curr Opin Microbiol* 28**,** 72-77. doi: 10.1016/j.mib.2015.07.016.

McBride, M.J., and Zhu, Y. (2013). Gliding motility and Por secretion system genes are widespread among members of the phylum bacteroidetes. *J Bacteriol* 195(2)**,** 270-278. doi: 10.1128/JB.01962-12.

Metcalf, W.W., Jiang, W., and Wanner, B.L. (1994). Use of the rep technique for allele replacement to construct new Escherichia coli hosts for maintenance of R6K gamma origin plasmids at different copy numbers. *Gene* 138(1-2)**,** 1-7.

Mougous, J.D., Cuff, M.E., Raunser, S., Shen, A., Zhou, M., Gifford, C.A., et al. (2006). A virulence locus of Pseudomonas aeruginosa encodes a protein secretion apparatus. *Science* 312(5779)**,** 1526-1530. doi: 10.1126/science.1128393.

Nguyen, L., Paulsen, I.T., Tchieu, J., Hueck, C.J., and Saier, M.H., Jr. (2000). Phylogenetic analyses of the constituents of Type III protein secretion systems. *J Mol Microbiol Biotechnol* 2(2)**,** 125-144.

Nielsen, M.E., Hoi, L., Schmidt, A.S., Qian, D., Shimada, T., Shen, J.Y., et al. (2001). Is Aeromonas hydrophila the dominant motile Aeromonas species that causes disease outbreaks in aquaculture production in the Zhejiang Province of China? *Dis Aquat Organ* 46(1)**,** 23-29. doi: 10.3354/dao046023.

Nivaskumar, M., and Francetic, O. (2014). Type II secretion system: a magic beanstalk or a protein escalator. *Biochim Biophys Acta* 1843(8)**,** 1568-1577. doi: 10.1016/j.bbamcr.2013.12.020.

Pan, H.J., Wu, S.Q., Dong, C.F., Shi, C.B., Ye, M.X., Lin, T.L., et al. (2004). Identification, virulence, hemolytic activity of GYK1, a strain of pathogenic Aeromonas hydrophila isolated from mandarinfish. *J. Shanghai Fish. Univ.* 13**,** 23–29.

Pang, M., Jiang, J., Xie, X., Wu, Y., Dong, Y., Kwok, A.H., et al. (2015). Novel insights into the pathogenicity of epidemic Aeromonas hydrophila ST251 clones from comparative genomics. *Sci Rep* 5**,** 9833. doi: 10.1038/srep09833.

Peabody, C.R., Chung, Y.J., Yen, M.R., Vidal-Ingigliardi, D., Pugsley, A.P., and Saier, M.H., Jr. (2003). Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. *Microbiology* 149(Pt 11)**,** 3051-3072. doi: 10.1099/mic.0.26364-0.

Pell, L.G., Kanelis, V., Donaldson, L.W., Howell, P.L., and Davidson, A.R. (2009). The phage lambda major tail protein structure reveals a common evolution for long-tailed phages and the type VI bacterial secretion system. *Proc Natl Acad Sci U S A* 106(11)**,** 4160-4165. doi: 10.1073/pnas.0900044106.

Perez-Cheeks, B.A., Planet, P.J., Sarkar, I.N., Clock, S.A., Xu, Q., and Figurski, D.H. (2012). The product of tadZ, a new member of the parA/minD superfamily, localizes to a pole in Aggregatibacter actinomycetemcomitans. *Mol Microbiol* 83(4)**,** 694-711. doi: 10.1111/j.1365-2958.2011.07955.x.

Podladchikova, O., Antonenka, U., Heesemann, J., and Rakin, A. (2011). Yersinia pestis autoagglutination factor is a component of the type six secretion system. *Int J Med Microbiol* 301(7)**,** 562-569. doi: 10.1016/j.ijmm.2011.03.004.

Pridgeon, J.W., Zhang, D., and Zhang, L. (2014a). Complete Genome Sequence of a Moderately Virulent Aeromonas hydrophila Strain, pc104A, Isolated from Soil of a Catfish Pond in West Alabama. *Genome Announc* 2(3). doi: 10.1128/genomeA.00554-14.

Pridgeon, J.W., Zhang, D., and Zhang, L. (2014b). Complete Genome Sequence of the Highly Virulent Aeromonas hydrophila AL09-71 Isolated from Diseased Channel Catfish in West Alabama. *Genome Announc* 2(3). doi: 10.1128/genomeA.00450-14.

Pukatzki, S., Ma, A.T., Sturtevant, D., Krastins, B., Sarracino, D., Nelson, W.C., et al. (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(5)**,** 1528-1533. doi: 10.1073/pnas.0510322103.

Qin, Y., Lin, G., Chen, W., Huang, B., Huang, W., and Yan, Q. (2014). Flagellar motility contributes to the invasion and survival of Aeromonas hydrophila in Anguilla japonica macrophages. *Fish Shellfish Immunol* 39(2)**,** 273-279. doi: 10.1016/j.fsi.2014.05.016.

Rao, P.S., Yamada, Y., Tan, Y.P., and Leung, K.Y. (2004). Use of proteomics to identify novel virulence determinants that are required for Edwardsiella tarda pathogenesis. *Mol Microbiol* 53(2)**,** 573-586. doi: 10.1111/j.1365-2958.2004.04123.x.

Rasmussen-Ivey, C.R., Hossain, M.J., Odom, S.E., Terhune, J.S., Hemstreet, W.G., Shoemaker, C.A., et al. (2016). Classification of a Hypervirulent Aeromonas hydrophila Pathotype Responsible for Epidemic Outbreaks in Warm-Water Fishes. *Front Microbiol* 7**,** 1615. doi: 10.3389/fmicb.2016.01615.

Rasmussen, M., Olsen, M.S., Sunde, L., Pedersen, L., and Petersen, O.B. (2016). Positive predictive value and completeness of prenatally assigned International Classification of Disease-10 kidney anomaly diagnoses in the Danish National Patient Registry. *Clin Epidemiol* 8**,** 9-14. doi: 10.2147/CLEP.S94394.

Richter, M., and Rossello-Mora, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A* 106(45)**,** 19126-19131. doi: 10.1073/pnas.0906412106.

Roach, D.J., Burton, J.N., Lee, C., Stackhouse, B., Butler-Wu, S.M., Cookson, B.T., et al. (2015). A Year of Infection in the Intensive Care Unit: Prospective Whole Genome Sequencing of Bacterial Clinical Isolates Reveals Cryptic Transmissions and Novel Microbiota. *PLoS Genet* 11(7)**,** e1005413. doi: 10.1371/journal.pgen.1005413.

Rondelet, A., and Condemine, G. (2013). Type II secretion: the substrates that won't go away. *Res Microbiol* 164(6)**,** 556-561. doi: 10.1016/j.resmic.2013.03.005.

Rosenzweig, J.A., and Chopra, A.K. (2013). Modulation of host immune defenses by Aeromonas and Yersinia species: convergence on toxins secreted by various secretion systems. *Front Cell Infect Microbiol* 3**,** 70. doi: 10.3389/fcimb.2013.00070.

Russell, A.B., Peterson, S.B., and Mougous, J.D. (2014a). Type VI secretion system effectors: poisons with a purpose. *Nat Rev Microbiol* 12(2)**,** 137-148. doi: 10.1038/nrmicro3185.

Russell, A.B., Wexler, A.G., Harding, B.N., Whitney, J.C., Bohn, A.J., Goo, Y.A., et al. (2014b). A type VI secretion-related pathway in Bacteroidetes mediates interbacterial antagonism. *Cell Host Microbe* 16(2)**,** 227-236. doi: 10.1016/j.chom.2014.07.007.

Sato, K., Yukitake, H., Narita, Y., Shoji, M., Naito, M., and Nakayama, K. (2013). Identification of Porphyromonas gingivalis proteins secreted by the Por secretion system. *FEMS Microbiol Lett* 338(1)**,** 68-76. doi: 10.1111/1574-6968.12028.

Seshadri, R., Joseph, S.W., Chopra, A.K., Sha, J., Shaw, J., Graf, J., et al. (2006). Genome sequence of Aeromonas hydrophila ATCC 7966T: jack of all trades. *J Bacteriol* 188(23)**,** 8272-8282. doi: 10.1128/JB.00621-06.

Sha, J., Rosenzweig, J.A., Kozlova, E.V., Wang, S., Erova, T.E., Kirtley, M.L., et al. (2013). Evaluation of the roles played by Hcp and VgrG type 6 secretion system effectors in Aeromonas hydrophila SSU pathogenesis. *Microbiology* 159(Pt 6)**,** 1120-1135. doi: 10.1099/mic.0.063495-0.

Shrivastava, A., Johnston, J.J., van Baaren, J.M., and McBride, M.J. (2013). Flavobacterium johnsoniae GldK, GldL, GldM, and SprA are required for secretion of the cell surface gliding motility adhesins SprB and RemA. *J Bacteriol* 195(14)**,** 3201-3212. doi: 10.1128/JB.00333-13.

Tan, W.S., Yin, W.F., and Chan, K.G. (2015a). Insights into the Quorum-Sensing Activity in Aeromonas hydrophila Strain M013 as Revealed by Whole-Genome Sequencing. *Genome Announc* 3(1). doi: 10.1128/genomeA.01372-14.

Tan, W.S., Yin, W.F., Chang, C.Y., and Chan, K.G. (2015b). Whole-Genome Sequencing Analysis of Quorum-Sensing Aeromonas hydrophila Strain M023 from Freshwater. *Genome Announc* 3(1). doi: 10.1128/genomeA.01548-14.

Tekedar, H.C., Karsi, A., Akgul, A., Kalindamar, S., Waldbieser, G.C., Sonstegard, T., et al. (2015). Complete Genome Sequence of Fish Pathogen Aeromonas hydrophila AL06-06. *Genome Announc* 3(2). doi: 10.1128/genomeA.00368-15.

Tekedar, H.C., Kumru, S., Kalindamar, S., Karsi, A., Waldbieser, G.C., Sonstegard, T., et al. (2017). Draft Genome Sequences of Three Aeromonas hydrophila Isolates from Catfish and Tilapia. *Genome Announc* 5(3). doi: 10.1128/genomeA.01509-16.

Tekedar, H.C., Kumru, S., Karsi, A., Waldbieser, G.C., Sonstegard, T., Schroeder, S.G., et al. (2016a). Draft Genome Sequence of Aeromonas hydrophila TN97-08. *Genome Announc* 4(3). doi: 10.1128/genomeA.00436-16.

Tekedar, H.C., Kumru, S., Karsi, A., Waldbieser, G.C., Sonstegard, T., Schroeder, S.G., et al. (2016b). Draft Genome Sequences of Four Virulent Aeromonas hydrophila Strains from Catfish Aquaculture. *Genome Announc* 4(4). doi: 10.1128/genomeA.00860-16.

Tekedar, H.C., Waldbieser, G.C., Karsi, A., Liles, M.R., Griffin, M.J., Vamenta, S., et al. (2013). Complete Genome Sequence of a Channel Catfish Epidemic Isolate, Aeromonas hydrophila Strain ML09-119. *Genome Announc* 1(5). doi: 10.1128/genomeA.00755-13.

Tomas, J.M. (2012). The main Aeromonas pathogenic factors. *ISRN Microbiol* 2012**,** 256261. doi: 10.5402/2012/256261.

Tomich, M., Planet, P.J., and Figurski, D.H. (2007). The tad locus: postcards from the widespread colonization island. *Nat Rev Microbiol* 5(5)**,** 363-375. doi: 10.1038/nrmicro1636.

Tran, N.T., Gao, Z.X., Zhao, H.H., Yi, S.K., Chen, B.X., Zhao, Y.H., et al. (2015). Transcriptome analysis and microsatellite discovery in the blunt snout bream (Megalobrama amblycephala) after challenge with Aeromonas hydrophila. *Fish Shellfish Immunol* 45(1)**,** 72-82. doi: 10.1016/j.fsi.2015.01.034.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., et al. (2012). Primer3--new capabilities and interfaces. *Nucleic Acids Res* 40(15)**,** e115. doi: 10.1093/nar/gks596.

Vilches, S., Urgell, C., Merino, S., Chacon, M.R., Soler, L., Castro-Escarpulli, G., et al. (2004). Complete type III secretion system of a mesophilic Aeromonas hydrophila strain. *Appl Environ Microbiol* 70(11)**,** 6914-6919. doi: 10.1128/AEM.70.11.6914-6919.2004.

Xu, L., Wang, H., Yang, X., and Lu, L. (2013). Integrated pharmacokinetics/pharmacodynamics parameters-based dosing guidelines of enrofloxacin in grass carp Ctenopharyngodon idella to minimize selection of drug resistance. *BMC Vet Res* 9**,** 126. doi: 10.1186/1746-6148-9-126.

Yang, W., Li, N., Li, M., Zhang, D., and An, G. (2016). Complete Genome Sequence of Fish Pathogen Aeromonas hydrophila JBN2301. *Genome Announc* 4(1). doi: 10.1128/genomeA.01615-15.

Yu, H.B., Rao, P.S., Lee, H.C., Vilches, S., Merino, S., Tomas, J.M., et al. (2004). A type III secretion system is required for Aeromonas hydrophila AH-1 pathogenesis. *Infect Immun* 72(3)**,** 1248-1256.

Zarivach, R., Vuckovic, M., Deng, W., Finlay, B.B., and Strynadka, N.C. (2007). Structural analysis of a prototypical ATPase from the type III secretion system. *Nat Struct Mol Biol* 14(2)**,** 131-137. doi: 10.1038/nsmb1196.

Zhang, X., Yang, W., Wu, H., Gong, X., and Li, A. (2014). Multilocus sequence typing revealed a clonal lineage of Aeromonas hydrophila caused motile Aeromonas septicemia outbreaks in pond-cultured cyprinid fish in an epidemic area in central China. *Aquaculture* 432**,** 1-6. doi: http://dx.doi.org/10.1016/j.aquaculture.2014.04.017.

Zhang, X.J., Yang, W.M., Li, T.T., and Li, A.H. (2013). Research on the genetic diversity and virulence characteristics of Aeromonas hydrophila isolated from fishponds with disease outbreaks in Hubei province. . 3**,** 458–466.

Zhao, Y.H., and Shaw, J.G. (2016). Cross-Talk between the Aeromonas hydrophila Type III Secretion System and Lateral Flagella System. *Front Microbiol* 7**,** 1434. doi: 10.3389/fmicb.2016.01434.

Zhou, C.E., Smith, J., Lam, M., Zemla, A., Dyer, M.D., and Slezak, T. (2007). MvirDB--a microbial database of protein toxins, virulence factors and antibiotic resistance genes for bio-defence applications. *Nucleic Acids Res* 35(Database issue)**,** D391-394. doi: 10.1093/nar/gkl791.

Zoued, A., Brunet, Y.R., Durand, E., Aschtgen, M.S., Logger, L., Douzi, B., et al. (2014). Architecture and assembly of the Type VI secretion system. *Biochim Biophys Acta* 1843(8)**,** 1664-1673. doi: 10.1016/j.bbamcr.2014.03.018.

CHAPTER V

COMPARATIVE GENOME ANALYSIS OF THE *AEROMONAS VERONII* STRAIN ML09-123 PROVIDES INSIGHTS INTO ITS PATHOGENICITY OF ORIGIN

5.1 Abstract

Aeromonas veronii infects various hosts including human and fish species. We sequenced the complete genome of an epidemic *A. veronii* strain ML09-123 from catfish and conducted comparative genome analysis to investigate the molecular basis of *A. veronii* pathogenesis. Based on our results, we noticed that the U.S. catfish isolate is highly similar to a recent Chinese isolate (*A. veronii* strain TH0426). Therefore, similar to the virulent *A. hydrophila* case, which is transferred from China to U.S., it seems that virulent *A. veronii* strain is also transferred from China to the U. S. aquaculture. Further, we extended our research by including all 41 *A. veronii* genomes available on the National Center for Biotechnology Information (NCBI). Comparative genome analysis of *A. veronii* from different locations and sources indicated that nonnative *A. veronii* species might have been transferred to different locations. Subsystem category distribution showed that majority of genomes show similar distribution. Type 1 Secretion System (T1SS), Type 2 Secretion System (T2SS), Type 4 pilus (T4P), and Flagellum core elements are conserved in all *A. veronii* genomes, whereas Type 3 Secretion System

(T3SS), Type 5 Secretion System (T5SS), Type 6 Secretion System (T6SS), and Tad adherence are the ones showing different distributions. Results revealed that mobile elements have different distributions depending on the location and organisms. China and U.S. isolates contain a large number of mobile elements. Overall, this study provides novel insights into the pathogenicity of *A. veronii* genomes.

5.2 Introduction

Aeromonas veronii is a Gram-negative, rod-shaped, mesophilic, motile bacterium in the *Aeromonadaceae* family, and it is widespread in aquatic environments (Janda and Abbott, 2010). *A. veronii* is one of the member of motile aeromonads that cause motile *Aeromonas* septicemia (MAS) in different organisms including including human (Janda and Abbott, 2010), ornamental and aquatic animals such as oscar cichlid (Sreedharan et al., 2011), tilapia (Hassan et al., 2017), sea bass (Uzun and Ogut, 2015), channel catfish (Liu et al., 2016), rainbow trout (Zepeda-Velazquez et al., 2015). Skin ulcers and systemic hemorrhagic septicemia are observed in sick fish (McGarey et al., 1991; Austin and Austin, 2012).

A. veronii cases have been reported from different regions of the world including USA (this study), China (Liu et al., 2016), and Saudi Arabia (Hassan et al., 2017). Previous studies have shown that *A. veronii* is known to be present in the gut of the host organisms. For example, metatranscriptomics (Bomar et al., 2011) and metagenomics (Maltz et al., 2014) studies from leech showed that *A. veronii* is the second symbiotic dominant organism in the gut microbiome.

Interestingly, the presence of *A. veronii* with other *Aeromonas* species has been reported from different cases such as *A. veronii-A. sobria* complex was identified from

soft shell disease of a turtle (Chen et al., 2013), and *A. veronii*-A. *jandai* complex was recovered from Nile tilapia (Dong et al., 2017). The significance of co-presence of two *Aeromonas* species in one host is still not clear, but it was reported that *A. veronii* might play a major role as a causative agent of the ulcerative syndrome in catfish (Cai et al., 2012).

It was reported that highly virulent clonal group of *A. hydrophila* has been causing severe economic losses in the cyprinid fish industry in China since 1989 (Zhang et al., 2014) and outbreaks in the channel catfish industry in the southeastern United States since 2009 (Pridgeon and Klesius, 2011). Further studies have shown that this clonal group was transferred from China to the USA recently (Hossain et al., 2014). A recent study has reported that *A. veronii* has been affecting cultured channel catfish industry in China (Liu et al., 2016). Similar to *A. hydrophila, A. veronii* might have been introduced to the USA from China.

In this study, we sequenced an epidemic *A. veronii* strain isolated from a catfish in the U.S. and investigated its genome. Total 41 publicly available *A. veronii* genomes, we applied comparative genomics approaches for a total 41 *A. veronii* genomes, which provides novel insights into the pathogenicity of *A. veronii* genomes from different hosts and geographical isolates. The virulences of epidemic *A. hydrophila* ML09-119 (Tekedar et al., 2013) and *A. veronii* ML09-123 were in catfish were also determined.

5.3 Material and methods

5.3.1 Bacterial strains and data source for comparative genome analysis

Virulent *A. veronii* ML09-123 and *A. hydrophila* ML09-119 are from outbreaks from commercial catfish farms. All *A. veronii* genomes (Table 5.1) were from the NCBI (as of 9/2/2017).

Table 5.1 *A. veronii* genomes used in comparative genomic analyses.

	Strain names	Country	Source	Level	Size (Mb)	GC%	S.	P	Accession	Ref
1	AMC34	USA	Homo sapiens	Scaffold	4.579	58.4	1		NZ_AGWU00000000.1	*
2	RU31B	N/A	N/A	Scaffold	4.534	58.7	93		NZ_FTMU00000000.1	N/A
3	TH0426	CHINA	Yellowhead catfish	Complete	4.923	58.3	1		NZ_CP012504.1	(Kang et al., 2016)
4	ML09-123	USA	Catfish	Contig	4.754	58.4	32		N/A	This study
5	Hm21	TURKEY	Hirudo verbana	Contig	4.685	58.7	50		NZ_ATFB00000000.1	(Bomar et al., 2013)
6	A29	S. AFRICA	Surface water	Scaffold	4.482	58.8	54		NJGB00000000.1	N/A
7	AER39	USA	Homo sapiens	Scaffold	4.421	58.8	4		NZ_AGWT00000000.1	*
8	CB51	CHINA	Grass carp	Complete	4.584	58.6	1		CP015448	N/A
9	Ae52	SRI LANKA	Goldfish	Contig	4.565	58.7	80		BDGY00000000.1	(De et al., 2017)
10	LMG 13067	USA	N/A	Scaffold	4.736	58.4	72		NZ_CDBQ00000000.1	N/A
11	AVNIH2	USA	Homo sapiens	Contig	4.523	58.9	50		NZ_LRBO00000000.1	N/A
12	AVNIH1	USA	Homo sapiens	Complete	4.955	58.5	2	1	NZ_CP014774.1 NZ_CP014775.1	N/A
13	AMC35	USA	Homo sapiens	Scaffold	4.566	58.5	2		NZ_AGWW00000000.1	*
14	CECT 4257	USA	Homo sapiens	Scaffold	4.516	58.9	52		NZ_CDDK00000000.1	N/A
15	CCM 4359	USA	Homo sapiens	Contig	4.511	58.9	56		NZ_MRZR00000000.1	N/A
16	B565	CHINA	pond sediment	Complete	4.552	58.7	1		NC_015424	(Li et al., 2011)
17	AER397	USA	Homo sapiens	Scaffold	4.497	58.8	5		NZ_AGWV00000000.1	*
18	TTU2014-143AME	USA	Dairy Cattle	Contig	4.681	58.6	59		NZ_LKKG00000000.1	(Webb et al., 2016)
19	TTU2014-134ASC	USA	Dairy Cattle	Contig	4.680	58.6	59		NZ_LKKB00000000.1	(Webb et al., 2016)
20	TTU2014-142ASC	USA	Dairy Cattle	Contig	4.681	58.6	45		NZ_LKKF00000000.1	(Webb et al., 2016)
21	TTU2014-130AME	USA	Dairy Cattle	Contig	4.679	58.6	64		NZ_LKJW00000000.1	(Webb et al., 2016)
22	TTU2014-125ASC	USA	Dairy Cattle	Contig	4.680	58.6	58		NZ_LKJV00000000.1	(Webb et al., 2016)
23	TTU2014-130ASC	USA	Dairy Cattle	Scaffold	4.680	58.6	49		NZ_LKJX00000000.1	(Webb et al., 2016)
24	TTU2014-141AME	USA	Dairy Cattle	Scaffold	4.680	58.6	48		NZ_LKKD00000000.1	(Webb et al., 2016)
25	TTU2014-134AME	USA	Dairy Cattle	Contig	4.681	58.6	50		NZ_LKKA00000000.1	(Webb et al., 2016)
26	TTU2014-143ASC	USA	Dairy Cattle	Contig	4.678	58.6	54		NZ_LKKH00000000.1	(Webb et al., 2016)

Table 5.1 (continued)

27	TTU2014-131ASC	USA	Dairy Cattle	Contig	4.675	58.6	70	NZ_LKJY00000000.1	(Webb et al., 2016)
28	TTU2014-141ASC	USA	Dairy Cattle	Contig	4.680	58.6	45	NZ_LKKE00000000.1	(Webb et al., 2016)
29	TTU2014-140ASC	USA	Dairy Cattle	Contig	4.676	58.6	81	NZ_LKKC00000000.1	(Webb et al., 2016)
30	TTU2014-113AME	USA	Dairy Cattle	Scaffold	4.663	58.6	122	NZ_LKJQ00000000.1	(Webb et al., 2016)
31	pamvotica	GREECE	Surface sentiment	Contig	4.919	58.1	21	NZ_MRUI00000000.1	N/A
32	TTU2014-115ASC	USA	Dairy Cattle	Contig	4.533	58.7	52	NZ_LKJS00000000.1	(Webb et al., 2016)
33	TTU2014-115AME	USA	Dairy Cattle	Scaffold	4.532	58.7	53	NZ_LKJR00000000.1	(Webb et al., 2016)
34	TTU2014-108AME	USA	Dairy Cattle	Contig	4.533	58.7	62	NZ_LKJN00000000.1	(Webb et al., 2016)
35	TTU2014-108ASC	USA	Dairy Cattle	Contig	4.533	58.7	58	NZ_LKJP00000000.1	(Webb et al., 2016)
36	VBF557	INDIA	Homo sapiens	Contig	4.697	58.4	526	LXJN00000000.1	N/A
37	ARB3	JAPAN	Pond water	Contig	4.543	58.8	63	NZ_JRBE00000000.1	(Kenzaka et al., 2014)
38	CIP 107763	USA	N/A	Contig	4.431	58.8	64	NZ_CDDU00000000.1	N/A
39	PhIn2	INDIA	Fish intestine	Contig	4.301	58.8	1899	ANNT01000667.1	N/A
40	CECT 4486	USA	Surface water	Scaffold	4.411	58.9	66	NZ_CDBU00000000.1	N/A
41	CCM 7244	GERMANY	Surface water	Contig	4.422	58.9	74	NZ_MRZQ00000000.1	N/A

N/A: Not available, S: Scaffold, P: Plasmid, Ref: Reference
*Human Microbiome U54 initiative, Broad Institute (broadinstitute.org)

5.3.2 Sequencing, assembly, and annotation

A. veronii ML09-123 strain was sequenced using an Illumina Genome Analyzer IIx. The total number of reads were 4,911,312 and coverage was 118X. Adaptor trimming, quality control of sequence reads, contig creation and *de novo* assembly were conducted using CLC Workbench version 6.5.1. (ClC Bio) and Sequencher version 5.4 (Gene Codes Corporation). For annotation, the draft genome was submitted to RAST (Aziz et al., 2008) and NCBI`s Prokaryotic Genome Automatic Annotation Pipeline (PGAAP).

5.3.3 Phylogenetic tree and ANI calculation

Phylogenetic tree for 41 *A. veronii* genomes, build from core genomes. The gene sets of the core genome were aligned using MUSCLE (Edgar, 2004) one by one, and alignments of the each genome were concatenated. This alignment was used to compute a Kimura distance matrix, which was used as input for the Neighbor-Joining algorithm as implemented in PHYLP (Felsenstein, 1989)

5.3.4 Pan-Core genome

The core- and pan-genome analyses were performed using the EDGAR 2.0 platform. It is a generic orthology criterion which in turn is based off BLAST score ratio values (SRV), i.e., BLAST scores are normalized in relation to the best hit possible, Based on the distribution of SRVs in the dataset, a cutoff is estimated as described in (Blom et al., 2016). Comparative subsystem coverage identification was conducted by submitting nucleotide sequences of 41 *A. veronii* genomes to RAST annotation server for SEED subsystem categorization (Aziz et al., 2008). Comparative subsystem coverage identification was conducted by submitting nucleotide sequences of 41 *A. veronii* genomes to RAST annotation server for SEED subsystem categorization (Aziz et al., 2008).

5.3.5 Subsystems coverage

Comparative subsystem coverage identification was conducted by submitting nucleotide sequences of 41 A. veronii genomes to RAST annotation server for SEED subsystem categorization (Aziz et al., 2008).

5.3.6 Secretion systems

Nucleotide sequences of 41 genomes were annotated by RAST annotation server (Aziz et al., 2008), then annotated protein files were submitted to MacSyFinder (Abby et al., 2014; Abby et al., 2016), which uses systems modeling and similarity search to identify protein secretion systems and related appendages. Default settings were used: the type of dataset to deal with: ordered/unordered replicon, the topology of the replicon (s): linear/circular, protein secretion systems to detect: all available systems, maximal e-value: 1.0, maximal independent e-value: 0.001, and minimal profile coverage: 0.5.

5.3.7 Insertion elements

Insertion elements were determined by submitting nucleotide files to Issaga (Varani et al., 2011). After the identification process, results were organized by including complete, partial and unknown regions and removing the false predicted insertion from the final results.

5.3.8 Phage elements

To detect the prophages in the evaluated *A. veronii* genomes, PHASTER (PHAge Search Tool Enhanced Release) was used (Arndt et al., 2016). First, nucleotides from all genomes were concatenated to serve as an input file and submitted to PHASTER server. Computed results were arranged into three categories: score > 90 was considered intact phage element; a score between 70 – 90 was found to be questionable; and score < 70 is considered incomplete phage region. Additionally, nucleotide sequences of all the identified phage regions were concatenated and performed MAUVE alignment (Darling et al., 2004) to identify conserved phage regions.

5.3.9 Integron identification

Identification of integron and their components for the evaluated *A. veronii* genomes were performed by IntegronFinder (Cury et al., 2016).

5.3.10 CRISPR (Clustered regularly interspaced short palindromic repeats) and Cas (CRISPR associated genes) elements analysis

To evaluate the CRISPR-Cas systems and their distribution in *A. veronii* genomes, CRISPRfinder (Grissa et al., 2007b) and MacSyFinder (Abby et al., 2014) were used. For the Cas elements identification, following criteria were used, maximal e-value was set to 1.0, independent e-value was set to 0.001, and minimal profile coverage was set to 0.5. Three different categories were obtained (*mandatory*, *accessory*, *forbidden*). Mandatory elements are identifiable and ubiquitous; accessory elements could be essential but not identifiable. If the identified element is partly homologous, it is considered forbidden elements.

5.3.11 Antibiotic resistance genes

Concatenated nucleotide files were submitted to CARD (Comprehensive Antibiotic Resistance Database) (Jia et al., 2017) to perform BLAST search. Results were organized, and duplicates were removed.

5.3.12 Fish challenge

Fish experiments were conducted under a protocol approved by the Institutional Animal Care and Use Committee at Mississippi State University. Virulences of *A. hydrophila* ML09-119 (Tekedar et al., 2013) and *A. veronii* ML09-123 were tested. Briefly, 12 month-old specific-pathogen-free (SPF) channel catfish (18.2 ± 0.53 cm, 82.3

± 6.39 g) were stocked into twenty-seven 40-liter flow-through tanks (10 fish/tank) and acclimated for four days. Tanks were assigned randomly to two treatment groups (*A. veroni* and *A. hydrophila*), and four doses (1×10^4, 1×10^5, 1×10^6, 1×10^7) Three replicate tanks were assigned to each dose, and a negative control group was also included. Water temperature was maintained at 32ºC (±2) during the experiments. Fish were fed twice a day with a commercial catfish feed. Fish mortalities were recorded daily for a total of seven days; the experiment was terminated when no fish mortality was observed for three consecutive days. Percent mortalities were calculated for each group.

5.4 Results

5.4.1 Genome features

A. veronii strains used in this study and their genome features are listed in table 1. G+C ratios range from 58.1-58.9, and only one genome (strain AVNIH1) carries a plasmid. Other genomes may have plasmids, which needs experimental confirmation. These *A. veronii* strains are isolated from different hosts including human, cattle, fish, water, and sediment and different geographical locations, including USA, China, Germany, Sri Lanka, Japan, India, South Africa, Turkey, and Greece (Table 5.1).

5.4.2 Phylogenetic tree and ANI calculation

To identify to taxonomic positions of evaluated *A. veronii* genomes, a phylogenetic tree was constructed based on the complete core genome of 41 *A. veronii* genomes. Phylogenetic tree for 41 *A. veronii* genomes, build out of a core of 2409 genes per genome, 8769 in total. The core has 771557 AA-residues/ bp per genome, 31633837 in total. The result did not show highly conserved branching except for a couple of them.

For example, U.S. (ML09-123) and China (TH0426) catfish isolates. Greece surface sediment isolate lands in between U.S. cattle isolates. U.S. human isolates (strains CECT 4257, CCM 4359, AER 397) and China pond sediment isolate strain B565 are grouped. Finally, U.S. surface water and Germany surface water isolates formed highly conservative groups. Importantly, these findings were confirmed and supported by Average Nucleotide Identity (ANI), conserved branches` ANI values were above 99.91. The Greece isolate strain pamvotica fall in between the cattle isolates, and their ANI value is slightly lower than other conserved strains (Figure 5.1). To be considered same species, ANI calculation ratio should be above 95. According to our ANI calculation, strain AMC34 slightly lower than 95, which shows that it might be another member of *Aeromonas spp.*

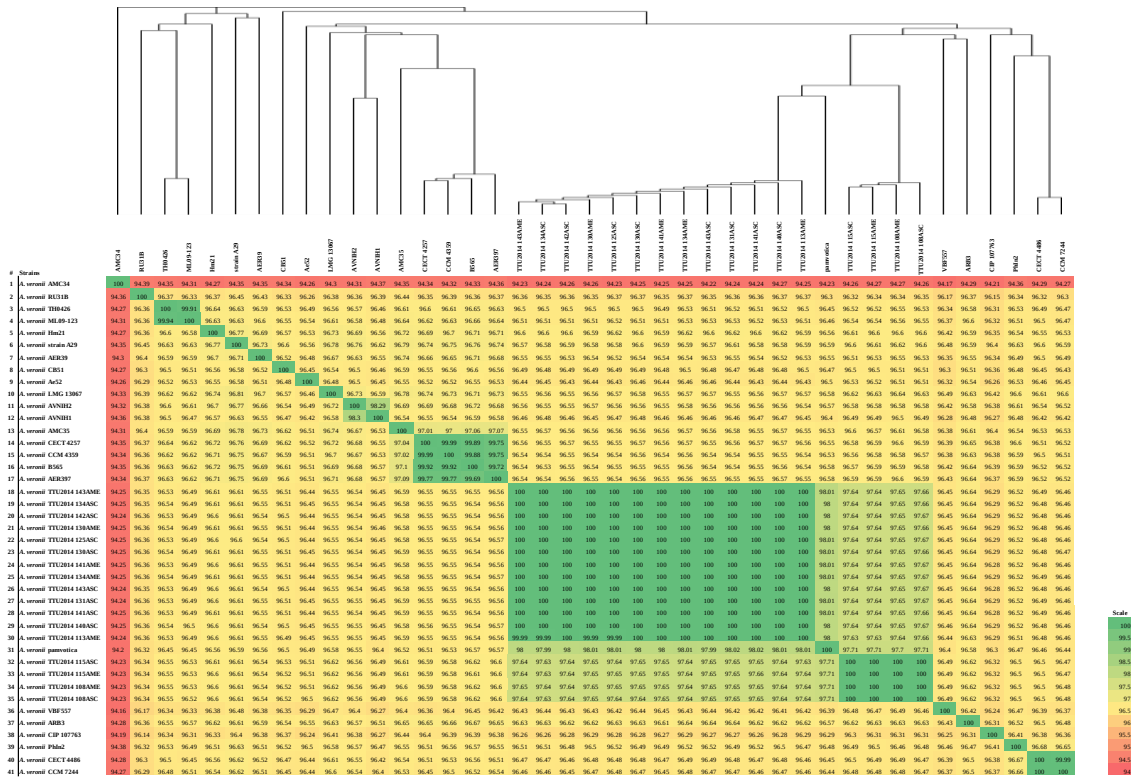


Figure 5.1 Average nucleotide identity (ANI) values of *A. veronii* genomes and phylogenetic tree analysis based on the core genomes.

5.4.3 Pan-Core genome analyses

Examination of Pan-core genome analysis revealed that addition of each genome to pan-genome analysis increased the starting number of pan-genome development from 4,072 genes to 9,120 total genes for 41 *A. veronii* genomes. On the other hand, core genome calculation showed the addition of each genome decreased the total number of core elements from 4,072 genes to 2,399 genes (Figure 5.2). Lack of a quality genome for strain Phln2 (has 1,899 contigs), core genome number seems to decrease the core genome number from around 3000 genes to 2400 genes.

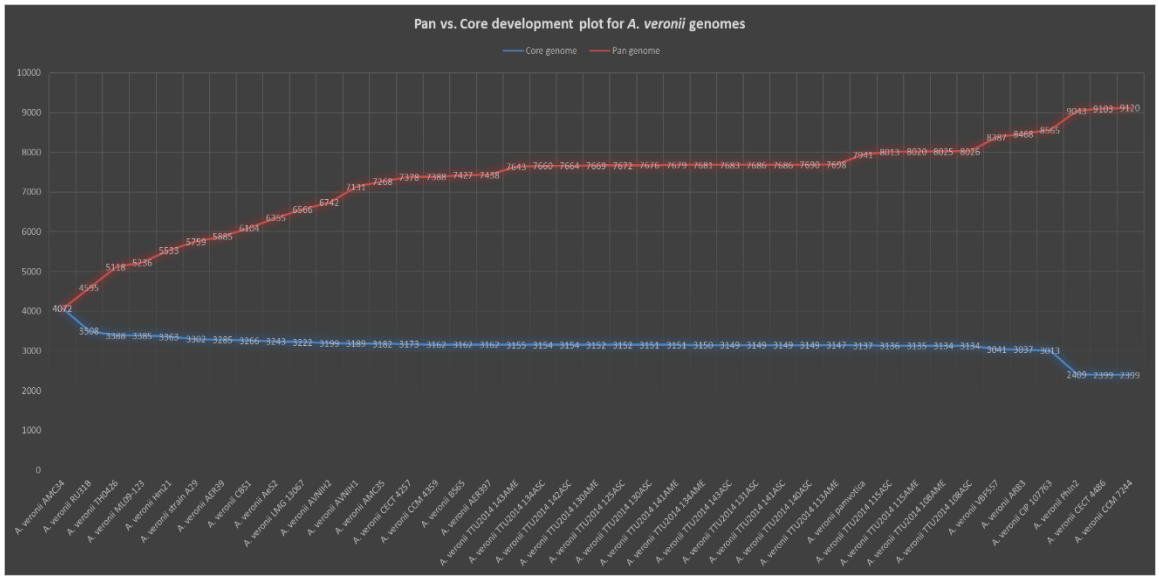


Figure 5.2 Pan vs. Core genome development plot for *A. veronii* genomes.

5.4.4 Subsystems coverage

The Rapid annotation Subsystem Technology (RAST) annotation server predicted all the evaluated 41 *A. veronii* genomes` subsystems listed in figure 5.3. SEED subsystem categorization analysis predicted that the Chinese isolate strain TH0426 carry the most elements (3559), whereas strain Phln2 carry the least of the elements (2576). This low number would be due to lack of assembly quality because this genome has 1899 contigs. Based on SEED subsystem, the most abundant system is that "amino acid and derivatives" biosynthesis and utilization. The second abundant system is "carbohydrates", followed by "cofactors, vitamins, prosthetic groups, and pigments". Interestingly, USA catfish isolate strain ML09-123 and Chinese catfish isolate strain TH0426 carry the most abundant "phages, prophages, transposable elements, and plasmids". SEED subsystem analysis enabled us to make coherent conclusions to understand the distribution of the subsystem for the *A. veronii* genomes.

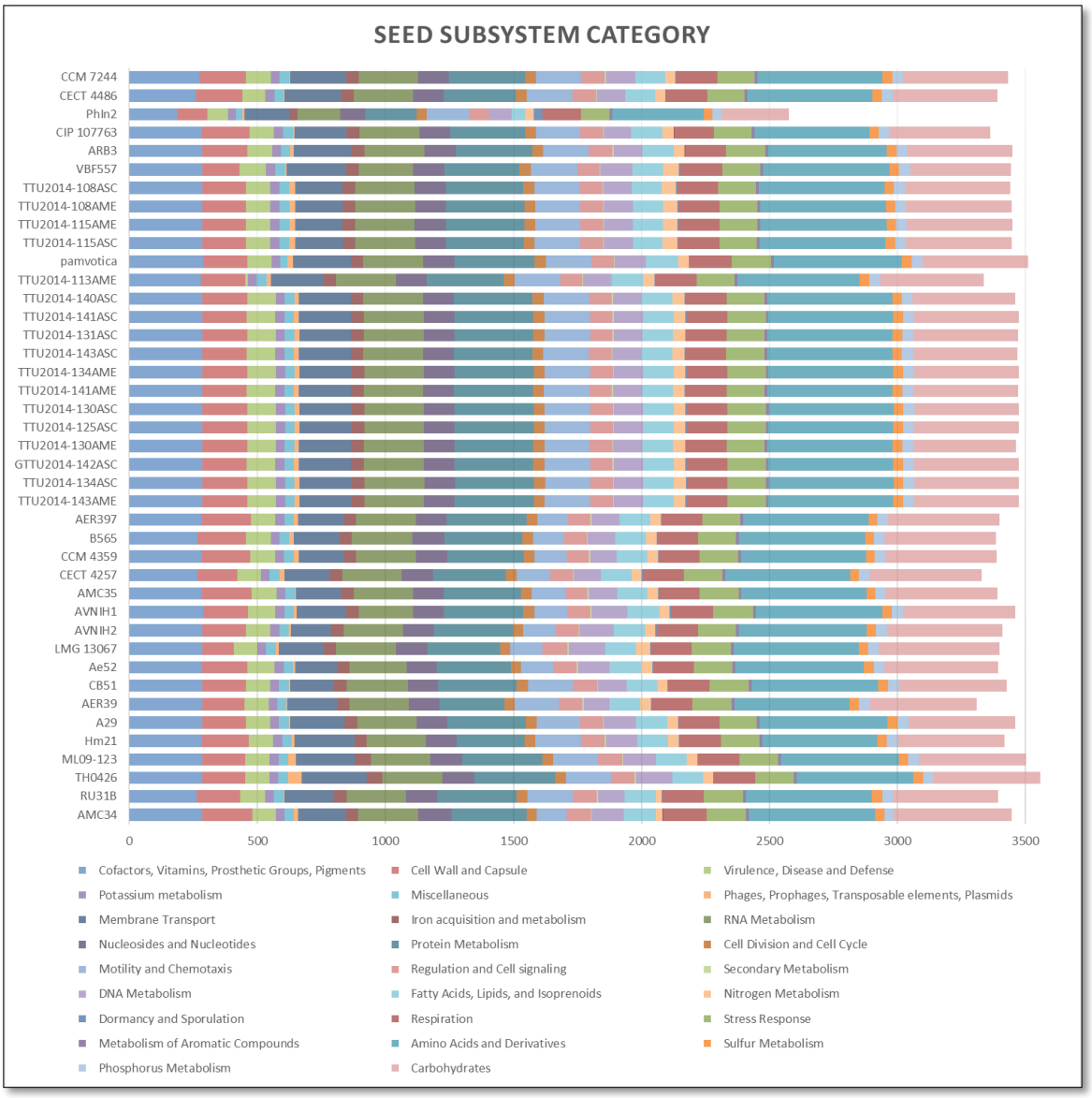


Figure 5.3 SEED subsystem category for *A. veronii* genomes.

Comparison of functional categories in 41 A. veronii genomes based on SEED. Functional categorization based on roles of annotated and assigned genes. Each colored bars represent the number of genes assigned to each category.

5.4.5 Secretion systems, Flagellum, TAD, T4P

In our silico secretion system analysis for *A. veronii* genomes, we identified that all the evaluated genomes encode T1SS, T2SS, T4P, Flagellum core components. In all these systems, a couple of genes were not present, but this could be due to lack of annotation or assembly. On the other hand, our genome (strain ML09-123) carry all the systems except TAD, which is similar to Chinese fish isolate (strain TH0426) (Figure 5.4). Additionally, the majority of *A. veronii* genomes encode T3SS. However, one Sri Lankan fish isolate (strain Ae52), one pond sediment isolate (strain B565) from China and seven human isolates (strains AMC34, AVNIH2, AVNIH1, AMC35, CECT4257, CCM 4359, AER397) and one unknown source isolate from the USA do not encode T3SS (Figure 4). For the T5aSS and T5bSS systems, only 17 dairy cattle isolates from the USA and one isolate (strain pamvotica) from Greece encode both systems, whereas rest of the genomes either carry one gene from each system or none (Figure 4). For the T4SS, 13 cattle diary isolates encode 8 of the T4SS-Type T accessory genes, whereas four strains (TTU2014-108AME, TTU2014-108ASC, TTU2014-115AME, TTU2014-115ASC) do not encode this system, also strain CIP107163, AVNIH1 and pamvotica encode some of the T4SS-type T components. Also, strain AVNIH1 possess a large size plasmid, and identification results showed that this plasmid encodes several of the elements of the T4SS-type F. Also, strains VBF557 and pamvotica genomes encode some of the T4SS-type F.

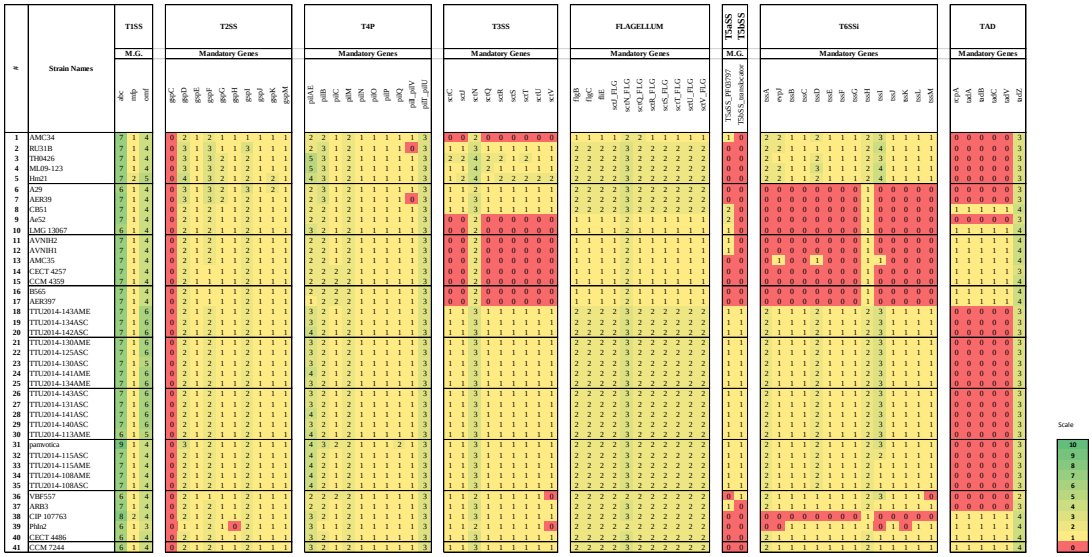


Figure 5.4 Core components of secretion systems elements, T4P, Tad and flagella availability in *A. veronii* genomes.

Numbers represent the presence of each gene.

Moreover, our results identified that *A. veronii* genomes either carry all the mandatory genes of T6SSi or one gene, *tssH*, from this system. T6SSi shows similar distribution as in T3SS. Strain A29, AER39, CB51 and CIP 107763 have T3SS do not encode T6SSi system, whereas AMC34 strain carries this T6SSi system (Figure 4). Tad system is another less conserved system among the evaluated genomes because only 13 out of 41 genomes encode this system (Figure 4).

5.4.6 Insertion elements

Our *in silico* insertion analyses identified that all the evaluated *A. veronii* genomes carry insertion elements. Number of the carried insertion elements differ from 16 to 163. Strain AVNIH1 has the highest number of insertion elements, whereas strain A29 carry the least number of insertions. Specifically, IS1595-subtype-ISPna2, IS51, IS3,

IS2, IS481, IS4, IS903, ISL3 insertion families, and subgroups are encoded by most of the evaluated *A. veronii* genomes.

5.4.7 Phage elements

Prophages in the 41 *A. veronii* genomes were investigated by PHAST; results showed that only two of the strains CECT4486 and CCM7244 do not carry any type (intact, questionable, incomplete) phages regions, whereas rest of the genomes have at least one type of identified phage regions. The general G+C regions for the *A. veronii* regions varies from 46.99-63.41. Specifically, since strains ML09-123 and TH0426 may have originated from the same clonal group as they tend to show similar G+C ratio in their phage elements, ranging from 57.55-63.41. Similarly, Greece isolate strain pamvotica, and U.S. cattle isolates show similar GC ratio distribution in their phage elements.

5.4.8 Integron elements

Integron regions consist of two main components, first integron-integrase (*intl*) and responsible promoter region (P_{intI}), attachment site of the integron (*attI*), and constitutive promoter (Pc) region to integrate the gene cassettes at the *attI* region. Secondly, gene cassettes (up to 200 bp), are opposite in the direction of the integron-integrase gene. Gene cassettes have an open reading frame (ORF) and followed by *attC* gene in their flanking regions, which helps them to attach to cassettes. Note that, ORF region is not mandatory. Briefly, complete integron elements include integrase and at least one *attC* region. ln0 elements consist of integrase without attC sites, and the clusters of the *attC* region without integron-integrase and encoding at least two *attC* regions is

considered (CALIN) (Collis and Hall, 1995; Boucher et al., 2007; Joss et al., 2009; Cury et al., 2016). According to our integron identification results, only two strains (AVNIH1 and Ae52) encode complete integron regions, whereas rest of the genomes either carry cluster of *attC* sites lacking integron-integrases (CALIN) or none of the integron or any part of integron regions. Moreover, only 4 strains CECT 4486, LMG 13067, CCM 7244, RU31B have two CALIN regions, and 14 strains (CCM 4359, ML09-123, CB-51, TH0426, CECT 4257, CIP 107763, AMC34, AMC35, AER397, B565, TTU2014-115ASC, TTU2014-108ASC, TTU2014-108AME, TTU2014-115AME) do not encode any of the integron components, whereas rest of the strains carry only CALIN elements.

5.4.9 CRISPR-Cas elements

Our CRISPR analysis identified that all the evaluated *A. veronii* genomes carry at least questionable CRISPR regions, while strains A29, LMG 13067, AVNIH2, AVNIH1, AMC35 genomes have confirmed CRISPR regions. Cas systems are divided into three categories (Type-I, Type-II, and Type-III) and these three categories has 10 different subcategories (TypeI-A to F, TypeII-A and B, TypeIII-A and B) for the Cas regions. A few of the *A. veronii* genomes encode these elements. Strains AMC34, TH0426, LMG 13067, AVNIH1 and AVNIH2 genomes carry these genes, rest of the genomes do not carry Cas elements (Table 5.2).

Table 5.2 General features of CRISPR-Cas loci in 41 *A. veronii* genomes.

#	Strain names	CRISPR regions and their features					Cas elements availability	
		Confirmed	Questionable	Direct repeat length	Number of spacers	CRISPR length	Cas-Type I	CAS-Type III
1	AMC34	-	3	23-40	1	91-109	cas3-TypeI cas3-TypeI	-
2	RU31B	-	4	23-42	1	86-121		-
3	TH0426	-	7	23-42	1	99-139	cas6-TypeIF csy3-TypeIF	- -
4	ML09-123	-	7	23-48	1	98-133	-	-
5	Hm21	-	3	24-35	1	99-113	-	-
6	A29	2	4	24-43	1-9	81-684	-	csx16-TypeIIIU cas1-TypeII
7	AER39	-	4	24-34	1	105-107	-	-
8	CB51	-	3	45-55	1	147-166	-	-
9	Ae52	-	3	24-45	1	106-117	-	-
10	LMG 13067	1	3	32-45	1-42	106-2804	cas1-TypeIC cas4-TypeI-II cas7c-TypeIC cas8c-TypeIC cas5c-TypeIC	-
11	AVNIH2	1	5	23-45	1-72	80-4430	cas2-TypeIE cas1-TypeIE cas5-TypeIE cas7-TypeIE cse2-TypeIE cse1-TypeIE cas3-TypeI cas6-TypeIE	-
12	AVNIH1	1	7	23-46	1-4	80-250	cas3-TypeI cas3-TypeI	-
13	AMC35	1	8	23-51	1-5	80-371	-	-
14	CECT 4257	-	6	23-53	1	95-158	-	-
15	CCM 4359	-	4	23-53	1	95-158	-	-
16	B565	-	5	23-53	1	95-158	-	-
17	AER397	-	7	24-53	1	81-127	-	-
18	TTU2014-143AME	-	5	24-40	1	96-118	-	-
19	TTU2014-134ASC	-	5	24-40	1	96-118	-	-
20	TTU2014-142ASC	-	5	24-40	1	96-118	-	-
21	TTU2014-130AME	-	5	24-40	1	96-118	-	-
22	TTU2014-125ASC	-	5	24-40	1	96-118	-	-
23	TTU2014-130ASC	-	7	24-40	1	96-118	-	-
24	TTU2014-141AME	-	5	24-40	1	96-118	-	-
25	TTU2014-134AME	-	5	2440	1	96-118	-	-
26	TTU2014-143ASC	-	5	24-40	1	96-118	-	-
27	TTU2014-131ASC	-	5	24-40	1	96-118	-	-
28	TTU2014-141ASC	-	5	24-40	1	96-118	-	-
29	TTU2014-140ASC	-	5	24-40	1	96-118	-	-
30	TTU2014-113AME	-	6	25-42	1-2	96-258	-	-
31	pamvotica	-	6	23-34	1-4	72-235	-	-
32	TTU2014-115ASC	-	4	24-39	1	96-124	-	-
33	TTU2014-115AME	-	4	24-48	1	96-133	-	-
34	TTU2014-108AME	-	4	24-39	1	96-124	-	-
35	TTU2014-108ASC	-	4	24-39	1	96-124	-	-
36	VBF557	-	5	24-34	1	96-107	-	-
37	ARB3	-	5	24-35	1	96-108	-	-
38	CIP 107763	-	2	25-34	1	96-105	-	-
39	PhIn2	-	2	30-35	1	106-115	-	-
40	CECT 4486	-	7	23-37	1	76-115	-	-
41	CCM 7244	-	7	23-37	1	79-124	-	-

5.4.10 Antibiotic resistance genes

According to antibiotic resistance database search, AVNIH1 genome has the largest number of the elements. Also, strain ML09-123 and TH0426 show similar distribution along with the majority of all the evaluated all the *A. veronii* genomes.

Following are the conserved antibiotic resistance elements in 41 *A. veronii* genomes that we identified in our study; MdtC (yegO), OXA-12 (AsB1), carbapenem elements, smeE element, edeine, kanamycin, tobramycin, neomycin, amikacin, gentamicin, rifampicin, aminocoumarin, fluoroquinolone, kasugamycin, kirromycin, Pulvomycin, Enacyloxin IIa, spectinomycin, gentamicin C, G418, paromomycin, streptomycin, tetracycline, hygromycin B, viomycin, neomycin, macrolide, clarithromycin, linezolid.

5.4.11 Fish challenge

Results of the injection challenge in catfish indicated that *A. veronii* ML09-123 shows a typical disease progress, showing mortality for each dose as following 10^4(3.33%), 10^5(36.67%), 10^6(86.67%), 10^7(93.33%). *A. hydrophila* ML09-119 result is different from *A. veronii* because there was no mortality at dose 10^4 and 10^5 but we lost majority or all of the fish at dose 10^6 (90%) and 10^7 (100%). Results are shown in Figure 5.5.

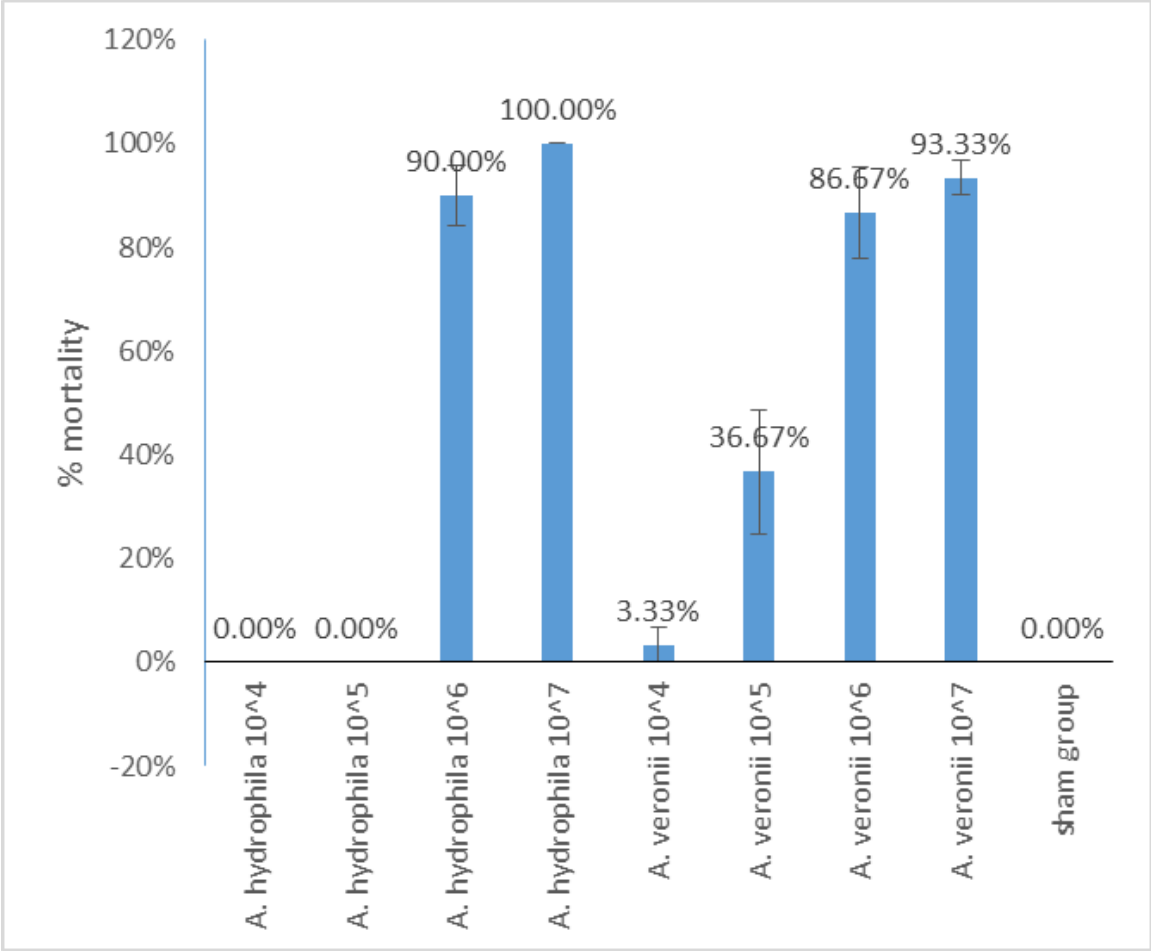


Figure 5.5 Percent mortalities in catfish challenged with *A. hydrophila* ML09-119 and *A. veronii* ML09-123.

5.5 Discussion

In our study, we sequenced one of the epidemic *A. veronii* isolates. To investigate the molecular mechanisms, we applied several aspects of comparative genomic approaches. For the first time, we report the presence of *A. veronii* strain ML09-123 in the U.S. aquaculture industry. This may be as a result of transfer from China to U.S. As similarly reported by several researchers, clonal group of *A. hydrophila* has been negatively affecting the U.S. aquaculture due to the transfer from China to U.S. Our

findings suggests that during this transfer not only *A. hydrophila* but also several *Aeromonas spp.* may have been introduced to U.S. aquaculture. Additionally, we extended our analysis using comparative genomics approaches to understand the molecular mechanism of *A. veronii* by including all the available *A. veronii* genomes on NCBI. Moreover, to investigate the virulence of *A. veronii* ML09-123 strain, we did fish challenge against to well-studied epidemic isolate *A. hydrophila* ML09-119. All in all, our findings strongly suggest that *A. veronii* has potential to cause serious problems in the aquaculture, and our comparative genomics will provide a basis for further studies to understand the molecular mechanism of *A. veronii* strains.

As it has been recommended to actively use ANI calculation for accurate digital investigation for genome distances (Kim et al., 2014), we applied this approach to see how our *A. veronii* ML09-123 genome fall into position along with all the publicly available *A. veronii* genomes. Our ANI calculation result showed that there are four groups that had a highly conservative similarity (above %99). Interestingly, all the four group members are isolated from different parts of the world suggesting potential transfer from one location to another location. To confirm the ANI calculation results, we performed core-genome based phylogenetic tree analysis, and our results indicated that ANI calculation and phylogenetic tree creation are consistent.

In our Pan-Core genome analysis, we observed that there is no drastic reduction also to each genome except strain Phln2, but this would be the lack of genome quality. Meanwhile pan-genome ratio is increased by the addition of each genome. This result would be explained by the acquisition of genetic material elements intake for their

genomes. Therefore we could conclude that as expected genome structure of each *A. veronii* strain, is being changed evolutionary.

According to SEED subsystem analysis, Chinese catfish isolate strain TH0426 carry the most identified elements in its genome. This would be the result of abundant mobile elements presented in its genome because U.S. catfish isolate strain ML09-123 and China catfish isolate strain TH0426 carry the most abundant category "phages, prophages, transposable elements and plasmids" (Figure-2). SEED subsystem analysis enabled us to come to coherent conclusions to understand the distribution of the subsystem for the *A. veronii* genomes and help us to understand that mobile elements affect general structure of the genome.

Comparative secretion system analysis provided valuable information about the distribution of secretion systems and their components available in the evaluated *A. veronii* genomes. Due to the importance of several roles in the organism, including invading hosts, damaging tissues, secreting a variety of substrates, bacterial competition secretion systems, these systems play an important role for the bacteria (Green and Mecsas, 2016). Interestingly, in our secretion systems analysis, we observed that they are four conserved groups according to ANI calculation and phylogenetic tree analysis. These four groups' secretion system distributions are similar, which support the ANI calculation and phylogenetic tree creation (Figure-4).

T1SS is known for one-step transportation for unfolded substrates directly into extracellular space by bypassing the periplasm (Kanonenberg et al., 2013), also made up of three main proteins (abc, omp, mfp) which span the cell envelope. In our analysis, we identified that all the evaluated *A. veronii* genomes possess the T1SS, but number of

mandatory elements vary depending on the strains. Since, this system is responsible for delivering destructive enzymes and toxins, encoding this system may give *A. veronii* species an advantage to invade the host and compete with other niche members. For instance, encoding this system has been confirmed in different bacterial pathogen organisms that presence of T1SS increases the virulence as observed in *Serrattia marcescens* (Letoffe et al., 1996) and *Vibrio cholera* (Dolores et al., 2015) (Figure 4).

Another important well-conserved secretion system in all the evaluated 41 genomes is the T2SS system, this system is known for exporting hydrolytic enzymes, some of the virulence factors and toxins from the studies of *A. hydrophila* (Ast et al., 2002; Cianciotto, 2005). One of the *A. veronii* studies revealed that presence of *A. veronii* in the gut of Leech helps to colonize because export of hemolysin by T2SS is an important step for the initial colonization of the gut environment (Maltz et al., 2011) (Figure 4). Our findings also suggest that regardless of the isolation location or organisms of evaluated all the genomes, T2SS is one of the well-conserved systems, which may suggest that *A. veronii* strains heavily and actively may use the system for colonization or other purposes. On the other hand, T2SS components are homologous to Tad pilus and T4P, therefore, they show similarity in their structural components (Peabody et al., 2003; Korotkov et al., 2012). Moreover, also T4P play a role in the motility, adhesion and signaling process (Nivaskumar and Francetic, 2014). According to our results, all the mandatory genes of T4P are carried by the evaluated *A. veronii* genomes.

Since Flagellum is evolutionarily similar to T3SS, they show a similarity in their components and our comparative genomics approach, we identified that all the *A. veronii* genomes encode the mandatory genes of the flagellum. Interestingly, some of the *A.*

veronii genomes do not carry T3SS; strains AMC34, AMC35, AVNIH2, AVNIH1, CECT 4257, CCM 4359, AER397 (human isolates), Ae52 (goldfish isolate) and B565 (pond sediment isolate), LMG 13067 (unknown source), while the rest of the evaluated genomes carry this system (Figure 4). Even though T3SS is considered an important virulence factor, we could hypothesize that these isolates could be using other secretion systems for invasion and competition purposes. Also, previous research reported that *A. veronii* uses T3SS to protect itself from leech immune cells (Silver et al., 2007), which shows that majority of the evaluated genomes could be using the similar strategy to avoid cellular immune system of the host organisms during the encounter of immune cell attacks.

In our analyses, we identified that some of the cattle isolates, one human isolate (strain VBF557) and one surface sentiment (strain pamvotica) encode some of the component of two different types of T4SS type T and F. T4SS is known for helping mobile genetic elements exchange across different organisms, which is observed in our study that strain AVNIH1 and pamvotica possess more insertion elements. T5SS is one of the most widespread and abundant systems in bacterial genomes and divided into five types (Abby et al., 2016), and they have a unique characteristic way of secreting the substrates including toxins and receptor proteins (Green et al., 2016). In our analyses, we identified that all the cattle isolates and strain pamvotica encode mandatory genes of T5aSS and T5bSS and rest of the evaluated genomes either carry only one or none of these mandatory gene of subtypes.

Another important result of the secretions systems analyses is that majority of the genomes encode the T6SSi system, but some of the *A. veronii* genomes lack this system.

This system has been studied extensively and well-known for their ability to interact with environment, host and competitor organisms by using the system as a versatile weapon (Cianfanelli et al., 2016). Interestingly, our *in silico* analysis showed that only one gene (*TssH* also known as *ClpV*) is conserved in all the evaluated genomes regardless of the presence of the other component of this system. Moreover, our results showed that T6SS availability in different organisms and location vary including carp, human, pond sediment, fish intestine and surface water, which may suggest that there is no specific pattern to carry this system based on a location or organism. Finally, we evaluated the Tad system, which is known for an essential colonization, biofilm formation and virulence of some species (Tomich et al., 2007) and results indicated that majority of the 28 out of 41 genomes do not encode this system.

Insertion elements are known as an important factor for evolutionary role in the gene inactivation or changing the genome structure (Touchon et al., 2007). They can undergo extreme expansion by flanking regions, gene inactivation and decay, genome rearrangement and reduction, and also incorporate additional genes they could contribute to pathogenicity (Siguier et al., 2014). In our analysis, interestingly, we observe that U.S. and China catfish potentially clonal isolates and U.S. cattle and Greece surface sentiment isolates, they encode similar types of insertion families, whereas rest of the *A. veronii* isolates show a more scattered pattern of distribution. As we discussed before, this may suggest that two different isolates may have been transferred from different locations and introduced to U.S. Also, we can speculate that carrying the large number insertion elements could be the result of encoding some or all the components of T4SS, because this system is well-known for DNA intake or exchange from the other organisms. In our

analysis, we found that strains pamvotica and AVNIH1 encode either all the accessory genes of T4SS elements or some of them, which could be playing an important role in the intake of insertion elements. Because these two strains encode the largest number of insertion elements.

Other important mobile elements are bacteriophage regions, which are well known for providing specific mechanisms to affected bacteria, attachment, invasion and survival capabilities (Boyd et al., 2012). However, this may end up affecting the genome structure by either increasing or decreasing the pathogenicity of the organism by acquiring virulence factors from the different organisms. According to our results, we could speculate that *A. veronii* genomes somehow tend to have phage elements because except two genomes rest of the genomes carry different levels of bacteriophage. Even though there is a different number of phage elements on U.S. and China catfish isolates; they share only one intact phage region which this could be due to the incomplete genome of strain ML09-123. Meanwhile, we did not identify a conversed phage elements for the 41 *A. veronii* genomes.

CIRSPIR-Cas systems are consist of two different main components, CRISPR array and CRISPR associated genes (Cas), these two systems are separated from each other by spacers (Grissa et al., 2007a). They play a role as an immune system to withstand against viral predation, foreign DNA invasion such as prophages, plasmids or transposons (Horvath et al., 2010). Our results showed that all the evaluated *A. veronii* genomes have either confirmed or questionable CRISPR elements but only strains AMC34, TH0426, LMG 13067, AVNIH1 and AVNIH2 encode the Cas elements along

with CRISPR regions, which may suggest that CRISPR regions could undergo dynamic and rapid evolutionary changes. Results are listed in table 2.

Interestingly, strain AVNIH1 is the only confirmed encoding one large size plasmid, which carries three integron regions. On the other hand, strain Ae52 encodes complete integron region on its genome. This is a very good example of promoting genetic diversity in the *A. veronii* genomes, by vertical inheritance from different organism because many of the *A. veronii* genomes either carry only one, two or none of the CALIN elements in their genome. For the functional integron, only simple core elements (*intl* gene and *attl* integration site) are required (Boucher et al., 2007). Since integrons are known for the major group genetic element for facilitating the spread of antibiotic resistance genes. We can speculate that the evaluated *A. veronii* genomes may have adapted some of the components of integron regions over the time, but they have not completed or acquired the full core components so that they can function properly. For instance, from the same study, 4 of 17 cattle isolates do not carry any integron regions, whereas 13 of them carry only CALIN regions. This is a good example of promoting to genetic material changes over the time. Also, as we will discuss in the upcoming section, the AVNIH1 strain has more antibiotic resistance elements more than other evaluated genomes, this would be the result of encoding three integron regions.

As we discussed previously, AVNIH1 strain encodes most of the antibiotic resistance elements, which could be explained by encoding three integrons elements. Due to heavily usage of antibiotics, we would expect to see different types of antibiotic distribution in evaluated *A. veronii* genomes, but antibiotic resistance database search

showed a majority of the antibiotic resistance elements are shared by all the *A. veronii* genomes with some exceptions.

Additionally, we investigated the disease progression in different doses by doing the fish challenge. To do this, we did fish challenge experiment using one of the epidemic isolate *A. hydrophila* ML09-119 against *A. veronii* ML09-123. Results were quite interesting because even in lower dose we observed mortality in *A. veronii* injection, whereas there was not any mortality at dose 10^4 and 10^5 in *A. hydrophila* injection (Figure-6). This result suggests that *A. veronii* injection shows a typical disease progression, which could suggest that even the little dose may lead to mortality. On the other hand, *A. hydrophila* may a different molecular mechanism to cause the mortality. For instance, to cause mortality, optical density should reach a certain point, so that produced toxin and destructive enzymes reach a certain level and start affecting the organism.

In conclusion, our comparative *A. veronii* genomes analyses shows that another *A. veronii* may have been transferred from China to U.S. Also, this observation was confirmed by ANI calculation and Phylogenetic tree, subsystem category distribution, secretion systems, phage elements, insertion elements, genomics islands, integron, antibiotic resistance elements. They showed the same type of distribution for the U.S. catfish isolate strain ML09-123 and China catfish isolate TH0426. Meanwhile, we see the similar pattern for the newly reported Greece isolates strain pamvotica, against to U.S. cattle isolates. These evidence clearly suggest non-native species may have been introduced to different geographical locations by different activities. Also, we report an extended evaluation of comparative genomics approaches, which will provide novel

insights into the pathogenicity of *A. veronii* genomes from different hosts and geographical locations. We expect that these findings will provide a base platform to investigate in detail the molecular mechanism of *A. veronii* strains for the further studies. Also, this will be the first report to raise an awareness of potential epidemics of *A. veronii*.

5.6 References

Abby, S.S., Cury, J., Guglielmini, J., Neron, B., Touchon, M., and Rocha, E.P. (2016). Identification of protein secretion systems in bacterial genomes. *Sci Rep* 6**,** 23080. doi: 10.1038/srep23080.

Abby, S.S., Neron, B., Menager, H., Touchon, M., and Rocha, E.P. (2014). MacSyFinder: a program to mine genomes for molecular systems with an application to CRISPR-Cas systems. *PLoS One* 9(10)**,** e110726. doi: 10.1371/journal.pone.0110726.

Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., Liang, Y., et al. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 44(W1)**,** W16-21. doi: 10.1093/nar/gkw387.

Ast, V.M., Schoenhofen, I.C., Langen, G.R., Stratilo, C.W., Chamberlain, M.D., and Howard, S.P. (2002). Expression of the ExeAB complex of Aeromonas hydrophila is required for the localization and assembly of the ExeD secretion port multimer. *Mol Microbiol* 44(1)**,** 217-231.

Austin, B., and Austin, A.D. (2012). *Bacterial Fish Pathogens Disease of Farmed and Wild Fish.* Springer.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., et al. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9**,** 75. doi: 10.1186/1471-2164-9-75.

Blom, J., Kreis, J., Spanig, S., Juhre, T., Bertelli, C., Ernst, C., et al. (2016). EDGAR 2.0: an enhanced software platform for comparative gene content analyses. *Nucleic Acids Res* 44(W1)**,** W22-28. doi: 10.1093/nar/gkw255.

Bomar, L., Maltz, M., Colston, S., and Graf, J. (2011). Directed culturing of microorganisms using metatranscriptomics. *MBio* 2(2)**,** e00012-00011. doi: 10.1128/mBio.00012-11.

Bomar, L., Stephens, W.Z., Nelson, M.C., Velle, K., Guillemin, K., and Graf, J. (2013). Draft Genome Sequence of Aeromonas veronii Hm21, a Symbiotic Isolate from the Medicinal Leech Digestive Tract. *Genome Announc* 1(5). doi: 10.1128/genomeA.00800-13.

Boucher, Y., Labbate, M., Koenig, J.E., and Stokes, H.W. (2007). Integrons: mobilizable platforms that promote genetic diversity in bacteria. *Trends Microbiol* 15(7)**,** 301-309. doi: 10.1016/j.tim.2007.05.004.

Boyd, E.F., Carpenter, M.R., and Chowdhury, N. (2012). Mobile effector proteins on phage genomes. *Bacteriophage* 2(3)**,** 139-148. doi: 10.4161/bact.21658.

Cai, S.H., Wu, Z.H., Jian, J.C., Lu, Y.S., and Tang, J.F. (2012). Characterization of pathogenic Aeromonas veronii bv. veronii associated with ulcerative syndrome from chinese longsnout catfish (Leiocassis longirostris Gunther). *Braz J Microbiol* 43(1)**,** 382-388. doi: 10.1590/S1517-838220120001000046.

Chen, C., Zhu, N., Kong, L., Bei, Y., Zheng, T., Ding, X., et al. (2013). First case of soft shell disease in Chinese soft-shelled turtle (Trionyx sinens) associated with Aeromonas sobria–A. veronii complex. *Aquaculture* 406-407**,** 62-67.

Cianciotto, N.P. (2005). Type II secretion: a protein secretion system for all seasons. *Trends Microbiol* 13(12)**,** 581-588. doi: 10.1016/j.tim.2005.09.005.

Cianfanelli, F.R., Monlezun, L., and Coulthurst, S.J. (2016). Aim, Load, Fire: The Type VI Secretion System, a Bacterial Nanoweapon. *Trends Microbiol* 24(1)**,** 51-62. doi: 10.1016/j.tim.2015.10.005.

Collis, C.M., and Hall, R.M. (1995). Expression of antibiotic resistance genes in the integrated cassettes of integrons. *Antimicrob Agents Chemother* 39(1)**,** 155-162.

Cury, J., Jove, T., Touchon, M., Neron, B., and Rocha, E.P. (2016). Identification and analysis of integrons and cassette arrays in bacterial genomes. *Nucleic Acids Res* 44(10)**,** 4539-4550. doi: 10.1093/nar/gkw319.

Darling, A.C., Mau, B., Blattner, F.R., and Perna, N.T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 14(7)**,** 1394-1403. doi: 10.1101/gr.2289704.

De, S.J.S.S., Honein, K., Arulkanthan, A., Ushio, H., and Asakawa, S. (2017). Genome sequencing and annotation of Aeromonas veronii strain Ae52, a multidrug-resistant isolate from septicaemic gold fish (Carassius auratus) in Sri Lanka. *Genom Data* 11**,** 46-48. doi: 10.1016/j.gdata.2016.11.011.

Dolores, J.S., Agarwal, S., Egerer, M., and Satchell, K.J. (2015). Vibrio cholerae MARTX toxin heterologous translocation of beta-lactamase and roles of individual effector domains on cytoskeleton dynamics. *Mol Microbiol* 95(4)**,** 590-604. doi: 10.1111/mmi.12879.

Dong, H.T., Techatanakitarnan, C., Jindakittikul, P., Thaiprayoon, A., Taengphu, S., Charoensapsri, W., et al. (2017). Aeromonas jandaei and Aeromonas veronii caused disease and mortality in Nile tilapia, Oreochromis niloticus (L.). *J Fish Dis*. doi: 10.1111/jfd.12617.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5)**,** 1792-1797. doi: 10.1093/nar/gkh340.

Felsenstein, J. (1989). PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics* 5**,** 164-166. doi: citeulike-article-id:2344765.

Green, E.R., and Mecsas, J. (2016). Bacterial Secretion Systems: An Overview. *Microbiol Spectr* 4(1). doi: 10.1128/microbiolspec.VMBF-0012-2015.

Grissa, I., Vergnaud, G., and Pourcel, C. (2007a). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* 8**,** 172. doi: 10.1186/1471-2105-8-172.

Grissa, I., Vergnaud, G., and Pourcel, C. (2007b). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* 35(Web Server issue)**,** W52-57. doi: 10.1093/nar/gkm360.

Hassan, M.A., Noureldin, E.A., Mahmoud, M.A., and Fita, N.A. (2017). Molecular identification and epizootiology of Aeromonas veronii infection among farmed Oreochromis niloticus in Eastern Province, KSA. *The Egyptian Journal of Aquatic Research* 43(2)**,** 161-167. doi: http://dx.doi.org/10.1016/j.ejar.2017.06.001.

Horvath, P., and Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327(5962)**,** 167-170. doi: 10.1126/science.1179555.

Hossain, M. J., Sun, D., McGarey, D. J., Wrenn, S., Alexander, L. M., Martino, M. E., Xinge, Y., Terhune, J. S., Liles, M. R. (2014). An Asian origin of virulent Aeromonas hydrophila responsible for disease epidemics in United States-farmed catfish. *MBio* 5(3)**,** e00848-00814. doi: 10.1128/mBio.00848-14.

Janda, J.M., and Abbott, S.L. (2010). The genus Aeromonas: taxonomy, pathogenicity, and infection. *Clin Microbiol Rev* 23(1)**,** 35-73. doi: 10.1128/CMR.00039-09.

Jia, B., Raphenya, A.R., Alcock, B., Waglechner, N., Guo, P., Tsang, K.K., et al. (2017). CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res* 45(D1)**,** D566-D573. doi: 10.1093/nar/gkw1004.

Joss, M.J., Koenig, J.E., Labbate, M., Polz, M.F., Gillings, M.R., Stokes, H.W., et al. (2009). ACID: annotation of cassette and integron data. *BMC Bioinformatics* 10**,** 118. doi: 10.1186/1471-2105-10-118.

Kang, Y., Pan, X., Xu, Y., Siddiqui, S.A., Wang, C., Shan, X., et al. (2016). Complete genome sequence of the fish pathogen Aeromonas veronii TH0426 with potential application in biosynthesis of pullulanase and chitinase. *J Biotechnol* 227**,** 81-82. doi: 10.1016/j.jbiotec.2016.04.009.

Kanonenberg, K., Schwarz, C.K., and Schmitt, L. (2013). Type I secretion systems - a story of appendices. *Res Microbiol* 164(6)**,** 596-604. doi: 10.1016/j.resmic.2013.03.011.

Kenzaka, T., Nakahara, M., Higuchi, S., Maeda, K., and Tani, K. (2014). Draft Genome Sequences of Amoeba-Resistant Aeromonas spp. Isolated from Aquatic Environments. *Genome Announc* 2(5). doi: 10.1128/genomeA.01115-14.

Kim, M., Oh, H.S., Park, S.C., and 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(Pt 2)**,** 346-351. doi: 10.1099/ijs.0.059774-0.

Korotkov, K.V., Sandkvist, M., and Hol, W.G. (2012). The type II secretion system: biogenesis, molecular architecture and mechanism. *Nat Rev Microbiol* 10(5)**,** 336-351. doi: 10.1038/nrmicro2762.

Letoffe, S., Delepelaire, P., and Wandersman, C. (1996). Protein secretion in gram-negative bacteria: assembly of the three components of ABC protein-mediated exporters is ordered and promoted by substrate binding. *EMBO J* 15(21)**,** 5804-5811.

Li, Y., Liu, Y., Zhou, Z., Huang, H., Ren, Y., Zhang, Y., et al. (2011). Complete genome sequence of Aeromonas veronii strain B565. *J Bacteriol* 193(13)**,** 3389-3390. doi: 10.1128/JB.00347-11.

Liu, D., Geng, Y., Wang, K., Chen, D., Huang, X.L., Ou, Y., et al. (2016). Aeromonas veronii infection in cultured channel catfish, ictalurus punctatus, in Southwest China. *The Israeli Journal of Aquaculture* 68.

Maltz, M., and Graf, J. (2011). The type II secretion system is essential for erythrocyte lysis and gut colonization by the leech digestive tract symbiont Aeromonas veronii. *Appl Environ Microbiol* 77(2)**,** 597-603. doi: 10.1128/AEM.01621-10.

Maltz, M.A., Bomar, L., Lapierre, P., Morrison, H.G., McClure, E.A., Sogin, M.L., et al. (2014). Metagenomic analysis of the medicinal leech gut microbiota. *Front Microbiol* 5**,** 151. doi: 10.3389/fmicb.2014.00151.

McGarey, D.J., Milanesi, L., Foley, D.P., Reyes, B.J., Frye, L., and Lim, D.V. (1991). The role of motile aeromonads in the fish disease, ulcerative disease syndrome (UDS). *Experientia* 47(5)**,** 441-444.

Nivaskumar, M., and Francetic, O. (2014). Type II secretion system: a magic beanstalk or a protein escalator. *Biochim Biophys Acta* 1843(8)**,** 1568-1577. doi: 10.1016/j.bbamcr.2013.12.020.

Peabody, C.R., Chung, Y.J., Yen, M.R., Vidal-Ingigliardi, D., Pugsley, A.P., and Saier, M.H., Jr. (2003). Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. *Microbiology* 149(Pt 11)**,** 3051-3072. doi: 10.1099/mic.0.26364-0.

Pridgeon, J.W., and Klesius, P.H. (2011). Molecular identification and virulence of three Aeromonas hydrophila isolates cultured from infected channel catfish during a disease outbreak in west Alabama (USA) in 2009. *Dis Aquat Organ* 94(3)**,** 249-253. doi: 10.3354/dao02332.

Siguier, P., Gourbeyre, E., and Chandler, M. (2014). Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol Rev* 38(5)**,** 865-891. doi: 10.1111/1574-6976.12067.

Silver, A.C., Kikuchi, Y., Fadl, A.A., Sha, J., Chopra, A.K., and Graf, J. (2007). Interaction between innate immune cells and a bacterial type III secretion system in mutualistic and pathogenic associations. *Proceedings of the National Academy of Sciences* 104(22)**,** 9481-9486. doi: 10.1073/pnas.0700286104.

Sreedharan, K., Philip, R., and Singh, I.S. (2011). Isolation and characterization of virulent Aeromonas veronii from ascitic fluid of oscar Astronotus ocellatus showing signs of infectious dropsy. *Dis Aquat Organ* 94(1)**,** 29-39. doi: 10.3354/dao02304.

Tekedar, H. C., Waldbieser, G. C., Karsi, A., Liles, M. R., Griffin, M. J., Vamenta, S., Sonstegard, T., Hossain, M., Schroeder, S, Khoo, L., Lawrence, M. L. (2013). Complete Genome Sequence of a Channel Catfish Epidemic Isolate, Aeromonas hydrophila Strain ML09-119. *Genome Announc* 1(5). doi: 10.1128/genomeA.00755-13.

Tomich, M., Planet, P.J., and Figurski, D.H. (2007). The tad locus: postcards from the widespread colonization island. *Nat Rev Microbiol* 5(5)**,** 363-375. doi: 10.1038/nrmicro1636.

Touchon, M., and Rocha, E.P. (2007). Causes of insertion sequences abundance in prokaryotic genomes. *Mol Biol Evol* 24(4)**,** 969-981. doi: 10.1093/molbev/msm014.

Uzun, E., and Ogut, H. (2015). The isolation frequency of bacterial pathogens from sea bass (Dicentrarchus labrax) in the Southeastern Black Sea. *Aquaculture* 437(1)**,** 30-37.

Varani, A.M., Siguier, P., Gourbeyre, E., Charneau, V., and Chandler, M. (2011). ISsaga is an ensemble of web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. *Genome Biol* 12(3)**,** R30. doi: 10.1186/gb-2011-12-3-r30.

Webb, H.E., Bugarel, M., den Bakker, H.C., Nightingale, K.K., Granier, S.A., Scott, H.M., et al. (2016). Carbapenem-Resistant Bacteria Recovered from Faeces of Dairy Cattle in the High Plains Region of the USA. *PLoS One* 11(1)**,** e0147363. doi: 10.1371/journal.pone.0147363.

Zepeda-Velazquez, A.P., Vega-Sanchez, V., Salgado-Miranda, C., and Soriano-Vargas, E. (2015). Histopathological findings in farmed rainbow trout (Oncorhynchus mykiss) naturally infected with 3 different Aeromonas species. *Can J Vet Res* 79(3)**,** 250-254.

Zhang, X., Yang, W., Wu, H., Gong, X., and Li, A. (2014). Multilocus sequence typing revealed a clonal lineage of Aeromonas hydrophila caused motile Aeromonas septicemia outbreaks in pond-cultured cyprinid fish in an epidemic area in central China. *Aquaculture* 432**,** 1-6.

CHAPTER VI

CONSTRUCTION AND EVALUATION OF SIX *AEROMONAS HYDROPHILA* IN-FRAME DELETION MUTANTS

6.1 Abstract

Catfish aquaculture is one of the major aquaculture commodities in the United States. Since 2009, epidemic *Aeromonas hydrophila* (EAh) cases have had a serious impact on the commercial channel industry in the USA. Even though, significant progress has been made to understand molecular mechanisms of *A. hydrophila* infection, effective control of EAh epidemics has not been achieved. Live attenuated vaccines have the potential as an alternative method against EAh epidemics. Here, we present the vaccine potential of novel *A. hydrophila* mutants. In our study, we successfully constructed six *A. hydrophila* mutants by in-frame deletion method, which is followed by virulence testing on channel catfish fingerlings by i.p. injection. Virulence challenges indicated that sialidase (*vAh*Δ*sia*) and enterotoxin (*vAh*Δ*ent*) mutants could be used as vaccine candidate. On the other hand, mutant *vAh*Δ*col* was not attenuated. Our results clearly support the importance of the role of invasins and toxins in the pathogenesis of *A. hydrophila*. Lastly, interestingly, none of the *hfq* and –related mutants were not statistically significant. Deletion of RNA-binding protein (*vAh*Δ*hfq1*) increased the

virulence, whereas RNA chaperone Hfq (*vAhΔhfq2*) mutant and their double mutant (*vAhΔhfq1Δhfq2*) slightly reduced virulence compared to wild-type.

6.2 Introduction

The aquaculture industry is one of the important agricultural sectors and provides almost half of the animal food for human consumption around the world (Bostock et al., 2010). Specifically, the aquaculture industry in United States of America has been well established for the last 40 years. The majority of sales in aquaculture come from the catfish industry (FAO, 2011). Because of the importance of the catfish industry, multistate research institutes and universities have collaborated to address emerging diseases and have been trying to find solutions for current or potential problems. Since 2009, catfish producers lost more than 5.5 million pounds of market size fish because of *Aeromonas hydrophila* epidemics. Previous comparative genomics research revealed that these outbreaks were caused by an emergent Asian origin clonal group of *A. hydrophila* strains (Hossain et al., 2014; Hossain et al., 2013).

It is well-known that *A. hydrophila* produces a large number of extracellular enzymes and toxins that contribute to host inflammation and promote spreading in the tissues. *A. hydrophila* infection is characterized by secreted proteins involved in bacterial invasion and spreading (Grizzle et al., 1993), and exotoxins are important in *A. hydrophila* infection (Ponnusamy et al., 2016). Also, secreted proteins from *Aeromonas* degrade major components in the extracellular matrix on the fish surface. They also decrease viscosity of mucin, contributing to host invasion and facilitating adhesion to host cells (Engibarov et al., 2014). However, there is not sufficient information available on the virulence of invasins and toxins in EAh and their contribution to pathogenesis has

not been studied extensively. To understand the mechanisms of invasin- and toxins-mediated inflammation and pathology in catfish at the molecular level, we chose three genes, sialidase, collagenase and enterotoxin to understand their pathological role in disease progression. Studies on importance of exotoxins in *A. hydrophila* infection (Ponnusamy et al., 2016) and secreted proteins such as sialidase in *Edwardsiella tarda* (Jin et al., 2012), collagenase in *Aeromonas veronii* (Han et al., 2008) indicate that they have potential to be viable vaccine candidates.

The master regulators of gene expression in bacteria are *hfq* and its related genes (Feliciano et al., 2016) . It is also an essential component for the activity and stability of most small RNAs (sRNAs). The capability of *hfq* genes in bacterial pathogenesis was evaluated, and results showed that deletion of *hfq* genes reduced the virulence of several different organisms. For instance, *Borrelia burgdorferi* (Lybecker et al., 2010), *Brucella abortus* (Robertson et al., 1999), *Agrobacterium tumefaciens* (Wilms et al., 2012), *Edwardsiella tarda* (Hu et al., 2014), *Vibrio cholerae* (Ding et al., 2004), *Yersinia pestis* (Geng et al., 2009), *Listeria monocytogenes* (Christiansen et al., 2004; Nielsen et al., 2010) are examples of decreased virulence due to deletion of *hfq* genes. On the other hand, it also proved that deletion of *hfq* genes did not affect the virulence in *Legionella pneumophila* (McNealy et al., 2005) and *Neisseria gonorroeae* (Dietrich et al., 2009). A recent study also showed that the *hfq* gene is responsible for colony morphology, nutrient utilization and oxidative and envelope stress response in *Vibrio alginolyticus* (Deng et al., 2016).

In spite of the widely appreciated magnitude of EAh problems, an effective preventive method is not still available. To date, complete control of EAh epidemics has

not been achieved due to lack of essential information on the pathogenic mechanisms of EAh. Due to the importance of invasion, toxin and *hfq* genes, understanding their roles may allow design of novel control measures for EAh by limiting its ability to disseminate and induce inflammation. In our study, we successfully constructed six *A. hydrophila* mutants by in-frame deletion method, which was followed by virulence test on channel catfish fingerlings by injection method. Our results clearly support the importance of the role of invasins and toxin in the pathogenesis of *A. hydrophila* because virulence challenge indicated that *vAh*Δ*sia* and *vAh*Δ*ent* mutants could be used as vaccine candidates.

6.3 Materials and methods

6.3.1 Bacterial strains and plasmids

A. hydrophila ML09-119 strain was used as a parent strain to construct *A. hydrophila* mutants. *A. hydrophila* strain ML09-119 was cultured in Brain Heart Infusion (BHI) agar and broth (Difco, Sparks,MD) at 37ºC throughout the study. *E. coli* strains were cultured on Luria-Bertani (LB) agar and in broth from Difco and incubated at 37ºC. Antibiotics from Sigma (St. Louis, MO, USA) were added to culture medium when required containing at the following concentrations: ampicillin (Amp: 100 µg/ml), Colistin (Col: 25 µg/mL), Chloramphenicol (Chl: 30 µG/mL), sucrose (5%), and mannitol (0.35%). *E. coli* strain CC118 and strain BW10 λpir (lambda-pir) were used in cloning and conjugal transfer of suicide plasmid (pMEG-375) into *A. hydrophila* respectively. *E. coli* strains were grown using LB broth and agar at 37ºC for 18 h in a shaker incubator at 200 rpm. As a vector, suicide plasmid pMEG-375 was used to

generate and introduce in-frame deletion in wild-type strain *A. hydrophila* ML09-119. Bacterial strains and plasmid used in this study are listed in table 6.1.

Table 6.1 Bacterial strains and plasmid used in this study.

(Abbreviations: Sialidase: sia, Collagenase: col, Enterotoxin: ent)

Strain or plasmid	**Description**	**Source**
A. hydrophila ML09-119	Isolate from a disease outbreak on a commercial catfish farm	(Griffin et al., 2013; Tekedar et al., 2013)
vAhΔsia	*A. hydrophila* ML09-119; Δ*sia*	This study
vAhΔcol	*A. hydrophila* ML09-119; Δ*col*	This study
vAhΔent	*A. hydrophila* ML09-119; Δ*ent*	This study
vAhΔhfq1	*A. hydrophila* ML09-119; Δ*hfq1*	This study
vAhΔhfq2	*A. hydrophila* ML09-119; Δ*hfq2*	This study
vAhΔhfq1Δhfq2	*A. hydrophila* ML09-119; Δ*hfq1*; Δ*hfq2*	This study
E. coli		
CC118λ*pir*	Δ(*ara-leu*); *araD*; Δ*lacX74*; *galE*; *galK*; *phoA20*; *thi-1*; *rpsE*; *rpoB*; *argE*(Am); *recAl*; λ*pir*R6K	(Herrero et al., 1990)
BW19851	*RP4-2 (Km::Tn7, Tc::Mu-1), DuidA3::pir+, recA1, endA1, thi-1, hsdR17, creC510*	(Metcalf et al., 1994)
Plasmid		
pMEG-375	8,142 bp, Amp^r, Cm^r, *lacZ*, R6K *ori*, *mob incP*, *sacR sacB*	(Dozois et al., 2003)
pAhΔ*sia*	10,049 bp, Δ*sia*, pMEG-375	This study
pAhΔ*col*	10,176 bp, Δ*col*, pMEG-375	This study
pAhΔ*ent*	9,947 bp, Δ*ent*, pMEG-375	This study
pAhΔ*hfq1*	10,158 bp, Δ*hfq1*, pMEG-375	This study
pAhΔ*hfq2*	9,978 bp, Δ*hfq2*, pMEG-375	This study

6.3.2 Genomic DNA and plasmid DNA isolation

A. hydrophila genomic DNA was isolated and purified using Promega (Madison, WI, USA) wizard genomic DNA purification kit. The integrity and concentration of the isolated genomic DNA was measured with NanoDrop 1000 (Thermo Scientific) and run on an agarose gel electrophoresis using 0.8 % concentration. The purified genomic DNA was stored at 4 ºC for further studies. For the plasmid isolation, Qiaspin Miniprep Kit (Qiagen, Hilden, Germany) was used, which is followed by quality and concentration check using Nanodrop 1000 and 1% agarose gel electrophoresis.

6.3.3 Preparation of competent cells

To prepare competent cells, the chemical method was chosen. To do so, a single colony of E. coli strain CC118 λpir was inoculated into 10 ml of LB broth and grown overnight. 8 ml of overnight culture was transferred to 100 ml of fresh LB broth. OD at 600nm was checked regularly until it reached to 0.4 -0.5. The culture was placed in an ice cup for 10 min, after which the bacteria were collected by spinning at 3,000 rpm for 8 min at 4 ºC. The collected pellet was washed two times with cold 0.1 M CaCl2. In a second wash, bacteria were incubated on ice for 30 min before centrifugation. The remaining bacterial pellet after the supernatant removal process was resuspended in 5 ml in cold 0.1 M CaCl2 along with 15% sterilized glycerol. Approximately, 200 µl bacterial culture suspension was aliquoted into 1.5 ml microcentrifuge tubes and placed in liquid nitrogen immediately for quick-freezing. Prepared competent cells were stored at -80 ºC until later use.

6.3.4 Selection of potential genes in *A. hydrophila*

We compared the genomes of the epidemic and reference isolates of *A. hydrophila* and identified genes unique and shared with epidemic and reference isolates. According to our selection results, Sialidase is one of the unique genes for the epidemic isolates. Moreover, collagenase and enterotoxin genes are consistently found in epidemic and reference isolates. Even though *hfq* genes play a crucial role in many pathways of the bacteria, their role in *A. hydrophila* has not been studied so far. Therefore, sialidase, collagenase, enterotoxin, *hfq1*, *hfq2* and *hfq1+hfq2* genes were selected for the mutation purposes.

6.3.5 In-frame gene deletion in vitro

All genes sequences plus 2000 bp flanking upstream and downstream regions were downloaded from https://www.ncbi.nlm.nih.gov by searching specifically in the *A. hydrophila* ML09-119 genome. After Target gene nucleotide files were downloaded, nucleotide files were imported to SeqBuilder (Dnastar) software, and after choosing the desired regions, primers for the selected genes were designed with Primer3web version 4.0.0 (http://bioinfo.ut.ee/primer3/). Mutant construction was illustrated in Figure 6.1. In step 1, 4 primers were designed, external forward and reverse primers were designed 1000 bp away, up and down the flanking regions of the selected genes. Moreover, internal reverse and forward primers were designed at the very beginning of the start and stop codon for each selected gene. In order to produce the spliced DNA fragment by in-frame deletion, reverse complement of the internal primer was added to the other internal primer's 5` end, which is followed by overlap extension PCR procedure. Restriction enzyme sites were added to the 5` end for each of the external primers. All the primers and restriction enzymes used in this study are listed in Table 6.2.

Table 6.2 Primers used to generate and verify in-frame deletion of vAh genes.

	Primer ID		Sequence 5-3'	**RE**
Sialidase	AhSiaEF01	A	CGTAAAGCGAATGCAAACAG	*Xba*I
	AhSiaIR01	B	CGTTTTCGCAAAGGCTTTAAC	
	AhSiaIF01	C	GTTAAAGCCTTTGCGAAAACGTCTCCGAGCACTGAAATTAGG	
	AhSiaER01	D	CGCAATCCCAAGACGAATAG	*SmaI*
	SiaA1Seq		GCCTGAACAACCACATAAACC	
Collagenase	AhColEF01	A	ATTCTAGACTCTGCTATTGAGGGGTCAG	*Xba*I
	AhColIR01	B	CAATCACGCAGACAATCAAGCGGACGAGACGACTATGACAAGG	
	AhColIF01	C	GCTTGATTGTCTGCGTGATTG	
	AhColER01	D	ATGCATGGGAGCTGCCATCTGAGAAGAG	*sphI*
	ColA1Seq		GGATCTTGGCAAACTTCTTCA	
Enterotoxin	AhEntEF01	A	GGCCCTCTACCAGCAACATA	*Xba*I
	AhEntIR01	B	CCTCGGGGATTACAGCTTG	
	AhEntIF01	C	CAAGCTGTAATCCCCGAGGTGATCCGGACCAAGTTTGAG	
	AhEntER01	D	CTGCGCTAATCGAACCTACA	*SmaI*
	EntA1Seq		AGCTCAAGATCGCCTTCTCC	
Hfq-1	AHHfqEF01	A	CCCCTCTAGACCTCTTGAACCCATAAACCAG	*Xba*I
	AHHfq1IR01	B	ATCTGGGTTGACGAGTAGCCTCCCCTTAGCCATCTTCTTTTCC	
	AHHfq1IF01	C	AGGCTACTCGTCAACCCAGAT	
	AHHfq1ER01	D	CCCCGCATGCCCAGCTTGTCGATCTTGTTG	*sphI*
	Hfq1A1Seq		TGGAGTATGATGAGATGCGCT	
Hfq-2	AHHfq2EF01	A	CCCCTCTAGAGGCACGTAGAGGATCACAGAG	*xbaI*
	AHHfq2IR01	B	TTTGACGCAGCAATGACATCAAAAAGTGGCGTTTGATTGTG	
	AHHfq2IF01	C	GATGTCATTGCTGCGTCAAA	
	AHHfq2ER01	D	CCCCGAGCTCCCGCTGTTTCATCTGTTCCA	*sacI*
	Hfq2A2Seq		GTGAAACCGGATGCAAGATGA	
Hfq-1 and *Hfq-2*	N/A	N/A	*Hfq-1* and *Hfq-2* primers were used	N/A

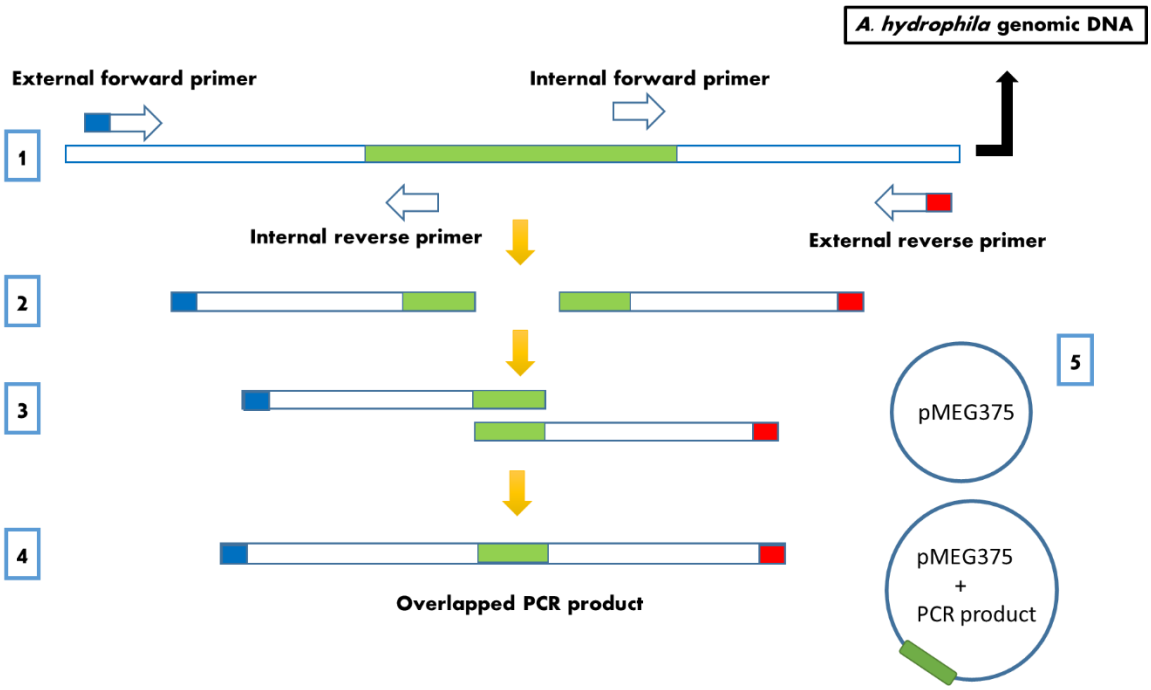


Figure 6.1 *A. hydrophila* in-frame deletion process.

In step 1, arrows stand for primers with the blue and red RE sites for cloning.

In the second step, the total PCR reaction volume was 25 µl, and the reaction was adjusted to 100 ng *A. hydrophila* wild-type genomic DNA (gDNA). Two µl of this prep was used with 0.5 µl of 10 mM internal and external forward and reverse primers, 0.5 µl 10 mM dNTPs, 5 µl 5x buffer, and 0.025 U goTaq DNA polymerase, and the remaining volume was filled with dd H_2O. In the first PCR, external forward-internal reverse primers and internal forward and external reverse primers were used to amplify two flanking regions of target gene. The PCR conditions: denaturation at 95 ºC for 2 min, 30 cycles of 95 ºC for 30 sec, 56 ºC for 1 min, and 72 ºC for 1 min, and extension at 72 ºC for 10 min. In the overlap extension PCR step, PCR products were used from the

previous steps as a template and following PCR conditions were used: denaturation at 94 ºC for 10 min, 30 cycles of 94 ºC for 30 sec, 56 ºC for 1 min, and 72 ºC for 2 min, and extension at 72 ºC for 10 min. After overlap PCR, PCR products were checked to see the correct amplification and cleaned by 1% agarose gel in 0.5 x TBE buffer. External forward and reverse primers were used to amplify the recombinant DNA fragment. In order to produce more recombinant DNA fragment, it was reamplified using the same primer sets.

Using specific restriction enzymes, the recombinant DNA and pMEG-375 plasmid were digested. To verify the cut, the digested DNA and pMEG-375 were checked on 1% agarose gel electrophoresis gel. Uncut plasmid and 1 kb plus ladder (Invitrogen, Thermofisher, Waltham, MA) were used as standards. To ligate the digested-DNA into digested-plasmid (pMEG-375), T4 ligase (Promega) was used, molar ratio was adjusted to vector: insert = 1: 3. Ligation procedure was performed at 16 ºC overnight. For the transformation, 8 µl of the ligation product was used to insert into *E. coli* CC118 λpir, using chemical competent cells. Ligation product (8µl) and competent cells (120 µl) were gently mixed in 1.5 ml microcentrifuge tube and placed on ice for 30-45 minutes. After this step, the mixture was immediately transferred into a water-bath at 42 ºC for 45 seconds, which was followed incubation on ice for another 2 minutes. 500 µl SOC medium was added to the competent cell and ligation product mixture, and incubated for 1 hour at 37 ºC, 200 rpm. Bacteria were spread on 2 LB+amp (100 µg/ml) agar plates and incubated at 37 ºC overnight. Next colony PCR was performed, using a bacterial colony as the genomic DNA template and adding 15 µl water, 0.5 µl primers, 0.5 µl 10 mM dNTPs, 5 µl 5x buffer, and 0.013 U goTaq DNA polymerase to the PCR reaction. The

PCR conditions used: denaturation at 94 ºC for 12 min, 30 cycles of 94 ºC for 30 sec, 56 ºC for 1 min, and 72 ºC for 2 min, and extension at 72 ºC for 10 min, 10 ºC infinite. Positive colonies were selected after the PCR step. Lastly, PCR product with in-frame deletion was ligated into pMEG-375 and generated pAhΔsia, pAhΔcol, pAhΔent, pAhΔhfq1, pAhΔhfq2 (Table 6.1). Inserted DNA in each plasmid was confirmed by enzyme digestion as well as sequencing.

The pMEG-375 plasmids with in-frame target genes were transferred into *E. coli* BW10 by chemical method and were mobilized into *A. hydrophila* ML09-119 by conjugation. The recipient cells were spread on BHI plates containing colistin (25 µg/ml) and chloramphenicol (30 µg/ml) to detect the clone with integrated vector by single crossover through allelic exchange. The colistin and chloramphenicol resistant colonies were propagated on BHI plates to allow for the second crossover allelic exchange. Next, colonies were streaked on BHI plates with 5% sucrose, 0.35% mannitol, and colistin (12.5 µg/ml). These plates selected for loss of pMEG-375 with target genes. Colonies from the selective plates were tested on BHI+ chloramphenicol sensitivity to ensure loss of the plasmid. Growing colonies were selected as a potential mutant and this step is followed by performing one more PCR to confirm the mutants using external primers sets. Finally, verified mutants were sent for sequencing for double verification.

6.3.6 Construction of double mutant

A. hydrophila ML09-119 *vAhΔhfq1Δhfq2* mutant were constructed using deletion mutagenesis. Mutant *vAhΔhfq1* was used as a wild type and *vAhΔhfq2* was introduced to this strain by applying the same procedure as discussed above.

6.3.7 Growth kinetics

Mutants *vAhΔsia*, *vAhΔcol*, *vAhΔent*, *vAhΔhfq1*, *vAhΔhfq2*, *vAhΔhfq1Δhfq2* and wild-type *A. hydrophila* strain ML09-119 were grown in brain heart infusion (BHI) broth at 37°C. To provide aerobic condition, shaking incubator was used at 37°C and 200 rpm. Bacterial growth was measured at 2 h, 4 h, 6 h, 8 h, 10 h, 12 h, 18 h, 24 h, 36 h, and 48 h post-inoculation using a spectrophotometer at OD_{600nm}.

6.3.8 Virulence assays

All the fish experiments conducted in this study followed a protocol approved by the Institutional Animal Care and Use Committee at Mississippi State University. The LD_{50} values of the wild-type strain were determined as 10^5 cfu in catfish. Virulence *vAhΔsia*, *vAhΔcol*, *vAhΔent*, *vAhΔhfq1*, *vAhΔhfq2*, *vAhΔhfq1Δhfq2* was compared to *A. hydrophila* ML09-119 wild-type by injection method. Briefly, twelve-month-old specific-pathogen-free (SPF) channel catfish fingerlings (15.93 ± 0.49 cm, 48.22 ± 3.81 g) were stocked into twenty-four 40-liter flow-through tanks (10 fish/tank) and acclimatized for a week. Tanks were assigned randomly to seven treatment groups: *vAhΔsia*, *vAhΔcol*, *vAhΔent*, *vAhΔhfq1*, *vAhΔhfq2*, *vAhΔhfq1Δhfq2*, vAh WT (positive control for virulence), and BHI (negative control for virulence and sham control for vaccination). Each group included three replicate tanks. Water temperature was maintained at 32ºC (±2) throughout the experiments. Fish were fed twice a day with a commercial catfish feed. Fish mortalities were recorded daily for a total of eight days, and percent mortalities was calculated for each group. At the end of the experiments, percent survival was calculated for *vAhΔsia*, *vAhΔcol*, *vAhΔent*, *vAhΔhfq1*, *vAhΔhfq2*, *vAhΔhfq1Δhfq2* mutants.

6.3.9 Statistical Analysis

For each tank`s and treatment`s mean percent mortalities were calculated and subjected to arcsine transformation. To evaluate virulence of each mutant, the mean transformed percent mortality for the fish challenged with *vAhΔsia*, *vAhΔcol*, *vAhΔent*, *vAhΔhfq1*, *vAhΔhfq2* and *vAhΔhfq1Δhfq2*, mutant strains were compared with *A. hydrophila* WT group by Student's t-test. Differences between the growth of *A. hydrophila* mutants and *A. hydrophila* ML09-119 was conducted using Student's t-test ($p < 0.05$).

6.4 Results

6.4.1 Construction of the *A. hydrophila* in-frame deletion mutants

The five single mutant and one double mutant constructed in this study included *vAhΔsia*, *vAhΔcol*, *vAhΔent*, *vAhΔhfq1*, *vAhΔhfq2*, *vAhΔhfq1Δhfq2*. Briefly, flanking regions were amplified, and in-frame deleted overlap extension fragments were obtained. All the DNA fragments were successfully cloned into pMEG-375 suicide plasmid, which is followed by conjugation step and mutants were confirmed with PCR and sequencing. *A. hydrophila* deletion mutants were obtained and listed in detail in table 6.3.

Table 6.3 The sizes of upstream (USF), downstream (DSF), and in-frame fused fragments (FF), deleted region (DR), and undeleted region (UD) by base pair (bp).

Gene name	**Locus tag (new/old)**	**USF (AB)**	**DSF (CD)**	**FF (AD)**	**DR**	**UD**
Sialidase (neurominadase)	AHML_RS2299 AHML_07175	934	973	1907	1767	33
Collagenase	AHML_RS02695 AHML_02655	1155	879	2034	2760	0
Enterotoxin	AHML_RS04140 AHML_04100	933	872	1805	2021	0
RNA-binding protein Hfq (*hfq1*)	AHML_RS04735 AHML_04675	1039	977	2016	264	12
RNA chaperone Hfq (*hfq2*)	AHML_RS20255 AHML_20165	947	889	1836	408	0
RNA-binding protein Hfq + RNA chaperone Hfq (double mutant)	AHML_RS04735 AHML_04675 AHML_RS20255 AHML_20165	N/A	N/A	N/A	672	12

6.4.2 Growth kinetics

Growth kinetic analysis revealed that there was no difference in the growth of *vAhΔsia*, *vAhΔcol*, *vAhΔent*, *vAhΔhfq1*, *vAhΔhfq2*, mutants and wild-type *A. hydrophila* and listed in Figure 6.2 A and B. However, double mutant, *vAhΔhfq1Δhfq2* showed a slightly lower growth starting from 10 h until 48 h (Figure 6.2)

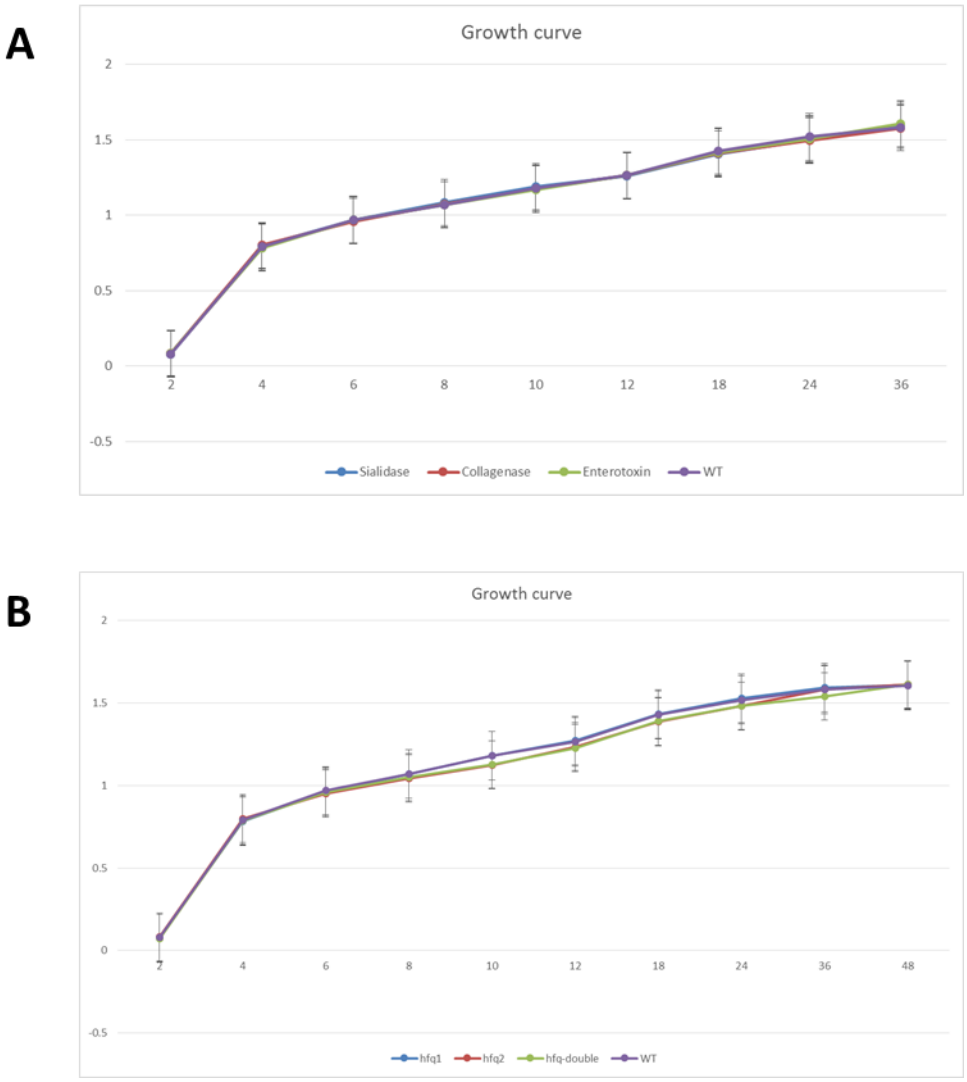


Figure 6.2 Growth kinetics in wild-type *A. hydrophila* ML09-119 and mutants.

(A) *vAhΔsia*, *vAhΔcol*, *vAhΔent*, and wild-type *A. hydrophila* ML09-119. (B) *vAhΔhfq1*, *vAhΔhfq2*, *vAhΔhfq1Δhfq2*, and wild-type *A. hydrophila* ML09-119.

6.4.3 Virulence and efficacy of *A. hydrophila* ML09-119 in-frame deletion mutants

Results of the injection challenge in catfish indicated that the mortality rate was lower in 4 mutants *vAhΔsia*, *vAhΔent*, *vAhΔhfq2*, *vAhΔhfq1Δhfq2* but only *vAhΔsia* and *vAhΔent* mutants were statistically significant, whereas two mutants *vAhΔcol*, *vAhΔhfq1* showed increased virulence and results were listed in Figure 6.3.

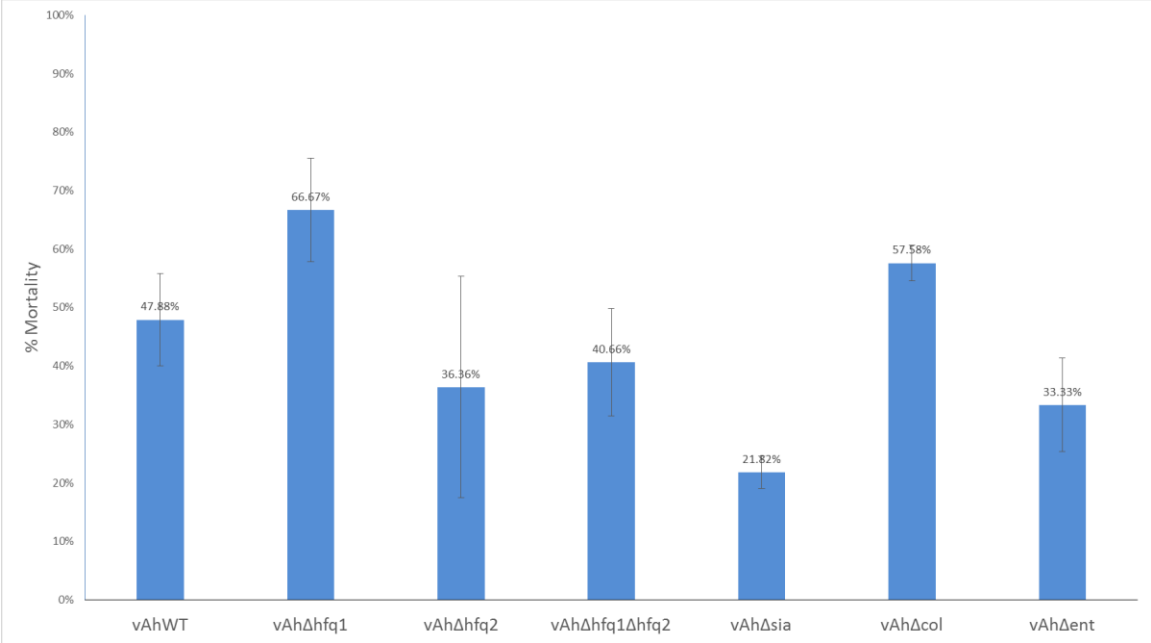


Figure 6.3 Percent mortalities of channel catfish fingerlings challenged with the 5 single mutants (*vAhΔsia*, *vAhΔcol*, *vAhΔent*, *vAhΔhfq*, *vAhΔhfq2*) and 1 double mutant (*vAhΔhfq1Δhfq2*) in-frame deletion mutants and wild type.

6.5 Discussion

The primary objective of this study was to construct live attenuated *A. hydrophila* strains based on mutation of genes responsible for invasion (sialidase, collagenase), toxin (enterotoxin) and *hfq* mutants. Further, we aimed to understand the role of each mutant's

attenuation in catfish. In this study, we successfully deleted all or a large portion of a target gene (more than 90% of each gene was deleted) using splicing overlap extension combined with allelic exchange method. For the first time, we report that these mutants were successfully constructed only in *A. hydrophila*.

The result of virulence of *A. hydrophila* in-frame deletion mutants in catfish trial showed that mutation of Sialidase (*vAh*Δ*sia*) and Enterotoxin (*vAh*Δ*ent*) are statistically significant and they are important in the pathogenesis of *A. hydrophila*. Sialic acid-based pattern recognition contributes to microbial virulence, which strongly supports sialidase inhibition for dampening inflammation caused by infection (Chen et al., 2011). It has been showed that Sialidase plays an important role in the pathogenesis of *E. tarda*, it can be used as a potential vaccine candidate (Chigwechokha et al., 2015; Jin et al., 2012). Our results clearly from fish trials has shown that Enterotoxin is also an important gene for the virulence. Exotoxins are also important virulence factors produced by *A. hydrophila*; deletion of enterotoxin in a human isolate caused attenuation, and it was suggested that it could be used as a potential vaccine candidate (Sha et al., 2002). Moreover, it was shown that cytotoxic enterotoxin in *A. hydrophila* induces proinflammatory cytokine production as well as activates arachidonic acid metabolism in macrophages (Chopra et al., 2000). It is also responsible for activating mitogen-activated protein kinases and induction of apoptosis in murine macrophages and human intestinal epithelial cells (Galindo et al., 2004)

In our study, mutation of collagenase (*vAh*Δ*col*) gene was not attenuated the virulence. *Aeromonas* extracellular enzymes play a central role in stimulating host inflammation, and they promote spreading in the tissues. For example, collagenolytic

activity in *Aeromonas* strains (Duarte et al., 2015; Han et al., 2008). Another study in *Aeromonas veronii* suggested collagenase could play a role in disease progression with the help of adherence and colonization capability in the host cells (Han et al., 2008). Collagenase genes were shown as a critical virulence factor for *Leptospira* species in terms of its invasive and transmission capability (Kassegne et al., 2014). In *Setaria cervi*, collagenase could be used to form an effective molecular vaccine against human lymphatic filariasis (Pokharel et al., 2006).

Next, we determined whether the deletion of RNA-binding protein Hfq (*vAhΔhfq1*), RNA chaperone (*vAhΔhfq2*) and their double mutant (*vAhΔhfq1Δhfq2*) had decreased or increased virulence. None of the mutants` virulence challenge results were statistically significant. Even though mutant RNA-binding protein Hfq (*vAhΔhfq1*) increased the virulence (66.6 %), whereas RNA chaperone (*vAhΔhfq2*) and their double mutant (*vAhΔhfq1Δhfq2*) reduced the virulence compared to wild-type *A. hydrophila*, 36.36 %, 40.66 %, 47.88 %, respectively (Figure 6.3). *Hfq* gene is well-known as a master regulator of gene expression in bacteria (Feliciano et al., 2016) and is also an essential component for activity and stability of most small RNA`s (sRNAs). Their capability in bacterial pathogenesis in several different organisms was due to deletion of *hfq* genes which reduced their virulence. Even though, our mutants did not show statistically significant reduction in the virulence, but slight decrease still correlates with several previous studies. Importantly, *A. hydrophila* is known to produce large numbers of destructive enzymes and toxins. Due to damage to the host organism, we were not able to observe any drastic changes. Hence, it can be said that double mutants of *hfq* genes along

with destructive enzymes or toxins could give a better understanding of the role of this gene.

Despite the presence of several extracellular enzymes and toxins in *A. hydrophila*, we know little about their vaccine potential because they have not been studied systematically. In conclusion, results from the current study confirmed that invasions and toxins have potential to be vaccine candidates. Deletion of sialidase and enterotoxin results are promising. The current study provides a better understanding of the molecular pathogenesis of *A. hydrophila* and specifically, determines the role of *vAhΔsia*, *vAhΔcol*, *vAhΔent*, *vAhΔhfq1*, *vAhΔhfq2*, *vAhΔhfq1Δhfq2* mutants in vAh pathogenesis.

6.6 References

Bostock, J., McAndrew, B., Richards, R., Jauncey, K., Telfer, T., Lorenzen, K., Little, D., Ross, L., Handisyde, N., Gatward, I., Corner, R. (2010). Aquaculture: global status and trends. *Philos Trans R Soc Lond B Biol Sci, 365*(1554), 2897-2912. doi:10.1098/rstb.2010.0170

Chen, G. Y., Chen, X., King, S., Cavassani, K. A., Cheng, J., Zheng, X., Cao, H., Yu, H., Qu, J., Fang, D., Wu, W., Bai, X. F., Liu, J. Q., Woodiga, S. A., Chen, C., Sun, L., Hogaboam, C. M., Kunkel, S. L., Zheng, P., Liu, Y. (2011). Amelioration of sepsis by inhibiting sialidase-mediated disruption of the CD24-SiglecG interaction. *Nat Biotechnol, 29*(5), 428-435. doi:10.1038/nbt.1846

Chigwechokha, P. K., Tabata, M., Shinyoshi, S., Oishi, K., Araki, K., Komatsu, M., Itakura, T., Shiozaki, K. (2015). Recombinant sialidase NanA (rNanA) cleaves alpha2-3 linked sialic acid of host cell surface N-linked glycoprotein to promote Edwardsiella tarda infection. *Fish Shellfish Immunol, 47*(1), 34-45. doi:10.1016/j.fsi.2015.08.015

Chopra, A. K., Xu, X., Ribardo, D., Gonzalez, M., Kuhl, K., Peterson, J. W., & Houston, C. W. (2000). The cytotoxic enterotoxin of Aeromonas hydrophila induces proinflammatory cytokine production and activates arachidonic acid metabolism in macrophages. *Infect Immun, 68*(5), 2808-2818.

Christiansen, J. K., Larsen, M. H., Ingmer, H., Sogaard-Andersen, L., & Kallipolitis, B. H. (2004). The RNA-binding protein Hfq of Listeria monocytogenes: role in stress tolerance and virulence. *J Bacteriol, 186*(11), 3355-3362. doi:10.1128/JB.186.11.3355-3362.2004

Deng, Y., Chen, C., Zhao, Z., Zhao, J., Jacq, A., Huang, X., & Yang, Y. (2016). The RNA Chaperone Hfq Is Involved in Colony Morphology, Nutrient Utilization and Oxidative and Envelope Stress Response in Vibrio alginolyticus. *PLoS One, 11*(9), e0163689. doi:10.1371/journal.pone.0163689

Dietrich, M., Munke, R., Gottschald, M., Ziska, E., Boettcher, J. P., Mollenkopf, H., & Friedrich, A. (2009). The effect of hfq on global gene expression and virulence in Neisseria gonorrhoeae. *FEBS J, 276*(19), 5507-5520. doi:10.1111/j.1742-4658.2009.07234.x

Ding, Y., Davis, B. M., & Waldor, M. K. (2004). Hfq is essential for Vibrio cholerae virulence and downregulates sigma expression. *Mol Microbiol, 53*(1), 345-354. doi:10.1111/j.1365-2958.2004.04142.x

Dozois, C. M., Daigle, F., & Curtiss, R., 3rd. (2003). Identification of pathogen-specific and conserved genes expressed in vivo by an avian pathogenic *Escherichia coli* strain. *Proc. Natl. Acad. Sci. U. S. A., 100*(1), 247-252. doi:10.1073/pnas.232686799

Duarte, A. S., Cavaleiro, E., Pereira, C., Merino, S., Esteves, A. C., Duarte, E. P., Tomás, J. M., Correia, A. C. (2015). Aeromonas piscicola AH-3 expresses an extracellular collagenase with cytotoxic properties. *Lett Appl Microbiol, 60*(3), 288-297. doi:10.1111/lam.12373

Engibarov, S., Eneva, R., & Abrashev, I. (2014). Neuraminidase (sialidase) from Aeromonas sp. strain 40/02 – isolation and partial purification. *Ann. Microbiol, 65*(3), 1515-1523.

FAO. (2011). National Aquaculture Sector Overview. Retrieved from http://www.fao.org/fishery/countrysector/naso_usa/en

Feliciano, J. R., Grilo, A. M., Guerreiro, S. I., Sousa, S. A., & Leitao, J. H. (2016). Hfq: a multifaceted RNA chaperone involved in virulence. *Future Microbiol, 11*(1), 137-151. doi:10.2217/fmb.15.128

Galindo, C. L., Fadl, A. A., Sha, J., Gutierrez, C., Jr., Popov, V. L., Boldogh, I., Aggarwal, B. B., Chopra, A. K. (2004). Aeromonas hydrophila cytotoxic enterotoxin activates mitogen-activated protein kinases and induces apoptosis in murine macrophages and human intestinal epithelial cells. *J Biol Chem, 279*(36), 37597-37612. doi:10.1074/jbc.M404641200

Geng, J., Song, Y., Yang, L., Feng, Y., Qiu, Y., Li, G., Guo, J.,Bi, Y.,Qu, Y., Wang, W., Wang, X.,Guo, Z., Yang, R., Han, Y. (2009). Involvement of the post-transcriptional regulator Hfq in Yersinia pestis virulence. *PLoS One, 4*(7), e6213. doi:10.1371/journal.pone.0006213

Griffin, M. J., Goodwin, A. E., Merry, G. E., Liles, M. R., Williams, M. A., Ware, C., & Waldbieser, G. C. (2013). Rapid quantitative detection of Aeromonas hydrophila strains associated with disease outbreaks in catfish aquaculture. *J Vet Diagn Invest, 25*(4), 473-481. doi:10.1177/1040638713494210

Grizzle, J. M., & Kiryu, Y. (1993). Histopathology of gill, liver, and pancreas, and serum enzyme levels of channel catfish infected with Aeromonas hydrophila complex. *Journal of Aquatic Animal Health, 5*, 36-50.

Han, H. J., Taki, T., Kondo, H., Hirono, I., & Aoki, T. (2008). Pathogenic potential of a collagenase gene from Aeromonas veronii. *Can J Microbiol, 54*(1), 1-10. doi:10.1139/w07-109

Herrero, M., de Lorenzo, V., & Timmis, K. N. (1990). Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria. *J. Bacteriol., 172*(11), 6557-6567.

Hossain, M. J., Sun, D., McGarey, D. J., Wrenn, S., Alexander, L. M., Martino, M. E., Xinge, Y., Terhune, J. S., Liles, M. R. (2014). An Asian origin of virulent Aeromonas hydrophila responsible for disease epidemics in United States-farmed catfish. *MBio, 5*(3), e00848-00814. doi:10.1128/mBio.00848-14

Hossain, M. J., Waldbieser, G. C., Sun, D., Capps, N. K., Hemstreet, W. B., Carlisle, K., Griffin, M. J., Khoo, L., Goodwin, A. E., Sonstegard, T. S., Schroeder, S., Hayden, K., Newton, J. C., Terhune, J. S., Liles, M. R. (2013). Implication of lateral genetic transfer in the emergence of Aeromonas hydrophila isolates of epidemic outbreaks in channel catfish. *PLoS One, 8*(11), e80943. doi:10.1371/journal.pone.0080943

Hu, Y. H., Li, Y. X., & Sun, L. (2014). Edwardsiella tarda Hfq: impact on host infection and global protein expression. *Vet Res, 45*, 23. doi:10.1186/1297-9716-45-23

Jin, R. P., Hu, Y. H., Sun, B. G., Zhang, X. H., & Sun, L. (2012). Edwardsiella tarda sialidase: pathogenicity involvement and vaccine potential. *Fish Shellfish Immunol, 33*(3), 514-521. doi:10.1016/j.fsi.2012.06.002

Kassegne, K., Hu, W., Ojcius, D. M., Sun, D., Ge, Y., Zhao, J., Yang, X. F., Li, L., Yan, J. (2014). Identification of collagenase as a critical virulence factor for invasiveness and transmission of pathogenic Leptospira species. *J Infect Dis, 209*(7), 1105-1115. doi:10.1093/infdis/jit659

Lybecker, M. C., Abel, C. A., Feig, A. L., & Samuels, D. S. (2010). Identification and function of the RNA chaperone Hfq in the Lyme disease spirochete Borrelia burgdorferi. *Mol Microbiol, 78*(3), 622-635. doi:10.1111/j.1365-2958.2010.07374.x

McNealy, T. L., Forsbach-Birk, V., Shi, C., & Marre, R. (2005). The Hfq homolog in Legionella pneumophila demonstrates regulation by LetA and RpoS and interacts with the global regulator CsrA. *J Bacteriol, 187*(4), 1527-1532. doi:10.1128/JB.187.4.1527-1532.2005

Metcalf, W. W., Jiang, W., & Wanner, B. L. (1994). Use of the rep technique for allele replacement to construct new Escherichia coli hosts for maintenance of R6K gamma origin plasmids at different copy numbers. *Gene, 138*(1-2), 1-7.

Nielsen, J. S., Lei, L. K., Ebersbach, T., Olsen, A. S., Klitgaard, J. K., Valentin-Hansen, P., & Kallipolitis, B. H. (2010). Defining a role for Hfq in Gram-positive bacteria: evidence for Hfq-dependent antisense regulation in Listeria monocytogenes. *Nucleic Acids Res, 38*(3), 907-919. doi:10.1093/nar/gkp1081

Pokharel, D. R., Rai, R., Nandakumar Kodumudi, K., Reddy, M. V., & Rathaur, S. (2006). Vaccination with Setaria cervi 175 kDa collagenase induces high level of protection against Brugia malayi infection in jirds. *Vaccine, 24*(37-39), 6208-6215. doi:10.1016/j.vaccine.2006.05.103

Ponnusamy, D., Kozlova, E. V., Sha, J., Erova, T. E., Azar, S. R., Fitts, E. C., Kirtley, M.L., Tiner, B.L., Andersson, J.A., Grim, C.J., Isom, R.P., Hasan, N.A., Colwell, R.R., Chopra, A. K. (2016). Cross-talk among flesh-eating Aeromonas hydrophila strains in mixed infection leading to necrotizing fasciitis. *Proc Natl Acad Sci U S A, 113*(3), 722-727. doi:10.1073/pnas.1523817113

Robertson, G. T., & Roop, R. M., Jr. (1999). The Brucella abortus host factor I (HF-I) protein contributes to stress resistance during stationary phase and is a major determinant of virulence in mice. *Mol Microbiol, 34*(4), 690-700.

Sha, J., Kozlova, E. V., & Chopra, A. K. (2002). Role of various enterotoxins in Aeromonas hydrophila-induced gastroenteritis: generation of enterotoxin gene-deficient mutants and evaluation of their enterotoxic activity. *Infect Immun, 70*(4), 1924-1935.

Tekedar, H. C., Waldbieser, G. C., Karsi, A., Liles, M. R., Griffin, M. J., Vamenta, S., Sonstegard, T., Hossain, M., Schroeder, S, Khoo, L., Lawrence, M. L. (2013). Complete Genome Sequence of a Channel Catfish Epidemic Isolate, Aeromonas hydrophila Strain ML09-119. *Genome Announc, 1*(5). doi:10.1128/genomeA.00755-13

Wilms, I., Moller, P., Stock, A. M., Gurski, R., Lai, E. M., & Narberhaus, F. (2012). Hfq influences multiple transport systems and virulence in the plant pathogen Agrobacterium tumefaciens. *J Bacteriol, 194*(19), 5209-5217. doi:10.1128/JB.00510-1

CHAPTER VII

CONCLUSION

7.1 Summary of results

The *overall objectives* of the study were to sequence and annotate selected *E. piscicida*, *A. hydrophila*, and *A. veronii* genomes and analyze them using several different comparative genomics approaches. Then, we extended our study to chosen invasins, toxins, master regulators, and T6SS effector genes by mutating them with an in-frame deletion method. The present study reports the first comprehensive comparative genomics study of three fish pathogens, which will enhance our understanding of molecular mechanisms of these pathogens. These findings may also allow the design of novel control and/or treatment strategies.

In the second chapter, we conducted genome sequencing of two different bacterial pathogens (*E. piscicida* and *A. hydrophila*) affecting U.S. aquaculture to accelerate research on their pathogenic mechanisms. We evaluated the relatedness of all the available *Edwardsiella tarda* and *piscicida* genomes in NCBI by applying ANI. Interestingly, the evaluated *E. piscicida* strains share high sequence identity, yet they are from diverse host species and geographic regions. To accelerate pathogenesis research on vAh, we conducted genome sequencing of 10 *A. hydrophila* strains (6 vAh strains and 4

A. hydrophila strains from other fish species). ANI calculation results showed that vAh isolates share very high sequence identity, while the other *A. hydrophila* genomes are more distantly related to this clonal group. BRIG analysis of strain ML09-119 compared to the other *A. hydrophila* strains showed that vAh strains have several unique genomic regions, suggesting that mobilome elements contributed to the virulent phenotype of vAh. Overall, the complete and draft genomes of *E. piscicida* and *A. hydrophila* will provide valuable information to our current knowledge of virulence mechanisms of these two pathogens.

In the third chapter, after determining the relatedness of *E. piscicida* genomes against all the available *E. piscicida* genomes in NCBI, we extended our study to compare *E. piscicida* genomes and *E. ictaluri* genomes to other available *Edwardsiella* genus members. According to ANI calculation and core genome phylogenetic tree analyses, we found that the five *Edwardsiella* species separated from each other. Orthology clusters analysis indicated that *E. ictaluri* and *E. piscicida* genomes have the most shared clusters. In this study, we report for the first time that the *E. ictaluri* 93-146 genome encodes a Type IV secretion system (T4SS). Therefore, we speculate that the T4SS could contribute to the increased number of mobilome elements in *E. ictaluri* genomes compared to *E. piscicida* genomes. Moreover, we showed that the *E. piscicida* C07-087 genome encodes two different Type VI secretion systems (T6SS). Overall, this comparative genomics evaluation provides valuable information about unique and shared features of two important pathogens in the *Edwardsiella* genus. Also, it warrants further study to elucidate the biology and virulence of these important fish pathogens.

The fourth chapter provided comprehensive analysis of *A. hydrophila* secretion systems. As pointed out in chapter IV, vAh strains from the southeastern U.S. and China belong to the same clonal group. In our analysis, we showed that this clonal group`s secretion system distribution is different from other *A. hydrophila* genomes. Intriguingly, tight adherence (Tad) system is present consistently in all the vAh strains. It is possible that the Tad system is one of the vAh-specific adaptations that make this clade of *A. hydrophila* more virulent. Moreover, we found that the majority of the U.S. isolates do not possess a complete type VI secretion system (T6SS), but three core elements (TssD, TssH, and TssI) are encoded. We mutated two of the core elements of T6SS from vAh isolate ML09-119. Our results showed that deletion of the *hcpA1* (*tssD*), *hcpA2* (*tssD*), and *vgrG* (*tssI*) genes reduced vAh virulence 1.80-2.24 fold in catfish fingerlings compared to the parent strain ML09-119. Overall, comparative genomics revealed distribution of various secretion systems in *A. hydrophila* genomes and provided functional information on the role of T6SS components in vAh. We expect that these findings will improve our understanding of *A. hydrophila* secretion systems and their roles in *A. hydrophila* virulence.

In the fifth chapter, we sequenced the genome of virulent *A. veronii* strain ML09-123 from catfish. Our results showed that the U.S. catfish *A. veronii* strain ML09-123 is highly similar to a recent Chinese isolate (*A. veronii* strain TH0426). Using all 41 *A. veronii* genomes available in the National Center for Biotechnology Information (NCBI), we observed that *A. veronii* strains from different locations and sources are sometimes highly similar. The majority of *A. veronii* genomes show similar distribution of subsystems, including type I secretion system (T1SS), type II secretion system (T2SS),

type 4 pilus (T4P), and flagellum core elements, whereas T3SS, T5SS, T6SS, and Tad adherence show different distributions. Mobile elements have different distributions depending on the location and organisms. Overall, this study provides novel insights into genomic variation within the *A. veronii* species and its pathogenicity.

In the sixth chapter, mechanisms of vAh secreted proteins and regulatory proteins in catfish virulence were determined. We successfully constructed three *A. hydrophila* mutants by an in-frame deletion method, which was followed by virulence testing in channel catfish fingerlings by intraperitoneal injection. Sialidase (*vAh*Δ*sia*) and enterotoxin (*vAh*Δ*ent*) mutants had significantly reduced virulence in catfish, whereas collagenase mutant *vAh*Δ*col* did not have decreased virulence. The role of *hfq* genes in catfish virulence was also evaluated, and none of the *hfq* mutants had significantly reduced the virulence. This included a mutant defective in RNA-binding protein (*vAh*Δ*hfq1*) and a mutant defective in RNA chaperone (*vAh*Δ*hfq2*), as well as their double mutant (*vAh*Δ*hfq1*Δ*hfq2*).

The overall results of this study will contribute to our knowledge of virulence and environmental adaptations of these important fish pathogens. Our findings are also useful for comparison between fish- and other-pathogenic *A. hydrophila, A. veronii,* and *E. piscicida* strains. Our results support the importance of the role of sialidase and enterotoxin in the pathogenesis of *A. hydrophila*. Our findings will potentially enable new control measures to prevent these diseases impacting U.S. aquaculture.