ABSTRACT

Title of Document: POLYPHASIC ANALYSES ON THE NATURAL ECOLOGY OF HUMAN PATHOGENS.

Bradd Joseph Haley, PhD, 2012

Directed By: Prof. Rita R. Colwell, UMIACS

The focus of research concerning human pathogens has been primarily centered on virulence in the host, transmission between hosts, and treatment of the subsequent infections. Justifiably, our well-being relies on such research but it has erroneously resulted in the assumption that the role of these microbial pathogens is to infect and reproduce within or on our bodies and then pass to another human to follow this same cycle *ad infinitum*. Although this does represent a true optional lifestyle for many pathogens, it must be stated that this lifestyle is one of several life histories that a pathogen may follow and in many cases human infections represent a dead end. This study focuses on the natural ecology of several human pathogens, *V. cholerae*, *V. parahaemolyticus*, and *V. metecus*, and their associated virulence factors, in regions where they cause sporadic illness as well as a region where one of these pathogens, *V. cholerae*, has never caused a human illness. In this work we demonstrate the non-human environment as a natural ecosystem for several human pathogens as well as a reservoir of virulence factors. This was achieved by employing a combination of

high-throughput whole genome analyses focused on the nucleotide and amino acid level, combined with broader ecological studies evaluating the role of the environment with respect to presence of the pathogens and expression of their virulence factors. This work further demonstrates the ubiquity of virulence factors in the environment and the expression of these factors at temperatures found outside of the human host suggests their utility in the environment.

# POLYPHASIC ANALYSES ON THE NATURAL ECOLOGY OF HUMAN PATHOGENS.

By

Bradd Joseph Haley.

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2012

Advisory Committee:
Professor Rita R. Colwell, Chair
Professor Anwar Huq
Professor James Kaper
Dr. Ivor Knight
Assistant Professor Amy Sapkota
Professor Kevin McIver, Dean's Representative

# Dedication

I dedicate this work to my family.

# Acknowledgements

I would like to gratefully acknowledge my advisors at the University of Maryland Drs. Anwar Huq and Rita R. Colwell, and at the University of Iceland, Dr. Eva Benediktsdóttir for their generous assistance, as well as my laboratory partners Elisa Taviani, Christopher J. Grim, Guillaume Constantin de Magny, Nur A. Hasan, Seon Yonug Choi, Jinna Choi, Arlene Chen, Wessam Mahmoud Elnemr, Phillip Clark, Elizabeth Sancomb for all of their help. Much of this work would not have been possible without the help of my international collaborators, Balakrish G. Nair, Jongisk Chun, Shair Gurbanov, Rashid Akhemedov, Muxtar Rajabov, Sevinj Akhemova, Ofeliya Kadirova, Tukaz Khajiyeva, Tamuna Kokashvili, Nino Janelidze, Ana Tskshvediani, Nino Mitaishvili, Nino Lashkhi, Marina Tediashvili, Celia Municio Diaz, Muhammad A. Islam, Greta Caburlotto, Monica Stauder, as well as collaborators in the United States, Tiffiani J. Onifade, Hediye N. Cinar, Ben D. Tall, Laurenda Carter, Surasri N. Sahu, Mahendra S. Kothary, Ron Baker, Richard Hutchinson. The generous assistance of the Maryland Port Administration, US Maritime Administration, MERC Testing Team (Janet Barnes, Tim Mullulady, George Smith, Derrick Sparks, and Mario Tamburri), Maryland Department of Natural Resources (Rusty Mckay and Sarah Harvey), and Crew of M/V Cape Washington is gratefully acknowledged for assisting in collecting water samples for the study. I would especially like to thank Norma Brinkley and Victoria Lord for their tireless work to make the laboratory function.

Funding for these studies was provided by the Office of the Chief Scientist (USA), and NIAID Microbial Sequencing Centers (grant no. N01-AI-30001 and N01-

AI-40001), the KOSEF National Research Laboratory Program (grant no.R0A-2005-000-10110-0), National Institutes of Health (grant nos. 1RO1A139129-01 and 2RO1A1039129-11A2), National Oceanic and Atmospheric Administration, Oceans and Human Health Initiative (grant no. S0660009), the Biological Threat Reduction Program of the U.S. Defense Threat Reduction Agency (DTRA) though Bechtel National Inc., (grant no. 24914416HC4W00000006), the National Science Foundation (grant no. 0813066), the Bio-Industry Initiative (BII) (grant no. 8008) of the U.S. Department of State and implemented by CRDF, the National Institutes of Health-Fogarty International Center Challenge Grant (no. 1RC1TW008587-01). Student support (BJH) was provided by the Institute of International Education, Fulbright Fellowship program, the Biological Threat Reduction Program of the U.S. Defense Threat Reduction Agency (DTRA) though Bechtel National Inc., (grant no. 24914416HC4W00000006), and the Bio-Industry Initiative (BII) (grant no. 8008) of the U.S. Department of State and implemented by CRDF.

# Table of Contents


Dedication ..... ii
Acknowledgements ..... iv
Table of Contents ..... vi
Chapter 1: Introduction ..... 1
Ecological Roles of Microbial Virulence ..... 2
Objectives of this Study ..... 6
Chapter 2: Pre-7th Pandemic *Vibrio cholerae* BX 330286 El Tor Genome; Evidence for the Environment as a Genome Reservoir ..... 11
Abstract ..... 11
Introduction ..... 11
Materials and Methods ..... 14
Genome Sequencing ..... 14
Comparative Genomics ..... 15
Phylogeny of Mobile Genetic Elements ..... 16
Results ..... 17
General Genome Features, Virulence Factors, and Gen2omic Islands ..... 17
Conservation of the Superintegron Region among Clinical Strains ..... 19
Genome-Wide Nucleotide Similarity Shows High Similarity to Pre-7th Pandemic and 7th Pandemic Strains ..... 20
Operon Divergence in 7th Pandemic Clade ..... 22
Conclusions ..... 23
Reproduction License Agreement ..... 30
Chapter 3: Multiomic Analysis of US Gulf Coast Cholera (2010 to 2011) Isolates ..... 37
Abstract ..... 37
Introduction ..... 37
Materials and Methods ..... 39
Results and Discussion ..... 40
Phylogenomic Analysis of Florida Outbreak Strains ..... 40
Genomic Islands, Pathogenicity Islands, and Virulence Factors ..... 41
O-Antigen Coding Region ..... 43
Open Reading Frame Polymorphisms ..... 44
Virulence Assays ..... 46
Conclusions ..... 48
Chapter 4: *Vibrio* Pathogenicity Island 2 (VPI-2) Diversity ..... 63
Abstract ..... 63
Introduction ..... 64
Methods ..... 65
Strains ..... 65
Comparative Genomics ..... 66
Results and Discussion ..... 67
Gene Structure and Attachment Loci of VPI-2 Varaints ..... 67
Sialic Acid Metabolism ORFs in the Backbone of non-*V. cholerae* Genomes ..... 72
Nucleotide Diversity ..... 73

Evolution of VPI-2 Variants ..... 74
Conclusions ..... 76
Chapter 5: Comparative Genomics Reveals Evidence of Two Novel *Vibrio* Species Closely Related to *V. cholerae* ..... 94
Abstract ..... 94
Introduction ..... 95
Methods ..... 97
Genome Sequencing ..... 97
Comparative Genomics ..... 98
Identification and Annotation of Genomic Islands ..... 99
Phylogenomic Analyses Employing Genome Sequences ..... 100
Results and Discussion ..... 100
Strains ..... 100
General Genome Overview ..... 101
Genome Comparisons ..... 102
Evolution of *Vibrio* sp. RC341 and *Vibrio* sp. RC586 Lineages ..... 105
Virulence Factors ..... 106
Natural Competence ..... 107
Genomic Islands and Integration Loci for Exogenous DNA ..... 108
Horizontal Gene Transfer of Genomic Islands ..... 111
Unique Genomic Islands ..... 112
Conclusions ..... 113
Reproduction License Agreement ..... 139
Chapter 6: Seasonality of *Vibrio metecus* in the Chesapeake Bay, MD ..... 140
Abstract ..... 140
Introduction ..... 141
Materials and Methods ..... 142
Results ..... 145
Discussion ..... 148
Chapter 7: *Vibrio cholerae* in a Historically Cholera-Free Country ..... 160
Abstract ..... 160
Introduction ..... 160
Materials and Methods ..... 161
Results and Discussion ..... 163
Sampling ..... 163
Isolate Characterization and Diversity ..... 164
Expression of Virulence Factors ..... 168
Conclusions ..... 169
Reproduction License Agreement ..... 181
Chapter 8: Molecular Diversity and Predictability of *Vibrio parahaemolyticus* in the Black Sea ..... 188
Abstract ..... 188
Introduction ..... 189
Materials and Methods ..... 192
Sample Collection and Processing ..... 192
Virulence Factors ..... 193

Estimation of Molecular Diversity by REP-PCR ..... 193
Phenotypic Analyses ..... 193
Serotype Determination ..... 194
Statistical Analyses ..... 195
Results ..... 195
Detection of *V. parahaemolyticus* ..... 195
Serodiversity ..... 196
Virulence Factors and Markers of Pandemic Clones ..... 197
REP-PCR ..... 198
Predictive Modeling ..... 198
Discussion ..... 198
Chapter 9: Detection of *Vibrio cholerae* in Environmental Waters Including Drinking Water Reservoirs of Azerbaijan ..... 209
Abstract ..... 209
Introduction ..... 210
Materials and Methods ..... 212
Results and Discussion ..... 213
Chapter 10: Comparison of Methods for Quantifying Bacterial Indicators in an Urban Brackish Water Environment ..... 230
Abstract ..... 230
Introduction ..... 231
Materials and Methods ..... 233
Sample Collection ..... 233
Enumeration of *Escherichia coli* on Modified mTEC Agar ..... 234
Enumeration of Enterococci on mE and BE Agar ..... 234
Enumeration of *Escherichia coli* and Enterococci Using IDEXX Kits ..... 235
Enumeration of Heterotrophic Bacteria ..... 236
Statistical Analyses ..... 236
Results and Discussion ..... 236
Comparison of IDEXX Dilutions ..... 237
Enterococci Densities ..... 238
*Escherichia coli* Densities ..... 239
Correlation between Methods ..... 240
Correlation between Incubation Temperatures ..... 241
Correlation of HPC Counts with the Number Enterococci and *E. coli* ..... 241
Conclusions ..... 242
Chapter 11: Summary and Conclusions ..... 247
Bibliography ..... 251
*Curriculum vitae* ..... 300

# Chapter 1: Introduction

Human pathogens are globally ubiquitous and conventional wisdom holds that they are associated with human contamination or human fecal pollution. Although human contamination does lead to the dispersal of pathogens, itq cannot account for their total presence and it does not account for the initial cases in an epidemic. Further, human absence in an environment does not imply the absence of human pathogens in that environment; thereby demonstrating many human pathogens have an ecology that is not dependent on the human body. For example, *Vibrio cholerae*, the causative agent of cholera and extraintestinal infections, responsible for thousands of deaths annually, was once thought to be spread solely by human fecal contamination (Allen et al., 1975). Decades of research supported by several thousands of publications have definitively proven this organism to be a natural member of aquatic environments on a global scale, its presence being related to non-anthropogenic parameters such as water temperature, salinity, and presence and species composition of zooplankton communities. Still, debates concerning the ecology of genotypes most frequently isolated from human infections continue.

Members of the *Enterobacteriaceae*, such as *Salmonella enterica* and *Escherichia coli*, are known to be shed by the human body into the environment and some pathotypes within this family, such as *Salmonella enterica* subsp. *enterica* serovar Typhi, being subclinically carried and shed by humans for decades (Saphra, 1957; Nath et al., 2010). Thus, the association between these organisms and human fecal contamination has been

clearly and clinically demonstrated. However, human shedding does not account for the total presence of these organisms in the environment as they have been found in regions of little or no human fecal contamination and have been demonstrated to outlive non-pathogenic members of the human intestinal microbiome (Mezrioui et al., 1995). This is demonstrative of their adaptation to, viability in, and reproductive success in the natural (non-mammalian) environment. Thus, they are natural members of the environmental microbiome with the ability to be successful in both the mammalian and non-mammalian environments.

***Ecological Roles of Microbial Virulence***

Prokaryotic and eukaryotic microorganisms can be pathogens of humans and other mammals and both are ubiquitous in the natural environment. It is often environmental exposures that lead to infections with these pathogens. Many of these infections are considered ecological dead-ends in that after the infection, the pathogen is either killed by the immune system of the infected individual or it is shed out of the host body (Levin, 1994; Lainhart et al. 2011). In developed nations this shedding, particularly during enteric infections, occurs into confined sewage systems where human waste is treated and pathogens are effectively killed. Thus, these microorganisms are unable to re-enter the human environment as they are killed before they can encounter food, water, or mammals in/on which they can replicate. Thus, the human → environment → human transmission cycle may occur only for a few pathogens, namely viruses, which are not considered viable organisms. In developing nations where sanitation and effectively treated water is lacking or limited, this human → environment → human transmission

cycle can and does occur. Yet, many infectious diseases that occur in developing nations are also reported in developed nations demonstrating that the environment → human cycle is responsible for many infections in developed nations and the initiation of epidemics in developing nations.

Thus, the environment can be accurately thought of as a reservoir of pathogens from which epidemics arise. This then begs the questions why; are pathogens in the environment, and what is the role of virulence in the environment? To sufficiently address this question, it must be noted that for an organism to become pathogenic to humans it must encode functional suites of genes, many of which are mobile, that produce products that result in an immune response in the infected human. For many pathogens, such as *V. cholerae*, this process relies on the expression of several sets of mobile elements (CTXΦ, *tcpA* of *Vibrio* pathogenicity island-1, and *nanH* of *Vibrio* pathogenicity island-2) and many genes encoded in the backbone (*toxR*, *hapA*,etc.). Absence of one or more of these mobile elements may result in a decreased colonization and infection or no colonization or infection at all. However, in this organism, cells lacking one or more of these virulence factors are frequently isolated from the environment, and sometimes these elements are found in high frequency in a particular environment or in environments where humans do not live (Haley et al., 2012). An example of this is the *Vibrio* pathogenicity island-2, which encodes sialidase (NanH), a sialic acid scavenger that removes this oligosaccharide from the gangliosides of the human intestine making them more available to the cholera toxin. This scavenger also allows sialic acid to be catabolized by the sialic acid metabolism cluster encoded in VPI-2. This pathogenicity island is found in many epidemic strains of *V. cholerae* isolated

from clinical cholera cases (those of the O1 serogroup that are toxigenic and are associated with the current cholera pendemic) as well as non-toxigenic strains isolated in the environment. The island is also known to have many variants, some with major segments deleted and others with additional elements inserted, such as the type III secretion system homologous to TIIISS-2 of *Vibrio parahaemolyticus*. Its ubiquity in the environment suggests it has an ecological role outside of the human body. Further supporting this is the demonstration that the TIIISS of this island is known to induce severe diarrhea in humans (Shin et al., 2011) while its homolog in *V. parahaemolyticus* is known to have a similar function as well as an anti-predation role in the presence of grazing protozoa (Vp T3SS paper). A variant of this island has also been identified in *Vibrio* sp. Ex25, an organism recovered from a deep-sea hydrothermal vent, an environment well outside of the human ecology.

As stated earlier, many of these human infections are relatively brief (non-chronic) and self-limiting, while it is known and demonstrated that these organisms live in the environment and individual cells live longer in the environment that the length of a human infections. Thus, the time that a pathogenic microorganism spends in the environment, interacting with those biotic and abiotic features within that environment, is much greater than the time spent in the human body. Thus, those interactions with the biota and abiota, such as predation, nutrient limitation, temperature shifts, and ultraviolet irradiation shape the microbial genome (Woods et al., 2011). It is further believed that these interactions have led to the evolution of protein products or secondary metabolites occurred in response to these parameters while also coincidentally resulting in an immune response in the human body (Lainhart et al., 2011). While a microorganism can be the

prey of many grazing organisms it is genomically impossible to encode virulence factors that respond to each species of prey. Therefore it is likely that natural selection selects for the evolution and maintenance of those virulence proteins or metabolites that target factors or pathways conserved in eukaryotes as a way of limiting the amount of DNA needed to replicate in the cell cycle. Thus, many virulence factors are able to interact with structures of pathways in the human body as it contains elements conserved in other eukaryotes including protozoan grazers (Lainhart et al., 2011). Examples of virulence factors that cause disease in humans but have a primary role in the environment are listed in Table 1.1.

Many of these virulence factors are mobile while some are conserved in the backbone of all members of a species. In the case of *V. cholerae*, mobile elements are known to be horizontally transferred in the environment (Boyd and Waldor, 1999; Boyd et al., 2000; Faruque at al., 2003; Meibom et al., 2005, Blokesh and Schoolnik, 2007, Miller at al., 2007). Thus, highly pathogenic strains can arise in the environment and lead to massive epidemics and pandemics.

Essentially, clinical isolates are environmental in origin as are their integrated mobile virulence factors. However, it is often presumed that highly pathogenic strains are only present in regions due to the presence of a human carrier and their subsequent shedding of the organism into local and untreated water supplies. Although, proper sanitation that prevents fecal contamination of water and food supplies does reduce human amplification of a cholera outbreak, it does not account for initial cases in regions where autochthonous microorganisms are not removed from drinking water or raw or undercooked seafood is consumed. This assumption that highly pathogenic strains are

not autochthonous to the natural environment is often based on the difficulty in isolating *V. cholerae* strains with "epidemic cholera" features from the environment in regions where cholera does not occur annually. However, the difficulty in isolating such strains is mostly related to the vast genomic diversity of organism in the in the environment such that diversity dilutes any one specific genotype and the difficulty in isolating culturable bacteria from the environment as the majority of cells (> 99%) are in the viable but non-culturable (VBNC) state and cannot be recovered for genotypic analysis (Figure 1.2).

***Objectives of this Study***

Objectives of this study were to take a genomic and ecological approach to understand the ecology of human pathogens in the environment. In this work we demonstrate the non-human environment as a natural ecosystem for several human pathogens. This was achieved by employing a combination of high-throughput whole genome analyses focused on the nucleotide and amino acid level, combined with broader ecological studies evaluating the role of the environment with respect to presence, densities, and virulence expression of the pathogens.

Chapters 2 - 5 present results of genomic analyses of *V. cholerae* and two newly described species, *V. metecus* and *V. parilis*, previously believed to be a "variant" *V. cholerae* of the non-O1/non-O139 serogroups. *V. metecus* and *V. parilis* are biochemically identical to *V. cholerae* and *V. mimicus*, respectively,but genomically unique. These two new species encode many of the genes involved in pathogenesis in *V. cholerae* and *V. mimicus* and show evidence of horizontal exchange with them. Chapter 2 reports on the genomic analysis of an environmental *V. cholerae* strain with a high level

of genomic similarity with those isolated from clinical cases in Egypt 75 years earlier providing evidence of genomes with clinical features being stable in the environment over long periods of time. This further demonstrates the natural habitat of pathogenic *V. cholerae* strains is the aquatic environment. The overriding theme of the four chapters is that bacteria sharing genome-wide similarity with pathogenic *V. cholerae* can be isolated from the environment in regions where cholera and other *V. cholerae* infections are sporadic at most. The data provided in the chapters present genomic evaluation of virulence factors of *V. cholerae*, namely *Vibrio* pathogenicity island 2 (VPI-2) about which a separate chapter is provided.

Chapters 6 - 10 present findings of ecological studies on *V. cholerae*, *V. metecus*, and *V. parahaemolyticus* conducted over several years in the Chesapeake Bay, Iceland, the Georgian coast of the Black Sea, and waters bodies in Azerbaijan, including the Caspian Sea. These regions proved to be significant since vibrioses occur sporadically in these locations. In fact, in Iceland there has never been a recorded *Vibrio* infection and this is a reliable conclusion because the country is known for its highly detailed and uninterrupted health record keeping for the entire population. Furthermore, vibrioses are reportable infections in Iceland. The ecological study on *V. metecus* is the first for this organism and the results reveal an ecology different than that of its nearest phylogenetic neighbor, *V. cholerae*. This organism is an emerging pathogen and has been shown to share many genes with *V. cholerae* demonstrating that its emergence is of public health significance.

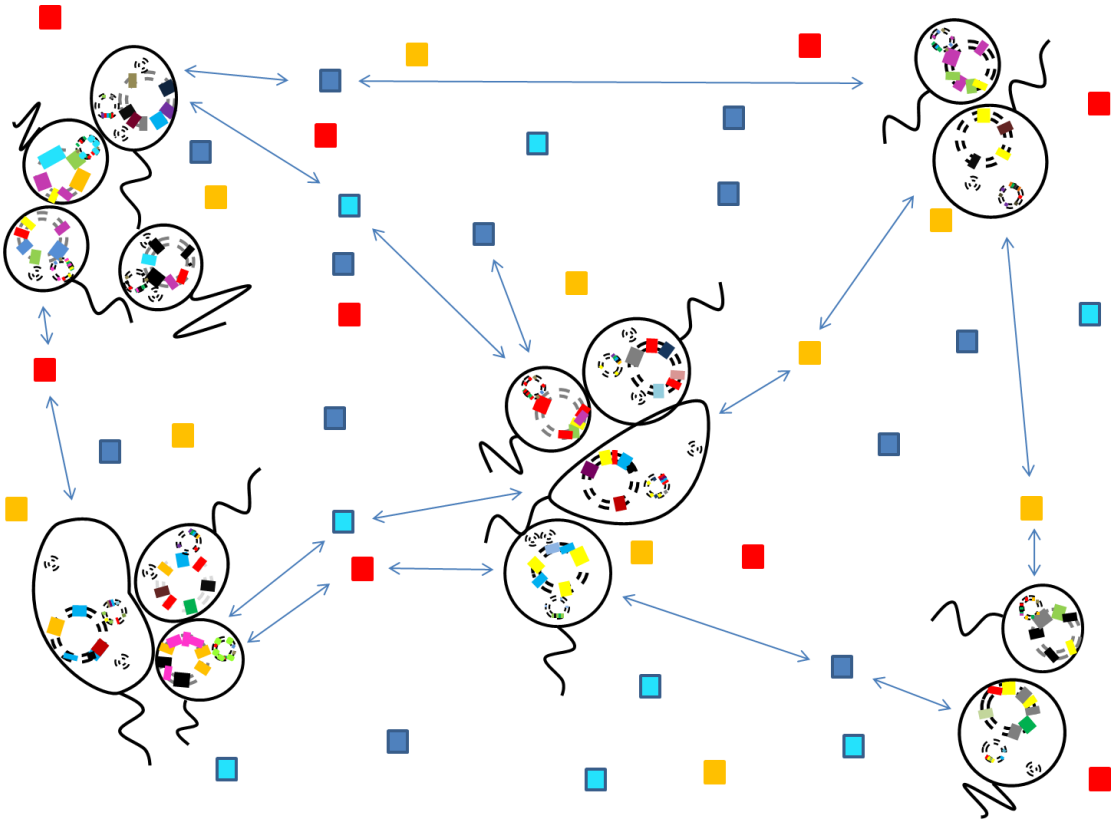


**Figure 1.1.** Genomic variation and gene flow in the environment. Colored squares represent transmissible genetic elements.

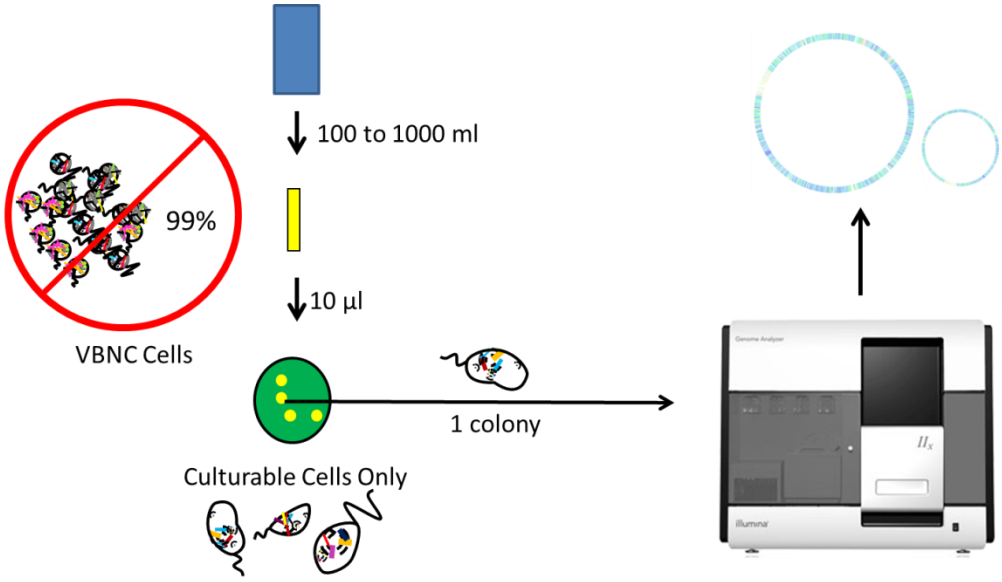



**Figure 1.2.** Schematic representation of sample processing to genotyping.

| Organism | Virulence Factor | Role in Virulence in Mammalian Hosts | Actions in the Environment | Reference |
|---|---|---|---|---|
| ***Vibrio cholerae*** | Type VI Secretion System | Induces an inflammatory diarrhea and facilitates replication within the intestine. | Protection from amoebae grazing. | (Pukatzki et al., 2006) |
| | Toxin Co-regulated Pilus (VPI-1) | Attachment to intestinal epithelial cells and formation of biofilms inside intestine. | Attachment to plankton. | (Reguera and Kolter, 2005) |
| | prtV (protease) | Cytotoxicity to human intestinal cells and fibronectin and fibrinogen degradation. | Protection from eukaryotic grazing. | (Vaitkevicius et al., 2006) |
| | HA/P (hemaglutinnin/protease) | Involved in release of *V. cholerae* cells from intestine. | Utilize Chironomid egg masses as a food source. | (Halpern et al., 2003) |

Table 1.1. Roles of virulence factors in the human body and in the non-human environment.

# Chapter 2: Pre-7th Pandemic *Vibrio cholerae* BX 330286 El Tor Genome; Evidence for the Environment as a Genome Reservoir


## ***Abstract***

*Vibrio cholerae* O1 El Tor BX 330286 was isolated from a water sample in Australia in 1986, nine years after an indigenous outbreak of cholera occurred in that region. This environmental strain encodes virulence factors highly similar to those of clinical strains, suggesting an ability to cause disease in humans. We demonstrate its high similarity in gene content and genome-wide nucleotide sequence to clinical *V. cholerae* strains, notably to pre-7th pandemic O1 El Tor strains isolated in 1910 (*V. cholerae* NCTC 8457) and 1937 (*V. cholerae* MAK 757), as well as 7th pandemic strains isolated after 1960 globally. Here we demonstrate that this strain represents a transitory clone with shared characteristics between pre- 7th and 7th pandemic strains of *V. cholerae*. Interestingly, this strain was isolated twenty-five years after the beginning of the 7th pandemic, suggesting the environment as a genome reservoir in areas where cholera does not occur in sporadic, endemic or epidemic form.


## ***Introduction***

*Vibrio cholerae* is an autochthonous member of marine, estuarine, and freshwater microbial communities worldwide and the causative agent of cholera, a profuse secretory-diarrhea that can lead to death by dehydration if left untreated. Cholera is

caused by the O1 and O139 serogroups of *V. cholerae*, while some of the more than 200 known non-O1/non-O139 serogroups are responsible for sporadic cases of diarrhea and extraintestinal *V. cholerae* infections (Safrin et al., 1988; Ko et al., 1998; Lukinmaa et al., 2006; Shannon et al., 2006; Chatterjee et al., 2009). Serogroup O1 strains can be further divided into either Classical or El Tor biotypes based on a distinct combination of biochemical and phenotypic traits, including phage and Polymyxin B susceptibility, acetoin production, and allelic variation of *rstR* of CTXΦ and *tcpA* of *Vibrio* Pathogenicity Island 1 (Mukerjee, 1963; Han and Khie, 1963; Davis et al., 1999; Jonson et al., 1991).

It has been observed that one biotype predominates in clinical cases in a pandemic. There have been seven cholera pandemics recorded since 1817, with the 5$^{th}$ and 6$^{th}$ pandemics caused by strains of the serogroup O1 biotype Classical and the current 7$^{th}$ pandemic by strains of serogroup O1 biotype El Tor. The interpandemic period between the 6$^{th}$ (ending in 1923) and 7$^{th}$ pandemics (beginning in 1961) was 38-years. During this period localized cholera epidemics occurred sporadically, but it is presumed that cholera did not occur in the form of a pandemic. Biotype El Tor strains isolated during this period are termed "pre-7$^{th}$ pandemic" (Byun et al., 1999). The first cluster of cases of the 7$^{th}$ pandemic is believed to have occurred in 1961 on the island of Sulawesi (formerly known as Celebes), an island of Indonesia, and rapidly spreading to Asia and Africa, and continuing globally. Recent studies have demonstrated that the flexible genome of *V. cholerae* allows genomic drift and, therefore, adaptation to novel environments, as well as avoidance of the host immune response (Chun et al., 2009). Lineages, however, may retain a relatively persistent genome over time that undergoes

little sequence divergence, suggesting that a part of the *V. cholerae* genome remains static.

*V. cholerae* strain BX 330286 is an O1 Inaba biotype El Tor strain isolated from a water sample in 1986 from Australia, in an area geographically close to Sulawesi (ca. 1100 km at their closest locations) (Safa et al., 2009). These waters were determined to be the source of Australia's first recorded indigenous cholera outbreak in 1977, when epidemiological investigations determined that surface water was the vehicle for *V. cholerae* O1 El Tor infections among the local populations (Rao and Stockwell, 1980) (Rogers et al., 1980). A PCR-based multilocus genetic analysis suggested it was similar to pre-7$^{th}$ pandemic strains as it did not encode VSP-I and II (Safa et al., 2009).

A phylogeny of 23 *V. cholerae* strains, including BX 330286, was determined by Chun et al. (2009), based on 1,676 homologous genes (1,370,469 bp). The evolutionary pathway based on these homologous regions of the genome revealed four major clades in the evolution of *V. cholerae*. A non-O1/non-O139 clade, dominated by environmental isolates, comprises the most heterogenous group of sequenced strains. Twelve other strains formed a Phylocore Genome (PG) clade, comprising O1 and O139 clinical strains, and another clinical strain that putatively received an O37 antigen via horizontal gene transfer. *V. cholerae* BX 330286 is a member of Phylocore Genome group-1 (PG-1), consisting mainly of El Tor clinical isolates, some of which had been isolated prior to the 7$^{th}$ pandemic, with a relatively short branch length to the ancestral node of the 7$^{th}$ pandemic clade, indicating a relatively short evolutionary distance from that ancestral or hypothetical progenitor clinical strain of the 7$^{th}$ pandemic strains isolated since 1961 (Chun et al., 2009). These data suggest that *V. cholerae* BX 330286 represents a

genomically persistent transitionary strain between the pre-7th pandemic El Tor strains and those of the 7th pandemic clade. *V. cholerae* O1 El Tor strains demonstrating similar pre-7th pandemic characteristics were also isolated from clinical cases in this region in 1998 (Nair et al., 2006).

The aim of this study was to investigate and characterize the genome of environmentally isolated *V. cholerae* BX 330286, a strain genetically similar to pre-7th pandemic El Tor strains (Safa et al., 2009), with a close phylogenetic relationship to the progeny of the hypothetical ancestor of the 7th pandemic clade (Chun et al., 2009). To accomplish this we compared the genome of *V. cholerae* BX 330286 to pre-7th pandemic and 7th pandemic El Tor, biotype Classical, and non-O1/non-O139 *V. cholerae* isolated globally from clinical cases and the environment. Here we present an analysis of this genome with an emphasis on its genome-wide similarity to two pre-7th pandemic clinical strains isolated during the 6th cholera pandemic in 1910 (*V. cholerae* NCTC 8457 from Saudi Arabia) and between the 6th and 7th pandemics in 1937 (*V. cholerae* MAK 757 from modern-day Sulawesi) and six 7th pandemic strains, particularly *V. cholerae* N16961, a strain isolated early in the 7th pandemic. Our findings suggest individual genomes demonstrating clinical strain characteristics and relatively little sequence divergence can persist in the environment over time where cholera is not endemic.

## ***Materials and Methods***

### Genome Sequencing

Draft sequences were obtained from a blend of Sanger and 454 sequences and involved paired end Sanger sequencing on 8kb plasmid libraries to 5X coverage, 20X

coverage of 454 data, and optional paired end Sanger sequencing on 35kb fosmid libraries to 1-2X coverage (depending on repeat complexity). To finish the genomes, a collection of custom software and targeted reaction types were used. In addition to targeted sequencing strategies, Solexa data in an untargeted strategy were used to improve low quality regions and to assist gap closure. Repeat resolution was performed using in-house custom software (Han and Chain, 2006). Targeted finishing reactions included transposon bombs, an *in vitro* transposon insertion strategy involving random insertion of a yeast transposable element into a gap followed by amplification off both ends of the element, primer walks on clones and PCR products, and adapter PCR reactions (Goryshin and Reznikoff, 1998). Gene-finding and annotation were achieved using an automated annotation server (Aziz et al., 2008). All genome sequences are available from the NCBI Genbank and accession numbers are listed in Table 2.1.

**Comparative Genomics**

Genome to genome comparison was performed using three approaches, since completeness and quality of nucleotide sequences varied from strain to strain among the set of strains examined in this study. First, nucleotide sequences as whole contigs were directly aligned using the MUMmer program (Kurtz et al., 2004). Second, ORFs of a given pair of genomes were reciprocally compared to each other, using BLASTN, BLASTP, and TBLASTX programs (ORF-dependent comparison). Third, a bioinformatic pipeline was developed to identify homologous regions of a given query ORF. Initially, a segment on target contig homologous to a query ORF was identified using the BLASTN program. This potentially homologous region was expanded in both

directions by 2,000 bp, after which the nucleotide sequences of the query ORF and selected target homologous region were aligned, using a pairwise global alignment algorithm (Myers and Miller, 1988) and the resultant matched region in the subject contig was extracted and saved as a homolog (ORF-independent comparison). Orthologs and paralogs were differentiated by reciprocal comparison. In most cases, both ORF-dependent and –independent comparisons yielded the same orthologs, though the ORF-independent method performed better for draft sequences of low quality, in which sequencing errors, albeit rare, hampered identification of correct ORFs. All genomic islands in the *V. cholerae* strains used in this study have been identified by Chun et al. (2009).

To estimate the genomic similarity among strains we determined the average nucleotide identity (ANI), by performing a reciprocal best match BLASTN analysis for each genome with *V. cholerae* BX 332086 (Konstantinidis and Tiedje, 2005).

**Phylogeny of Mobile Genetic Elements**

To infer phylogeny of the mobile elements among strains, neighbor-joining trees were developed, using homologous ORFs between the strains included in this study. Sequences were aligned using CLUSTALW2 (Larkin et al., 2007). Trees were developed using the Kimura-2-parameter of nucleotide substitution (Kimura, 1980). One hundred bootstrap replications were executed for each tree. Phylogenetic estimations were performed using MEGA software (Kumar et al., 2007).

## ***Results***

### General Genome Features, Virulence Factors, and Genomic Islands

The draft genome of *V. cholerae* BX 330286 spans 8 contigs and putatively encodes 3,663 coding sequences covering 4,000,672 bp. Of the 3,663 coding sequences, 778 (26.5%) are annotated as hypothetical proteins, 416 of which are located on the large chromosome and 362 on the small chromosome. This strain encodes 111 RNAs, 90 of which are annotated as tRNAs.

Similar to the clinical strains, *V. cholerae* BX 330286 encodes several pathogenicity islands; however, this strain does not encode *Vibrio* seventh pandemic islands (VSP) I and II, hallmarks of clinical strains of the 7$^{th}$ pandemic (Safa et al., 2009). *V. cholerae* BX 330286 encodes *Vibrio* pathogenicity islands (VPI) 1 and 2, toxin linked cryptic plasmid (TLC), cholera toxin prophage (CTXΦ), and the RS1Φ element. VPI-1 of *V. cholerae* BX 330286 has 100% nucleotide sequence similarity with *V. cholerae* MAK 757, 99.9% similarity with 7$^{th}$ pandemic strains, and 99% similarity with NCTC 8457. Based on the *tcpA* sequence, *V. cholerae* BX 330286 encodes VPI-1$^{\text{El Tor}}$ in that it has 100% sequence similarity with all O1 El Tor strains and 99% sequence similarity with *V. cholerae* O395 Classical. This is supported by a phylogenetic analysis which infers that the donor of this island is an El Tor strain as the *tcpA* sequences of the El Tor strains group together in an unresolved branch (Figure 2.1). Although the *tcpA* phylogeny is unresolved, its nucleotide sequence similarity suggests that *V. cholerae* BX 330286 and MAK 757 received VPI-1 from a similar source (Table 2.2).

VPI-2 of *V. cholerae* BX 330286 has 100% sequence similarity across the complete island (100% conserved) with *V. cholerae* N16961, 2740-80, and V52 and

99.9% with *V. cholerae* NCTC 8457, O395 Classical, and the other 7$^{th}$ pandemic strains, excluding *V. cholerae* MO10 which has a truncated island and *V. cholerae* MJ-1236 with which it has 99.7% sequence identity (Table 2.2). VPI-2 of *V. cholerae* BX 330286 has 98.2% sequence similarity across 66.6% of the ORFs with *V. cholerae* MAK 757 (Table 2.2). The divergence and low percent conserved ORFs between the two homologous islands can be attributed to VPI-2 variants encoded by the two strains suggesting a different source of the island for these two strains. A phylogenetic analysis of the phage integrase from this island (the only ORF conserved among all strains encoding this island) showed this ORF clustering with all other clinical strains, including the 7$^{th}$ pandemic and pre-7$^{th}$ pandemic strains (B.J. Haley, unpublished). Although it cannot be inferred by phylogenetic analysis, the sequence similarity and high percent of conservation across the ORFs suggests that *V. cholerae* BX 330286 acquired canonical VPI-2 from either a 7$^{th}$ pandemic strain or from an ancestral strain that donated VPI-2 to the 7$^{th}$ pandemic clade.

*V. cholerae* BX 330286 contains the TLC element, proximal to CTXΦ on the large chromosome, sharing 100% sequence similarity with that of all other strains carrying the element suggesting, an highly conserved state (Table 2.2). This strain also encodes tandem copies of CTXΦ$^{CL}$, based on the *rstR* sequence (Chun et al., 2009; Safa et al., 2009). Similarly, MAK 757 encodes CTXΦ$^{CL}$ but only a single copy (Chun et al., 2009). The tandem CTXΦ$^{CL}$ copies in *V. cholerae* BX 330286 are identical in sequence, most likely from duplication of the prophage. The CTXΦ of *V. cholerae* BX 330286 shows divergence from other CTXΦ, with higher sequence similarity to the CTXΦ of the *V. cholerae* 7$^{th}$ pandemic strains than all other strains (Table 2.2). This finding suggests

that although the prophage is a Classical biotype, based on *rstR* sequence, it has overall higher sequence similarity to CTXΦ of El Tor strains (Table 2.1). A phylogenetic analysis of a ca. 5.5 kb homologous region of the CTXΦ (covering *rstA* to *ctxB*) of strains included in this study infers CTXΦ of *V. cholerae* BX 330286 shares a common ancestor with 7$^{th}$ pandemic strains diverging from CTXΦ of *V. cholerae* MAK 757 (Figure 2.2). This strain also encodes a copy of RS1Φ$^{ENV}$ suggesting a non-7$^{th}$ pandemic or pre-7$^{th}$ pandemic source of this element (Chun et al., 2009; Safa et al., 2009).

*V. cholerae* BX 330286 encodes *V. cholerae* GIs-1, 2, and 5. GI-1 of *V. cholerae* BX 330286 has a higher sequence similarity with *V. cholerae* NCTC 8457, MAK 757, the 7$^{th}$ pandemic clade, V52, 2740-80, and O395 Classical than with other strains encoding the same island (Table 2.2). GI-2 of *V. cholerae* BX 330286 has a higher sequence similarity with *V. cholerae* NCTC 8457, MAK 757, 2740-80, V52, and O395 Classical, than with other strains encoding this island (Table 2.2). GI-5 of *V. cholerae* BX 330286 has 100% sequence similarity with all other strains encoding this island (Table 22).

**Conservation of the Superintegron Region among Clinical Strains**

The genome of *V. cholerae* BX 330286 encodes a superintegron region putatively encoding 236 ORFs, 114 of them are hypothetical proteins. Results of the analysis demonstrate that this region is highly conserved in clinical strains and *V. cholerae* BX 330286, sharing most of the ORFs with the superintegron region of *V. cholerae* NCTC 8457 (225 ORFs, 95% of ORFs in the superintegron) with an average of 99.3% nucleotide sequence identity between these ORFs (Table2. 2) and with only 6 ORFs

showing less than 100% nucleotide sequence identity (Data not shown). Interestingly, the superintegron region encodes two pre-7$^{th}$ pandemic regions. One region is a 6 ORF duplication spanning ca. 3 kb (NCBI Genbank locus tags VCF_001728 to VCF_001733) found in *V. cholerae* BX 330286, NCTC 8457, and MAK 757 (Figure 2.3). Four of the ORFs are annotated as hypothetical proteins, one as a plasmid stabilization system, and another as a microcin immunity protein MccF. Another region unique to these three strains is ca. 6.3 kb, containing 8 ORFs (VCF_001761 to VCF_001768), five of which show no match to 7$^{th}$ pandemic or *V. cholerae* O395 Classical and 2 to 3 which have a non-reciprocal match with ORFs from these strains (Figure 2.3). The 5 ORF unique region encodes 4 hypothetical proteins and one protein annotated as Sll1503 protein, and the non-reciprocal region encodes 2 hypothetical proteins and one annotated as lactoylglutathione lyase and related lyases (Figure 2.3). The superintegron region of *V. cholerae* BX 330286 also has high similarity to *V. cholerae* N16961 and RC9, in that it has 203 and 201 ORFs in common and 99.8 and 99.3% nucleotide sequence identity across homologous ORFs from the superintegron regions of these two strains, respectively (Table 2.2). These data indicate a clonal relationship of this region between *V. cholerae* BX 330286 and NCTC 8457, as well a significant similarity to current 7$^{th}$ pandemic strains.

### Genome-Wide Nucleotide Similarity Shows High Similarity to Pre-7$^{th}$ Pandemic and 7$^{th}$ Pandemic Strains

To estimate the genomic similarity among strains, we determined the average nucleotide identity (ANI) between protein-coding sequences of *V. cholerae* BX 330286

and 22 *V. cholerae* genomes (Table 2.1). By this analysis *V. cholerae* BX 330286 has the highest genomic similarity with six strains of the 7$^{th}$ pandemic clade (N16961, RC9, B33, CIRS101, MO10, and MJ-1236) and the pre-7$^{th}$ pandemic strains *V. cholerae* NCTC 8457 and MAK 757 (Table 2.1). *V. cholerae* BX 330286 shares 100% nucleotide sequence similarity with 3,272 of 3,627 (90%) protein-coding ORFs with *V. cholerae* N16961 but has more ORFs with 100% sequence similarity with *V. cholerae* NCTC 8457 (ca. 93.41%) and MAK 757 (ca. 92.53%) than any other strain in this analysis (Table 2.1). The ANI range for *V. cholerae* BX 330286 and NCTC 8457 and MAK757 was larger than the 7$^{th}$ pandemic clade strains, suggesting selective pressure over time may have caused sequence divergence of relatively few ORFs among *V. cholerae* BX 330286 and the pre-7$^{th}$ pandemic group.

There were 215 ORFs in *V. cholerae* BX 330286 that showed divergence with homologs in *V. cholerae* NCTC 8457 (ca. 6.6 % of shared ORFs) and 258 ORFs that showed divergence with homologs in MAK 757 (ca. 7.5% of shared ORFs). Fifty-six of these ORFs showed divergence between homologs in *V. cholerae* BX 330286 and both pre-7$^{th}$ pandemic strains and 105 of these ORFs also showed divergence with homologs of at least one 7$^{th}$ pandemic strain. The most notable of these are *hapR*, a quorum-sensing regulator of virulence and biofilm formation, putative RTX toxin transporter, putative RTX toxin, putative response regulator, RNA polymerase sigma factor *rpoS*, and general secretion pathway protein K involved in CT translocation. Interestingly, these ORFs are involved in *V. cholerae* virulence or environmental response.

Further, there are seven ORFs which show divergence between *V. cholerae* BX 330286 and both pre-7$^{th}$ pandemic strains, but not between *V. cholerae* BX 330286 and

the 7th pandemic strains. These are topoisomerase IV subunit A, predicted ATPase related to phosphate starvation-inducible protein PhoH, putative acetyltransferase, heat shock protein 60 family chaperone GroEL, spermidine putrescine ABC transporter permease component PotB, myo-inositol-1(or 4)-monophosphatase, and secreted trypsin-like serine protease. By our methods, several ORFs were determined to have diverged dramatically in *V. cholerae* BX 330286 from those of *V. cholerae* NCTC 8457 and MAK 757. These include phage replication protein *rstA*, hypothetical protein, methyl-accepting chemotaxis protein, and L-serine dehydratase. However, none of these ORFs have, as of yet, been implicated in environmental persistence or pandemic potential and most likely do not account for the persistence of *V. cholerae* BX 330286. RstA, however, is involved in CTXΦ prophage replication, but its homolog in *V. cholerae* NCTC 8457, a CTXΦ-negative strain, is part of the GI-19 phage-like element (Chun et al., 2009). It should be noted that *V. cholerae* BX 330286 has the highest ANI across the conserved core of all *V. cholerae* strains used in this study with MAK 757, RC9, and N16961 (Table 2.1).

**Operon Divergence in 7th Pandemic Clade**

Interestingly, there are 266 ORFs found in all 7th pandemic strains that show divergence with homologous ORFs of both the pre-7th pandemic strains and *V. cholerae* BX 330286 (7.6% of *V. cholerae* BX 330286 genome). One hundred ninety nine of these ORFs occur in contiguous sets of 2 or more ORFs, totaling 35 sets of contiguous ORFs with nucleotide sequence divergence suggesting that several operons of the 7th pandemic clade have undergone partial or complete sequence divergence with respect to those of *V. cholerae* BX 330286. These diverged ORFs include GI-2, a carbohydrate

phosphotransferase system (PTS), maltose regulon, iron-transport system, glycerol uptake and metabolism, tryptophan metabolism, D-gluconate and ketogluconates metabolism, pyruvate metabolism, deoxyribose/deoxynucleoside catabolism, and polyol metabolism operons. Interestingly, maltose metabolism has been shown to affect cholera toxin secretion (Lång, et al. 1994), suggesting that the observed divergence (98.7% sequence similarity) of *malT*, the transcriptional activator of the maltose regulon, in the $7^{th}$ pandemic strains may have contributed to an alteration of virulence and clinical success in the $7^{th}$ pandemic strains. Furthermore, divergence in three GGDEF domain protein coding sequences was observed. These proteins comprise a family involved in cyclic diguanylate degradation and synthesis, a signaling molecule involved in virulence regulation (Tischler et al., 2005; Matson et al., 2007). These data suggest that $7^{th}$ pandemic strains have undergone a uniform sequence divergence of these ORFs. The consistency of this sequence divergence across all strains of the $7^{th}$ pandemic clade suggests that this divergence occurred when this group evolved from the $7^{th}$ pandemic ancestor and may have conferred increased success in the human host.

## ***Conclusions***

*V. cholerae* BX 330286, an environmental O1 El Tor strain isolated from a water sample in Australia in 1986, demonstrates high genome-wide similarity to clinical strains, especially the pre-$7^{th}$ pandemic strains *V. cholerae* NCTC 8457 (isolated in Saudi Arabia in 1910) and *V. cholerae* MAK 757 (isolated in Sulawesi in 1937), and $7^{th}$ pandemic strains (isolated globally). This strain encodes all major virulence factors, including a tandemly repeated $CTX\Phi^{CL}$, but does not encode either VSP-I and II, two hallmarks of

the 7th pandemic clade. The high genomic similarity of this strain with pre-7th pandemic and 7th pandemic strains and the results of the phylogenetic analysis of this strain and other *V. cholerae* strains suggest that *V. cholerae* BX 330286 has a genomic structure similar to pre-7th pandemic strains and an early progeny of the 7th pandemic clade ancestor as was suggested by a multilocus genetic analysis (Safa et al., 2009). *V. cholerae* BX 330286 appears to be "frozen in time"; i.e., originating from the pre-7th pandemic period before divergence of 7th pandemic strains, but isolated 25 years after the onset of the 7th pandemic. This lineage has undergone divergence from the pre-7th pandemic El Tor strains in relatively a few ORFs compared to the 7th pandemic strains, which contributes to the finding of this genome being a genomically transitory clone between the pre-7th pandemic and 7th pandemic strains. However the *V. cholerae* BX 330286 genome shares a greater percent of ORFs with 100% similarity to the pre-7th pandemic strains, further grounding the idea that this strain is an environmental pre-7th pandemic clone. Furthermore, it appears that the 7th pandemic strains evolved from the pre-7th pandemic El Tor strains through a genomic divergence, not across the whole genome but rather localized to several operons, as well as an insertion of the VSP-I and II islands. It is concluded that the *V. cholerae* genome is highly plastic and flexible, but clonal strains persist in the environment with relatively little sequence divergence, while novel genotypes emerge. Essentially, lineages radiate in gene content and sequence identity while others exhibit minor divergence over generations.

| Strain | Phylogenetic clade | Pandemic period during which strain was isolated | Serogroup | Biotype | Source | Location | Year of Isolation | ANI(%) with BX 330286 | ANI range for ORFs shared with BX 330286 | No. of ORFs shared with BX 330286 | ANI(%) of conserved *V. cholerae* core with BX 330286 | Percent of ORFs with 100% ANI with BX 330286 | NCBI Genbank accession No. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| N16961 | 7P [A] | 7th | O1 | El Tor | clinical | Bangladesh | 1975 | 99.9 | 0.17 | 3627 | 99.87 | 90.21 | AE003852/AE003853 |
| RC9 | 7P | 7th | O1 | El Tor | clinical | Kenya | 1985 | 99.89 | 0.17 | 3625 | 99.87 | 90.01 | ACHX00000000 |
| CIRS101 | 7P | 7th | O1 | El Tor | clinical | Bangladesh | 2002 | 99.89 | 0.21 | 3590 | 99.86 | 88.69 | ACVW00000000 |
| B33 | 7P | 7th | O1 | El Tor | clinical | Mozambique | 2004 | 99.88 | 0.17 | 3590 | 99.86 | 88.36 | ACHZ00000000 |
| MAK 757 | PG1 | interpandemic | O1 | El Tor | clinical | Sulawesi | 1937 | 99.87 | 0.54 | 3480 | 99.87 | 92.53 | AAUS00000000 |
| MJ-1236 | 7P | 7th | O1 | El Tor | clinical | Bangladesh | 1994 | 99.87 | 0.42 | 3502 | 99.86 | 88.72 | CP001485/CP001486 |
| MO10 | 7P | 7th | O139 | | clinical | Madras, India | 1992 | 99.85 | 0.34 | 3505 | 99.85 | 88.73 | AAKF03000000 |
| NCTC 8457 | PG1 | 6th | O1 | El Tor | clinical | Saudi Arabia | 1910 | 99.77 | 0.58 | 3507 | 99.85 | 93.41 | AAWD01000000 |
| 2740-80 | PG1 | 7th | O1 | El Tor | water | US Gulf Coast | 1980 | 99.62 | 0.57 | 3446 | 99.71 | 85.55 | AAUT01000000 |
| O395 | PG2 | 7th | O1 | Classical | clinical | India | 1965 | 99.45 | 0.33 | 3582 | 99.33 | 58.63 | CP000626/CP000627 |
| V52 | PG2 | 7th | O37 | | clinical | Sudan | 1968 | 99.27 | 0.57 | 3478 | 99.23 | 61.85 | AAKJ02000000 |
| 12129(1) | | 7th | O1 | El Tor | water | Australia | 1985 | 98.37 | 0.47 | 3369 | 98.32 | 8.28 | ACFQ00000000 |
| AM-19226 | | 7th | O39 | | clinical | Bangladesh | 2001 | 98.37 | 0.52 | 3288 | 98.36 | 6.93 | AATY01000000 |
| MZ-O2 | | 7th | O14 | | clinical | Bangladesh | 2001 | 98.29 | 0.55 | 3224 | 98.26 | 6.64 | AAWF01000000 |
| 1587 | | 7th | O12 | | clinical | Lima, Peru | 1994 | 98.15 | 0.58 | 3248 | 98.27 | 6.74 | AAUR01000000 |
| 623-39 | | 7th | non-O1/O139 | | water | Bangladesh | 2002 | 98.13 | 0.57 | 3270 | 98.31 | 6.88 | AAWG00000000 |
| TMA21 | | 7th | non-O1/O139 | | water | Brazil | 1982 | 98.11 | 0.55 | 3295 | 98.16 | 6.95 | ACHY00000000 |
| MZ-O3 | | 7th | O37 | | clinical | Bangladesh | 2001 | 98.09 | 0.57 | 3222 | 98.36 | 9.40 | AAUU01000000 |
| TM-11079 | | 7th | O1 | El Tor | sewage | Brazil | 1980 | 97.95 | 0.48 | 3299 | 97.96 | 6.34 | ACHW00000000 |
| VL426 | | unknown | non-O1/O139 | | water | UK | Unknown | 97.83 | 0.48 | 3203 | 97.85 | 4.87 | ACHV00000000 |
| V51 | | 7th | O141 | | clinical | USA | 1987 | 97.78 | 0.58 | 2991 | 97.91 | 4.38 | AAKI02000000 |
| RC385 | | 7th | O135 | | plankton | Chesapeake Bay | 1998 | 97.1 | 0.58 | 2744 | 97.30 | 3.68 | AAKH02000000 |
| BX 330286 | PG1 | 7th | O1 | El Tor | water | Australia | 1986 | 100 | 0 | 3663 | 100 | 100 | ACIA00000000 |
| Mean [B] | | | | | | | | 99.75 | 0.37 | 3539 | 99.74 | 84.24 | |

A = Evolutionary clade as determined by Chun *et al.,* (2009)

B = Does not include values from BX 330286 versus BX 330286

Table 2.1. *V. cholerae* strains used in this study and results of their comparative nucleotide sequence analysis with *V. cholerae* BX 330286.

| | | Element | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Strain | Clade | CTXΦ [B] | TLC | VPI-1 | VPI-2 | GI-1 | GI-2 | GI-5 | Superintegron |
| N16961 | 7P [A] | 99.22[C] (100)[D] | 100 (100) | 99.9 (100) | 100 (100) | 100 (100) | 98.6 (100) | 100 (100) | 99.8 (86) |
| RC9 | 7P | 99.19 (100) | 100 (100) | 99.9 (100) | 99.9 (100) | 100 (100) | 98.6 (100) | 100 (100) | 99.2 (85) |
| CIRS101 | 7P | 99.19 (100) | 100 (100) | 99.9 (100) | 99.9 (100) | 100 (100) | 98.6 (100) | 100 (100) | 99.3 (72) |
| B33 | 7P | 99.19 (100) | NP | 99.9 (100) | 99.9 (100) | 100 (100) | 98.6 (100) | 100 (100) | 98.9 (73) |
| MAK 757 | PG1 | 98.51 (100) | 100 (100) | 100 (100) | 98.2 (66.6) | 100 (100) | 100 (100) | 100 (100) | 98.5 (67) |
| MJ-1236 | 7P | 99.22 (100) | NP | 99.9 (100) | 99.7 (100) | 100 (100) | 98.6 (100) | 100 (100) | 98.7 (74) |
| MO10 | 7P | 99.22 (100) | 100 (100) | 99.9 (100) | 100 (47.9) | 100 (100) | 98.6 (100) | 100 (100) | 98.3 (57) |
| NCTC 8457 | PG1 | NP | 100 (100) | 99 (100) | 99.9 (100) | 100 (100) | 100 (100) | 100 (100) | 99.3 (95) |
| 2740-80 | PG1 | NP | 100 (100) | 99.1 (100) | 100 (100) | 100 (100) | 100 (100) | 100 (100) | 98.7 (72) |
| O395 | PG2 | 98.50 (100) | 100 (100) | 98 (100) | 99.9 (100) | 100 (100) | 100 (100) | 100 (100) | 98.9 (70) |
| V52 | PG2 | 99.05 (100) | 100 (100) | 98 100) | 100 (100) | 100 (100) | 100 (100) | 100 (100) | 98.8 (70) |
| 12129(1) | | NP | NP | 98.8 (100) | 93.8 (54) | 98.9 (100) | NP | NP | 94.4 (34) |
| AM-19226 | | NP | NP | NP | 94.7 (47.9) | 99.6 (100) | NP | NP | 93.6 (30) |
| MZ-O2 | | NP | NP | NP | NP | 99.6 (100) | NP | NP | 93.7 (33) |
| 1587 | | NP | NP | NP | 95.5 (47.9) | 99.5 (100) | NP | NP | 93.9 (36) |
| 623-39 | | NP | NP | NP | 96.9 (45.8) | 99.6 (100) | NP | NP | 92.4 (37) |
| TMA21 | | NP | NP | NP | 91.8 (50) | 99.5 (100) | NP | NP | 95.1 (32) |
| MZ-O3 | | NP | NP | NP | NP | NP | NP | NP | 95.1 (50) |
| TM-11079 | | NP | NP | NP | 96.1 (97.9) | NP | NP | NP | 91.1 (29) |
| VL426 | | NP | NP | 93.8 (97) | NP | NP | 99.3 (100) | NP | 91.4 (23) |
| V51 | | 98.94 (100) | NP | 93.8 | 94.4 (47.9) | 99 (100) | NP | NP | 92.4 (25) |
| RC385 | | NP | NP | NP | NP | NP | 99.1 (100) | NP | 92.4 (33) |

A = Evolutionary clade as determined by Chun *et al.*, (2009)

B = Two identical tandem copies in *V. cholerae* BX 330286

C = Percent identity across conserved ORFs

D = Percent ORFs conserved compared with *V. cholerae* BX 330286

NP = Not present

Table 2.2. Percent identity and percent ORFs conserved between *V. cholerae* strains and *V. cholerae* BX 330286 genomes.

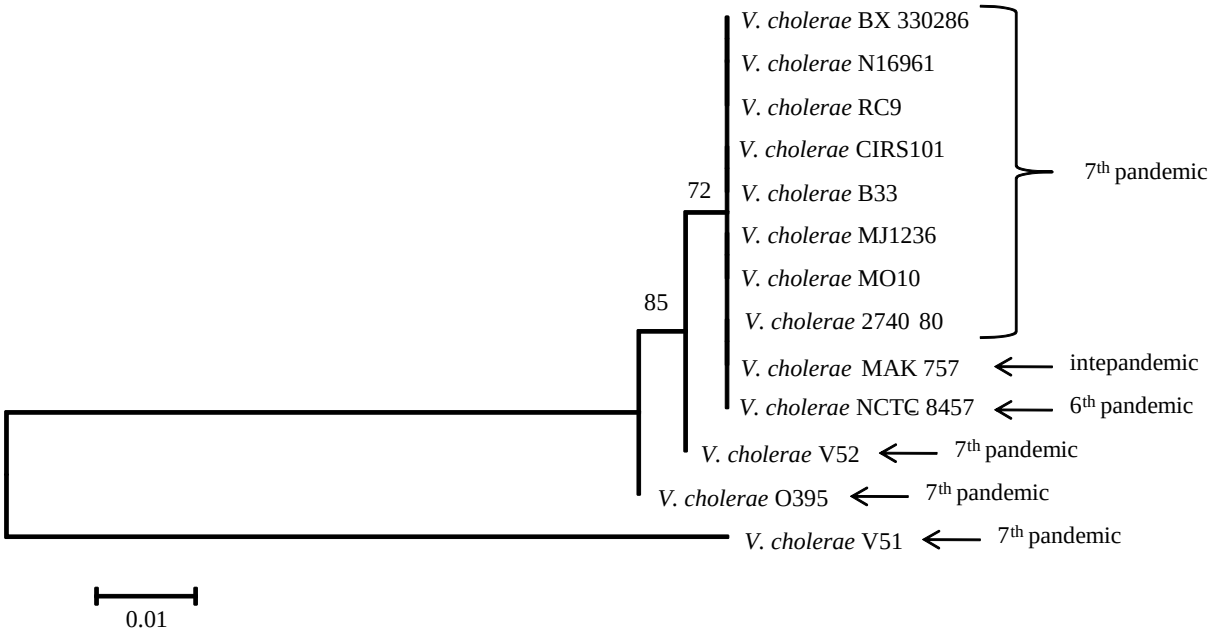



Figure 2.1. Neighbor-Joining tree of *tcpA* from the strains used is this study. Bootstrap values are (%) shown at nodes. Bar is equivalent to 0.01 substitutions per site. Pandemic during which each strain was isolated is listed on right.

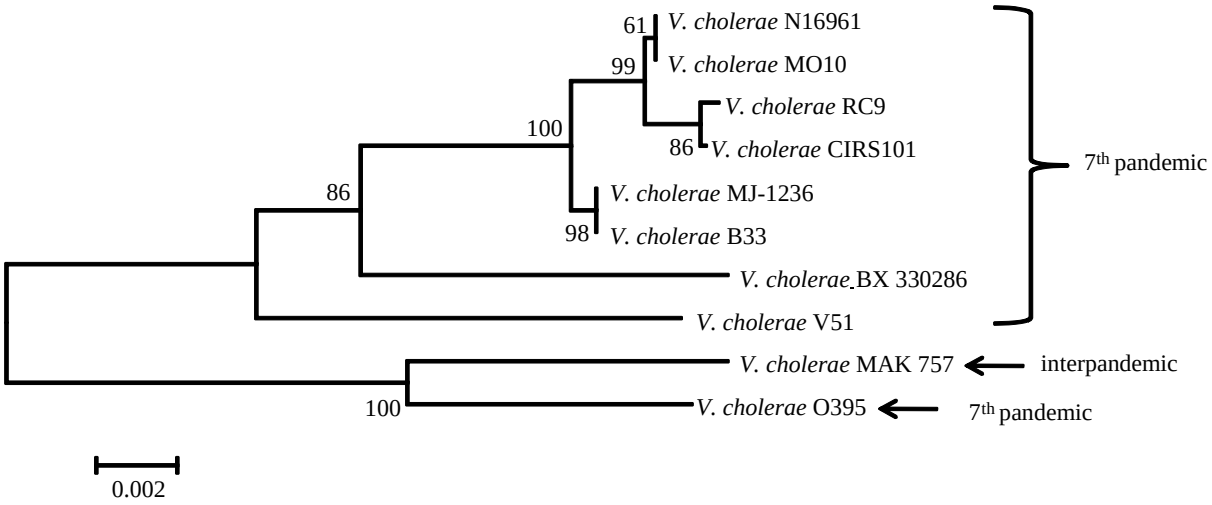



Figure 2.2. Neighbor-Joining tree of a ca. 5.5 kb (concatenated *rstA* to *ctxB* ORFs) homologous region of the CTXΦ from the strains used in this study. Bootstrap values (%) are shown at nodes. Bar is equivalent to 0.002 substitutions per site. Pandemic during which each strain was isolated is listed on right.

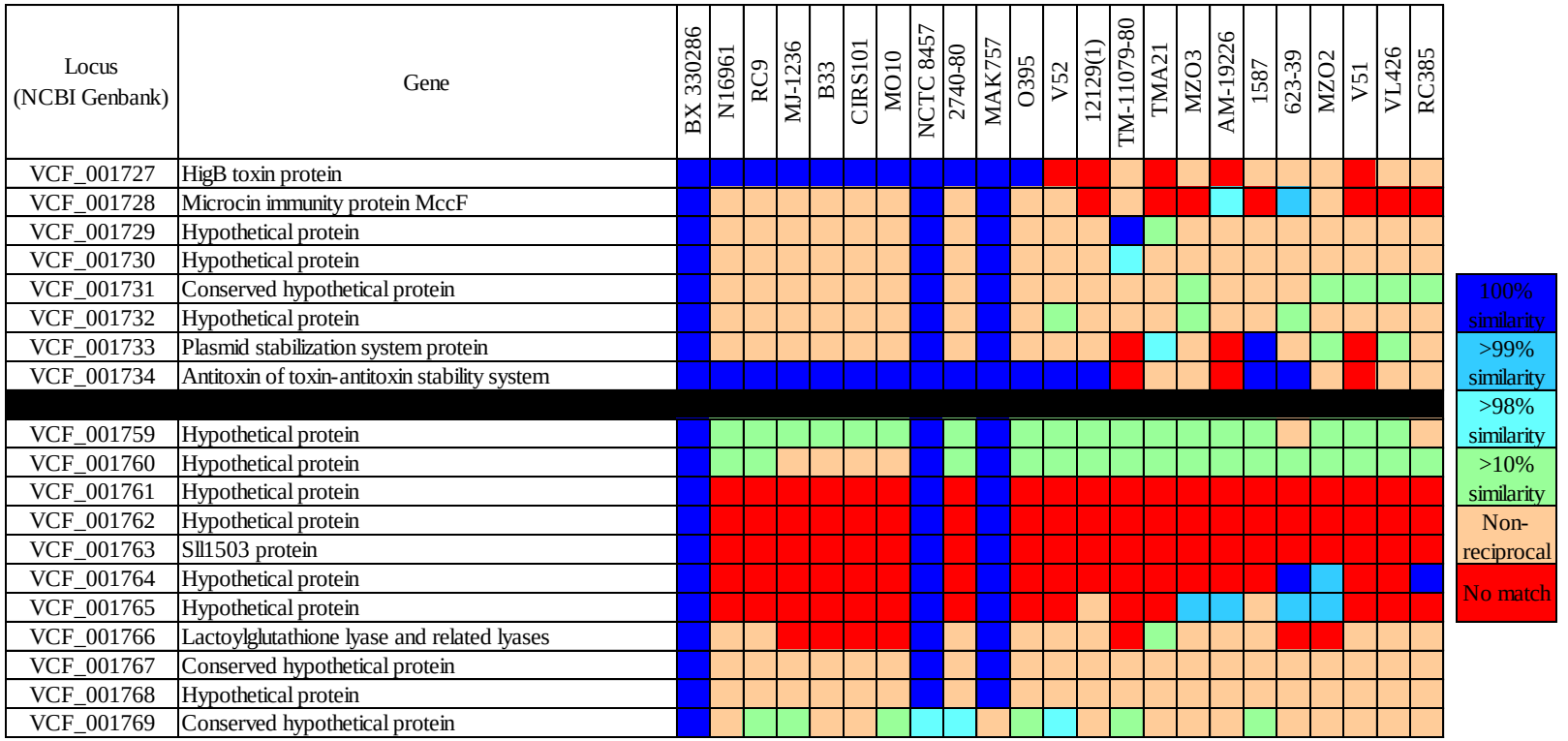



Figure 2.3. Duplicated region of the genome in *V. cholerae* BX 330286, and pre-7$^{th}$ pandemic strains NCTC 8457 and MAK 757 (top). Unique genomic regions in *V. cholerae* BX 330286, NCTC 8457, and MAK 757 (bottom). *V. cholerae* strains are listed in the top row. NCBI Genbank loci are located in column on far left. Nucleotide sequence similarity is noted by color (column on far right).

*Reproduction License Agreement*

Mar 24, 2012

John Wiley and Sons License Terms and Conditions

---

---

This is a License Agreement between Bradd J Haley ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number

2873811048127

License date

Mar 21, 2012

Licensed content publisher

John Wiley and Sons

Licensed content publication

Environmental Microbiology Reports

Licensed content title

The pre-seventh pandemic Vibrio cholerae BX 330286 El Tor genome: evidence for the environment as a genome reservoir

Licensed content author

Bradd J. Haley,Christopher J. Grim,Nur A. Hasan,Elisa Taviani,Jongsik Chun,Thomas S. Brettin,David C. Bruce,Jean F. Challacombe,J. Chris Detter,Cliff S. Han,Anwar Huq,G. Balakrish Nair,Rita R. Colwell

Licensed content date

Feb 1, 2010

Start page

208

End page

216

Type of use

Dissertation/Thesis

Requestor type

Author of this Wiley article

Format

Electronic

Portion

Full article

Will you be translating?

No

Order reference number

BX330286

Total

0.00 USD

Terms and Conditions

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or a society for whom a Wiley Company has exclusive publishing rights in relation to a particular journal (collectively WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your Rightslink account (these are available at any time at http://myaccount.copyright.com)

Terms and Conditions

1. The materials you have requested permission to reproduce (the "Materials") are protected by copyright.

2. You are hereby granted a personal, non-exclusive, non-sublicensable, non-transferable, worldwide, limited license to reproduce the Materials for the purpose specified in the licensing process. This license is for a one-time use only with a maximum distribution equal to the number that you identified in the licensing process. Any form of republication granted by this licence must be completed within two years of the date of the grant of this licence (although copies prepared before may be distributed thereafter). The Materials shall not be used in any other manner or for any other purpose. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Material. Any third party material is expressly excluded from this permission.

3. With respect to the Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Materials without the prior permission of the respective copyright owner. You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Materials, or any of the rights granted to you hereunder to any other person.

4. The Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc or one of its related companies (WILEY) or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or

to the Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto.

5. NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

6. WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

7. You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

8. IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

9. Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

10. The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and

condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

11. This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

12. Any fee required for this permission shall be non-refundable after thirty (30) days from receipt.

13. These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

14. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

15. WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

16. This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

17. This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

Wiley Open Access Terms and Conditions

All research articles published in Wiley Open Access journals are fully open access: immediately freely available to read, download and share. Articles are published under the terms of the Creative Commons Attribution Non Commercial License. which permits

use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes. The license is subject to the Wiley Open Access terms and conditions:
Wiley Open Access articles are protected by copyright and are posted to repositories and websites in accordance with the terms of the Creative Commons Attribution Non Commercial License. At the time of deposit, Wiley Open Access articles include all changes made during peer review, copyediting, and publishing. Repositories and websites that host the article are responsible for incorporating any publisher-supplied amendments or retractions issued subsequently.
Wiley Open Access articles are also available without charge on Wiley's publishing platform, Wiley Online Library or any successor sites.

Use by non-commercial users

For non-commercial and non-promotional purposes individual users may access, download, copy, display and redistribute to colleagues Wiley Open Access articles, as well as adapt, translate, text- and data-mine the content subject to the following conditions:

 The authors' moral rights are not compromised. These rights include the right of "paternity" (also known as "attribution" - the right for the author to be identified as such) and "integrity" (the right for the author not to have the work altered in such a way that the author's reputation or integrity may be impugned).
 Where content in the article is identified as belonging to a third party, it is the obligation of the user to ensure that any reuse complies with the copyright policies of the owner of that content.
 If article content is copied, downloaded or otherwise reused for non-commercial research and education purposes, a link to the appropriate bibliographic citation (authors, journal, article title, volume, issue, page numbers, DOI and the link to the definitive published version on Wiley Online Library) should be maintained. Copyright notices and disclaimers must not be deleted.
 Any translations, for which a prior translation agreement with Wiley has not been agreed, must prominently display the statement: "This is an unofficial translation of an article that appeared in a Wiley publication. The publisher has not endorsed this translation."
Use by commercial "for-profit" organisations

Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee. Commercial purposes include:

 Copying or downloading of articles, or linking to such articles for further redistribution, sale or licensing;
 Copying, downloading or posting by a site or service that incorporates advertising with such content;
 The inclusion or incorporation of article content in other works or services (other than normal quotations with an appropriate citation) that is then available for sale or licensing,

for a fee (for example, a compilation produced for marketing purposes, inclusion in a sales pack)
 Use of article content (other than normal quotations with appropriate citation) by for-profit organisations for promotional purposes
 Linking to article content in e-mails redistributed for promotional, marketing or educational purposes;
 Use for the purposes of monetary reward by means of sale, resale, licence, loan, transfer or other form of commercial exploitation such as marketing products
 Print reprints of Wiley Open Access articles can be purchased from: corporatesales@wiley.com

Other Terms and Conditions:

BY CLICKING ON THE "I AGREE..." BOX, YOU ACKNOWLEDGE THAT YOU HAVE READ AND FULLY UNDERSTAND EACH OF THE SECTIONS OF AND PROVISIONS SET FORTH IN THIS AGREEMENT AND THAT YOU ARE IN AGREEMENT WITH AND ARE WILLING TO ACCEPT ALL OF YOUR OBLIGATIONS AS SET FORTH IN THIS AGREEMENT.

v1.7

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK500744274.
Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

For suggestions or comments regarding this order, contact RightsLink Customer Support: customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

# Chapter 3: Multiomic Analysis of US Gulf Coast Cholera (2010 to 2011) Isolates


## ***Abstract***

Between October, 2010, and May, 2011, twelve cases of cholera, unrelated to a concurrent outbreak in Hispaniola, were recorded and the causative agent, *Vibrio cholerae* serogroup O75, was traced to oysters harvested from Apalachicola Bay, Florida. From the 11 diagnosed cases, eight isolates of *V. cholerae* were isolated (73% of all diagnosed cases of this outbreak) and their genomes were sequenced. Genomic analysis demonstrated the presence of a suite of mobile elements previously shown to be involved in the disease process of cholera and a phylogenomic analysis showed the isolates to be monophyletic with *V. cholerae* V51 serogroup O141, a clinical strain isolated 23 years earlier. Metabolic profiles, virulence gene expression, secretome analyses, and a *Caenorhabditis elegans* model of infection for the isolates were conducted and comparative analyses of the attributes revealed layers of diversity among these otherwise highly similar strains of *V. cholerae*.


## ***Introduction***

*Vibrio cholerae* non-O1/non-O139 are the causative agents of sporadic, yet significant, extraintestinal and gastrointestinal infections globally and it is well established that, like *Salmonella enterica*, all strains of this species are capable of causing human infections and represent a significant global health burden (Ko et al. 1998; Safrin

et al 1998; Shannon and Kimbrough 2006; Lukinmaa et al 2006; Chatterjee et al. 2009). Infection and subsequent illness caused by these organisms are linked to the presence of virulence factors in the core backbone of *V. cholerae* (hemolysins, RTX toxins, lipases) or mobile pathogenicity islands (VPIs-1 and -2, and CTXΦ) that are frequently found in clinical isolates from cholera patients suffering severe rice water diarrhea (Tacket et al. 1998; Vanden Broeck et al., 2007; Almagro-Moreno and Boyd, 2009). Epidemic cholera is typically ascribed to *V. cholerae* serogroup O1 or O139; however, it is now understood that, similar to disease caused by pathogenic *Escherichia coli*, a constellation of virulence factors and host immune and nutritional status, are responsible for the severity and characteristic infections caused by these organisms (Tacket et al. 1998; Vanden Broeck et al., 2007; Manning et al., 2008; Almagro-Moreno and Boyd, 2009; Chun et al., 2009). It is established that those *V. cholerae* which acquire and express genes carried on mobile elements (O-antigens, VPI-1, VPI-2 with or without a type III secretion system, CTXΦ, NAG-ST, etc.) are linked to epidemics of cholera. The scenario of mobile genetic element acquisition has been shown to have occurred within the 7$^{th}$ pandemic and CG-1 and -2 clades (Chun et al., 2009), but occurrence and persistence of such genetic constellations remains underappreciated in *V. cholerae* non-O1/non-O139 (non-CG) lineages. These elements, among many others, can be laterally transferred between strains of the same species or distantly related species in the environment (Meibom et al., 2005; Udden et al., 2008; Boucher et al., 2011) and give rise to virulent strains that potentially can cause epidemics. Further, these elements can be stable in *V. cholerae* non-O1/non-O139 isolates, as in strains of the 7$^{th}$ pandemic clade and persist in these conformations over time, ultimately conserved in the environment.

In developed nations, the leading cause of human disease caused by vibrios is consumption of raw or undercooked seafood, namely shellfish. In the United States, seafood-borne vibrioses have been traced to shellfish harvested from coastal (Atlantic and Pacific) regions, as far north as Alaska, but by far the majority of infections occur in the Gulf of Mexico, where the water is warm and highly productive, both associated with increased *Vibrio* spp. densities as well as increased risk of vibriosis (Hlady and Klontz, 1996; Shapiro et al., 1998; Tamplin, 2001; Lipp et al., 2002; Huq et al., 2005). Recent outbreaks of cholera traced to seafood consumption and many *V. parahaemolyticus* infections and deaths caused by *V. vulnificus* have been reported in this region.

*V. cholerae* O75 serogroup strains have been reported to cause sporadic shellfish-borne cholera cases in the southeastern United States (Tobin-D'Angelo et al., 2008; Onifade et al., 2011). Outbreaks caused by these strains are not continuous as outbreaks in developing nations because sanitation in the United States is such that untreated human waste is not discharged into water used for drinking, recreation, or harvesting of seafood. Further, *V. cholerae* O75 strains have been isolated from environmental waters in the southeastern United States in the absence of reported cholera cases (Tobin-D'Angleo et al., 2008). Here we present results of analysis of eight clinically recovered *V. cholerae* O75 isolates from an indigenous US Gulf Coast cholera outbreak that occurred in October, 2010, and during March and April, 2011 (Onifade et al., 2011).

## ***Materials and Methods***

Clinical *V. cholerae* isolates that were epidemiologically linked to consumption of oysters harvested from the Apalachicola Bay, FL were obtained from the Florida

Department of Health. Virulence assays were conducted following methods standardized for *V. cholerae* (Son and Taylor, 2011). BiOLOG phenotypic microarrays (PM1, PM2A, PM9, and PM10) were conducted following the manufacturers' instructions (BiOLOG, Hayward, CA). *C. elegans* virulence assays was conducted following the methods of Cinar et al. (2010). To infer phylogeny of the mobile elements among strains, neighbor-joining trees were developed, using homologous ORFs between the strains included in this study. Sequences were aligned using CLUSTALW2 (Larkin et al., 2007). Trees were developed using the Kimura-2-parameter of nucleotide substitution (Kimura, 1980). One hundred bootstrap replications were executed for each tree. Phylogenetic estimations were performed using MEGA software (Kumar et al., 2007).

## ***Results and Discussion***

### Phylogenomic Analysis of Florida Outbreak Strains

The eight isolates subjected to analysis in this study have been labeled by number (isolates CP1110, 1111, 1112, 1113, 1114, 1115, 1116 and 1117) and are referred to as the *V. cholerae* FL Group. Using 1104 homologous ORFs the phylogeny of 38 fully and partially sequenced *V. cholerae* strains, including the eight *V. cholerae* FL Group genomes, was inferred (Figure 3.1). Results of the analysis demonstrate that the *V. cholerae* FL Group are monophyletic with *V. cholerae* V51, a clinical *V. cholerae* O141 serogroup strain isolated from a human clinical case in the United States in 1987, suggesting a common ancestor after it had diverged from other *V. cholerae* lineages. This inferred phylogeny showed identical topology when both the Maximum Composite Likelihood and the Kimura-2-Parameter models of nucleotide substitution were used.

From a public health perspective, the results of the analysis demonstrate the group represents a phyletic lineage of non-O1/non-O139 strains that persist in the United States as a cause of morbidity

**Genomic Islands, Pathogenicity Islands, and Virulence Factors**

PCR analyses showed all isolates were CTXΦ$^{Classical}$-positive and encoded the three known RS1Φ elements (RS1$_{environmental}$, RS1$_{Calcutta}$, and RS1$_{El\ Tor}$). The small chromosome CTX-attachment site of all isolates, except CP1112, had an insertion, suggesting that whereas seven of the eight isolates had a CTX insertion or other genomic island at this site, isolate CP1112 does not. The presence of the three RS1 elements has not heretofore been described suggesting these isolates have a unique RS1-CTXΦ arrangement.

The genomes of the eight isolates encoded *Vibrio* pathogenicity island 1 (VPI-1) shown to be responsible for biofilm formation in the intestine and a receptor for CTXΦ phage. VPI-1 of the *V. cholerae* FL Group has highest amino acid identity with *V. cholerae* O395 (98.5%), and TcpA protein (often used as a marker of *V. cholerae* biotype) of *V. cholerae* O395 shared 98.7% amino acid identity with *V. cholerae* FL Group, higher than all other reciprocal comparisons in this study. However, *V. cholerae* O395 encodes 61% of the *V. cholerae* FL Group VPI-1 ORFs found in, determined by reciprocal BLASTP analysis (Table 3.1).

The genomes of all *V. cholerae* FL Group isolates also encoded VPI-2 with type III secretion system (TTSS). The TTSS in these genomes, similar to that of *V. cholerae* V51 and AM-19226, a non-O1 TCP-negative and CTX-negative isolate, has been shown to be essential in the colonization of the infant rabbit intestine and associated with severe

diarrhea in this model (Shin et al., 2011). These VPI-2 elements did not encode a sialidase (also known as neuraminidase), determined by genome sequence analysis and PCR. Sialidase is a sialic acid scavenger that acts on higher order gangliosides in the small intestine and facilitates the interaction of cholera toxin with these gangliosides. Absence of sialidase results in a decrease in binding of cholera toxin to GM1 gangliosides, as demonstrated in the suckling mouse model (Galen et al., 1992). However, all isolates encoded a functional sialic acid catabolism operon within VPI-2, determined be Biolog Phenotypic Microarray assays. The assays demonstrated that all strains utilized sialic acid 3 to 6 times background levels, indicating the operon in these isolates is functional. The mu-like phage region, the most variable region of the canonical VPI-2, is absent in these genomes.

The presence of other genomic islands comprising the *V. cholerae* mobilome described by Chun et al. (2009) was evaluated using BLASTN. Including VPI-1 and 2 and a VSP-II-like element, all genomes encoded sequences with high similarity to GIs-1, 2, 3, 4, 26, 37, 57, 58, and two genomic islands not yet described and designated here as FLGI-1 and FLGI-2 (Figure 3.2) The locus for VSP-I is empty and, therefore, not encoding any inserted genomic island. The VSP-II-like island discovered in the *V. cholerae* FL Group isolates shows variable similarity and conservation with other homologous sequence strings in the *Vibrionaceae* (Figure 3.3). Interestingly, the VSP-II-like element found in the *V. cholerae* FL Group isolates shares 12 of 20 contiguous ORFs (60%) with a contiguous sequence encoded in the genome of *Vibrio corallilyticus* ATCC BAA-450, having an average amino acid identity of 81.2%, suggesting the suite of VSP-II elements is distributed, not only among clinical *V. cholerae* isolates, but also among

those from the environment including non-cholera vibrios. Figure 3.4 depicts a proposed scenario of genomic island insertion and deletion in the *V. cholerae* V51/*V. cholerae* FL Group lineage before and after these two sets of isolates (*V. cholerae* V51 and *V. cholerae* FL Group isolates) diverged from a common ancestor.

**O-Antigen Coding Region**

This region in the *V. cholerae* FL Group is ca. 57.6 kb, with the LPS core region ca. 17.3 kb and O75 specific region ca. 40 kb. Of all *V. cholerae* serogroup data represented in NCBI GenBank, the core OS and the O141-antigen-specific coding regions are most similar to the homologous ORFs of the *V. cholerae* FL Group (Figures 3.5A and 3.5B). Of 56 identified ORFs in this region of the genomes of the *V. cholerae* FL Group isolates, *V. cholerae* V51 shares 20 (35.7%) with 100% nucleotide sequence similarity and 20 with at least 95% nucleotide sequence similarity. When the *wav** (Core OS) and *wbf** (O-antigen specific) clusters were compared, the *V. cholerae* FL Group isolates encode all ORFs in the homologous clusters of *V. cholerae* V51, with high nucleotide sequence similarity in the *wav** cluster and 100% nucleotide similarity in the *wbf** cluster. However, the *wbf** cluster of *V. cholerae* V51 does not encode the *wbfE* ORF of the *V. cholerae* FL Group.

Eight ORFs were found in the O75-antigen coding region of the *V. cholerae* FL Group isolates that have not yet been described in the O-antigen coding regions of other *V. cholerae* genomes and these ORFs may be specific to the O75 antigen (Figures 3.5A and 3.5B). When the O-antigen ORFs of *V. cholerae* V51 serogroup O141 are used as a reference, a similar pattern is observed with the only observed structural differences

being 7 of 10 ORFs missing in the regions homologous to VCV51_0176 to VCV51_0185 in the Florida isolates and 11 of 14 ORFs in *V. cholerae* V51 missing in the homologous region (CP1110_00269 to CP1110_00282). Although, it is well known that this region is a hot-spot for gene transfer, it can be assumed that the O141 and O75 O-antigen coding regions derived from a common source based on the high level of conservation between the two and that the difference between the two clusters arise from substitution of ORFs specific to the O-antigen region. A similar mechanism has been suggested for the relationship between O139 and O22 serogroups (Dumontier and Berche, 1998; Yamasaki et al., 1999). This substitution may have involved a ca. 18.2 kb region in the genomes of the *V. cholerae* FL Group isolates and a ca. 16.2 kb region in *V. cholerae* V51 flanked by the homologs CP1110_00268 (glucose-1-phosphate thymidylyltransferase (EC 2.7.7.24) and CP1110_00286 (lipid carrier:UDP-N-acetylgalactosaminyltransferase). Alternatively, three substitution events involving shorter sequences may have occurred between the flanking regions, indicated by absent ORFs (red squares) in reciprocal comparison. Interestingly, the serogroup with the next highest level of conservation with serogroup O141 and O75 is found in the O139 serogroup isolate *V. cholerae* MO10.

**Open Reading Frame Polymorphisms**

These 8 genomes proved to be highly clonal with very few polymorphisms at each of the conserved gene loci (Table 3.2). A total of 29 ORFs were found to be polymorphic across the entirety of the genomes, with 14 in the core backbone and 15 in mobile genetic elements (2 ORFs in GI-26 [TTSS] and 13 cassettes in the superintegron region). Eleven of the ORFs were annotated as hypothetical proteins with a median size

of 135 bp compared to 865 bp of non-hypothetical protein polymorphic ORFs, 828 bp of all polymorphic ORFs, and 831 bp of all ORFs of the *V. cholerae* FL Group. These data suggest that the polymorphic hypothetical proteins are pseudogenes arising from nonsense mutations.

One polymorphism was a SNP occurring in an ornithine decarboxylase gene (VCA1063); however, a benchtop assay of the decarboxylase activity demonstrated showed the system is functional in all eight isolates. Another SNP was found in strain VC417 in *vpsR* (VC0665), the regulator of *rbmA*, the rugosity and biofilm structure modulator subunit A. Interestingly this isolate formed rugose colonies and produced larger biofilm than all other *V. cholerae* FL Group isolates, suggesting SNP effects expression of the associated genes. One SNP was found in each ORF that was annotated as acetolactate synthase large subunit (VC0031) and dihydroxy-acid dehydratase (VC0028), part of the *ilv* operon involved in isoleucine and valine biosynthesis, shown to confer decreased virulence (Merrell et al., 2002). A SNP was found in an ORF annotated as RTX toxin and related $Ca^{2+}$-binding proteins (VCA0849) a gene discivered to be expressed > 2-fold in *V. cholerae* El Tor biotype isolates than in *V. cholerae* Classical isolates (Beyhan et al., 2006). A SNP was also found in hypothetical protein (VC0874), transcriptionally induced during an infant mouse model infection (Osorio et al., 2005). Taken together, these SNP data suggest there may be differences in virulence expression among these isolates that may result in subtle differences in disease outcomes.

**Virulence Assays**

The eight *V. cholerae* FL Group isolates were further evaluated for hemolysis, motility, and proteolysis, following standard methods for *V. cholerae* (Son and Taylor, 2011) and the results were compared among strains from this outbreak and with other strains whose genomes have been sequenced (Figures 3.6 and 3.7a - c). Biofilm formation was quantified and compared among *V. cholerae* FL Group isolates (Figure 3.6). All strains were hemolytic with wide zones of clearing on blood agar plates, except for strain CP1114 which demonstrated incomplete hemolysis. Interestingly, CP1114 had a wider zone of hemolysis, than all of the other *V. cholerae* FL Group isolates. Isolate CP1115 was significantly more hemolytic than the other isolates, not including CP1114 (analysis of variance, $P < 0.05$). Significant variation in proteolysis was observed among *V. cholerae* FL Group isolates, with isolate CP1114 significantly less proteolytic than other *V. cholerae* FL Group isolates. This low proteolytic activity in CP1114 may result in low hemolysis by this strain as hemolysin A must be processed by an extracellular protease for complete hemolysis to occur. All *V. cholerae* FL Group isolates were highly motile but displayed variability in degree of motility (Figure 3.6). Biofilm formation was variable across the *V. cholerae* FL Group isolates but these isolates did not significantly differ from each other. In a comparison of virulence expression with other *V. cholerae* isolates (clinical and environmental), the *V. cholerae* FL Group does not form a cohesive group, but rather their expression data is interspersed among all isolates (Figs. 3.7a - c). In all three assays isolates of the *V. cholerae* FL Group show some of the highest levels of virulence expression of all *V. cholerae* isolates assayed.

Using the *Caenorhabditis elegans* model of *V. cholerae* infection, which yields data on the strength of the hemolytic activity (*hlyA*) of a strain, (Cinar et al., 2010), nematodes fed three isolates of the *V. cholerae* FL Group (*V. cholerae* CP1112, 1114, and 1115) and the results demonstrated significantly faster die-off than nematodes fed a non-pathogenic *E. coli* control strain but a significantly slower die-off than nematodes fed *V. cholerae* El Tor strain E7946 ($P < 0.05$) (Figure 3.8). All three *V. cholerae* FL Group isolates resulted in similar *C. elegans* survival patterns when compared to each other. However, median survival of worms fed isolates *V. cholerae* CP1112 and CP1115 was 9 days but 11 days for worms fed CP1114, the isolate with incomplete hemolysis, suggesting a somewhat weaker hemolysis in this isolate than the others. Interestingly, the three isolates each caused a *C. elegans* die-off more similar to *V. cholerae* O1 biotype Classical isolates than to El Tor isolates (unpublished data). However, *hlyA* of the *V. cholerae* FL Group do not have the same 11 bp deletion that is the polymorphism linked to the decreased hemolytic activity of *V. cholerae* O1 Classical isolates but rather have higher nucleotide sequence similarity with *V. cholerae* O1 El Tor N16961 than Classical O395 and (98 and 97% respectively). The *hlyA* sequences of the *V. cholerae* FL Group are 100% similar at the nucleotide level and the identified nucleotide sequence polymorphisms among the *V. cholerae* FL Group have not been documented to result in differential hemolytic activity, suggesting genomic rearrangements or methylation patterns, two characteristics not analyzed in this work, may influence virulence expression. Results of *in vitro* and *in vivo* analyses support the conclusion that clonal strains of the same outbreak may potentially result in different levels of virulence expression in the human intestine. However, pathogenic potential of an infectious agent

is only one part of the equation, as host immune status in known to play a significant role in disease outcome.

## ***Conclusions***

There are no known or suggested human carrier(s) of *V. cholerae* O75 and no recent cases of *V. cholerae* O75 infections prior to the outbreak. It is concluded that the outbreak was caused by emergence virulent *V. cholerae* that grew to a density high enough over a region large enough of the environment (not only local one oyster) to cause multiple cases of cholera. A high level of clonality (only 14 polymorphisms in core backbone ORFs) of the isolates, covering 73% of all reported cholera cases from this source, demonstrates that there need not be a human vehicle of *V. cholerae* dispersal into a region prior to a cholera outbreak caused by highly clonal strains, as has been suggested for the ongoing Haitian cholera epidemic.

Because these isolates form a monophyletic lineage with *V. cholerae* V51 serogroup O141, a human clinical isolate from 1987, we postulate the clade may represent a new lineage of cholera-causing isolates, similar to those of the 7$^{th}$ pandemic clade, but smaller geographic scope. Other *V. cholerae* serogroup O141 isolates have been shown to cause significant global disease and many carry the CTX$^{classical}$ prophage (Dalsgaard et al., 2001; Udden et al., 2008) as do the *V. cholerae* FL Group serogroup O75 isolates. To assess this fully, clinical *V. cholerae* non-O1/non-O139 serogroup isolates need to be evaluated in a phylogenomic context. Pathogenic *V. cholerae* cells that cause a significant global disease burden should be characterized in their

phylogenomic context and genomic island constellations rather than simply as being isolates labeled as serogroups O1, O139, or non-O1/non-O139.

Data from this analysis and recently published studies of *V. cholerae* non-O1/non-O139 pathogenicity suggest monitoring the presence of TTSS genes, along with CTXΦ, in shellfish and shellfish harvesting waters should be conducted to assess public health safety of seafood and, thereby, prevent illness and economic loss. In conclusion, multi-dimensional approaches to analyzing bacterial isolates, such as the one presented here, can yield multiple layers of information.

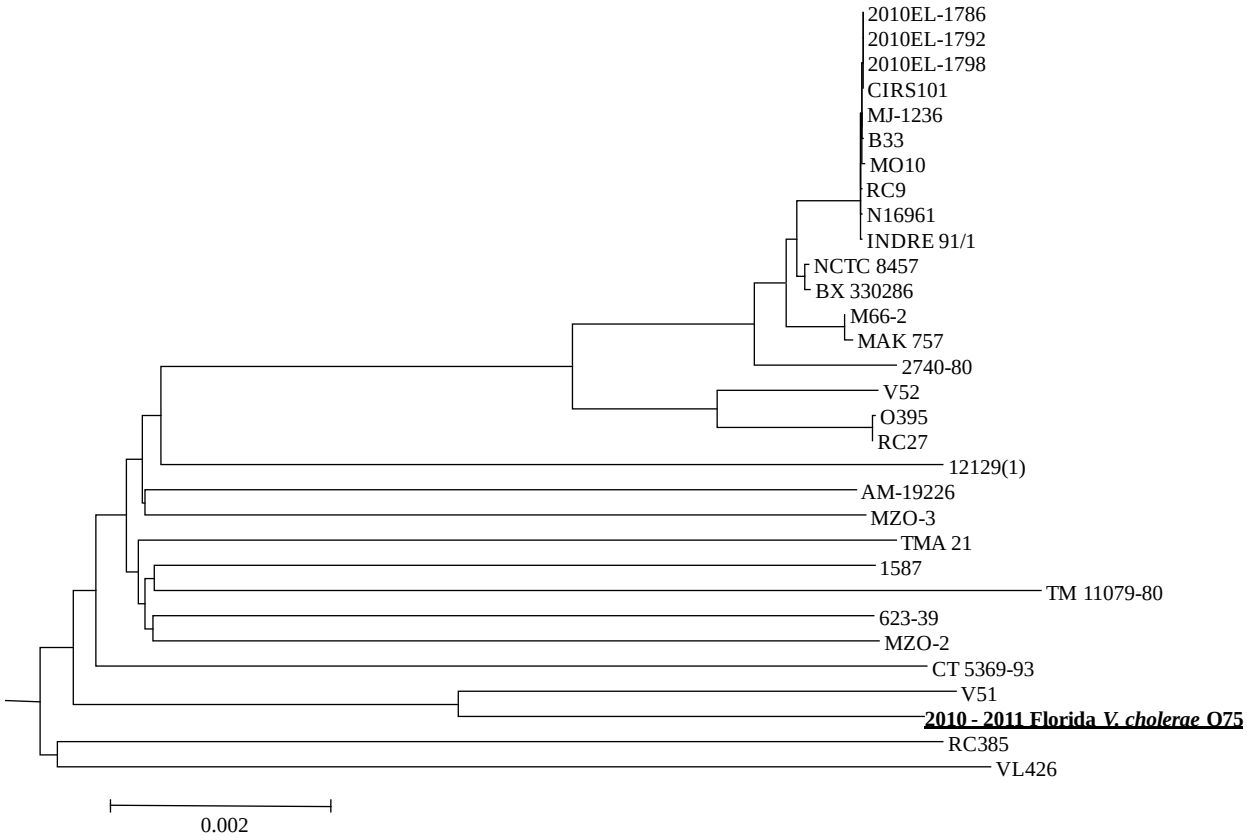



Figure 3.1. Neighbor-Joining tree of 1104 homologous from 38 fully and partially sequenced *V. cholerae* strains. Nucleotide substitution is the Kimura-2-parameter. Bar length = 0.002 nucleotide substitution.

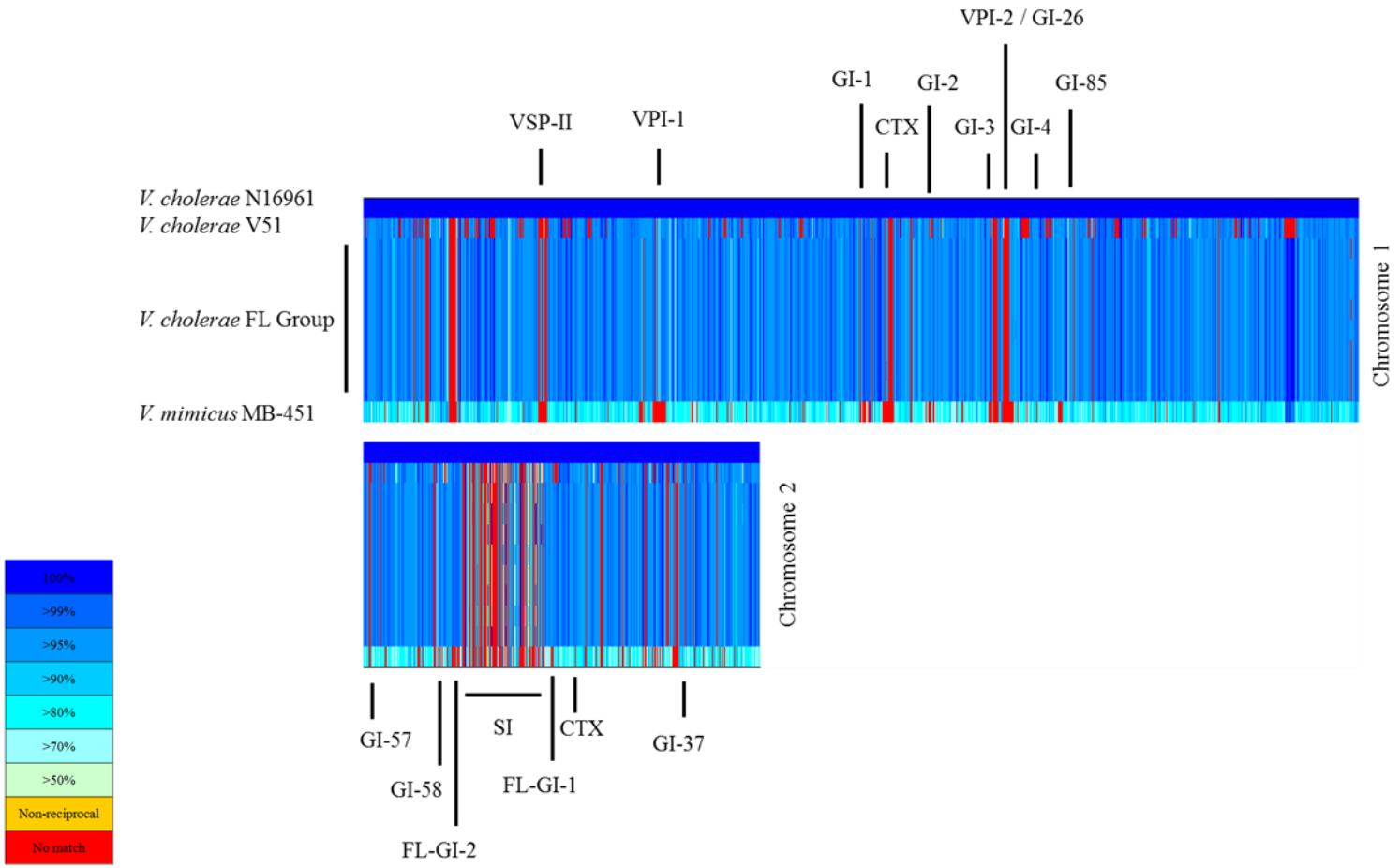



Figure 3.2. BLASTN atlas of *V. cholerae* FL Group genomes, *V. cholerae* V51, and *V. mimicus* MB-451 with *V. cholerae* N16961 as the reference and chromosomal locations of genomic islands in the *V. cholerae* FL Group genomes. Genomic islands with the prefix "GI" are described by Chun et al., (2009). SI = superintegron.

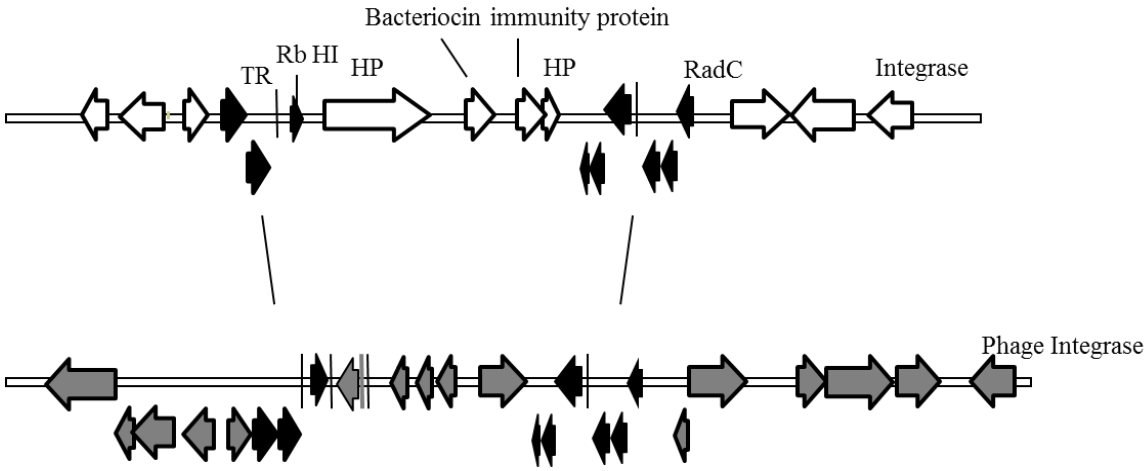



Figure 3.3. Novel VSP-II-like element in all *V. cholerae* FL Group isolates (top). Canonical VSP-II in *V. cholerae* N16961 (bottom). ORFs in black are conserved between the two and white or grey ORFs are not found in the other element. HP = hypothetical protein, TR = transcriptional regulator, Rb HI = Ribonucease HI.

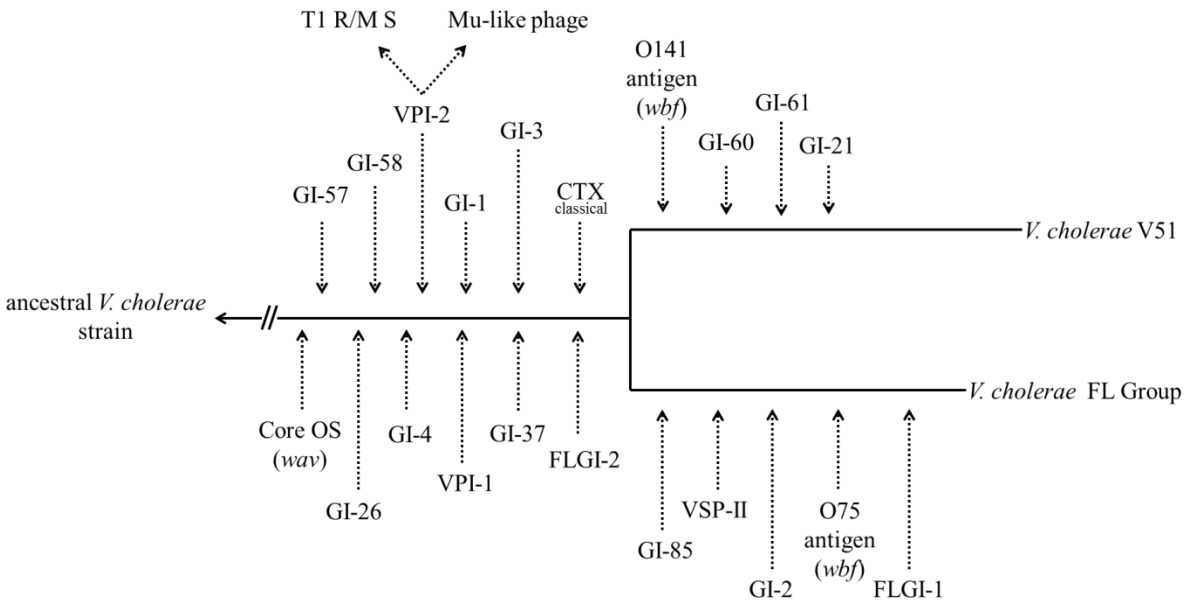



Figure 3.4. Proposed hypothetical insertions of genomic islands in the *V. cholerae* V51/*V. cholerae* FL Group clade.

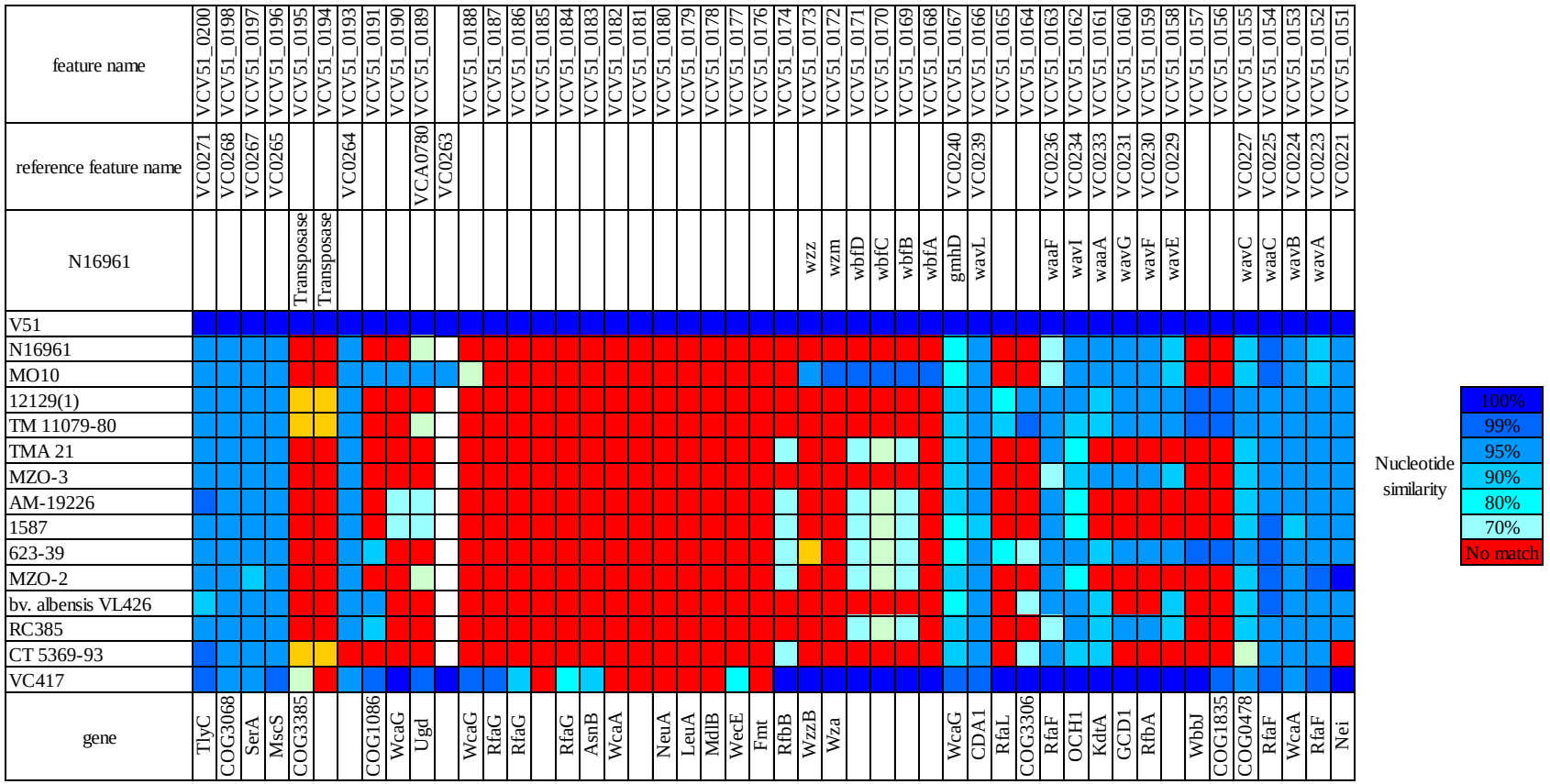



Figure 3.5A. Reciprocal BLASTN analysis of LPS coding region with *V. cholerae* V51 as a reference.

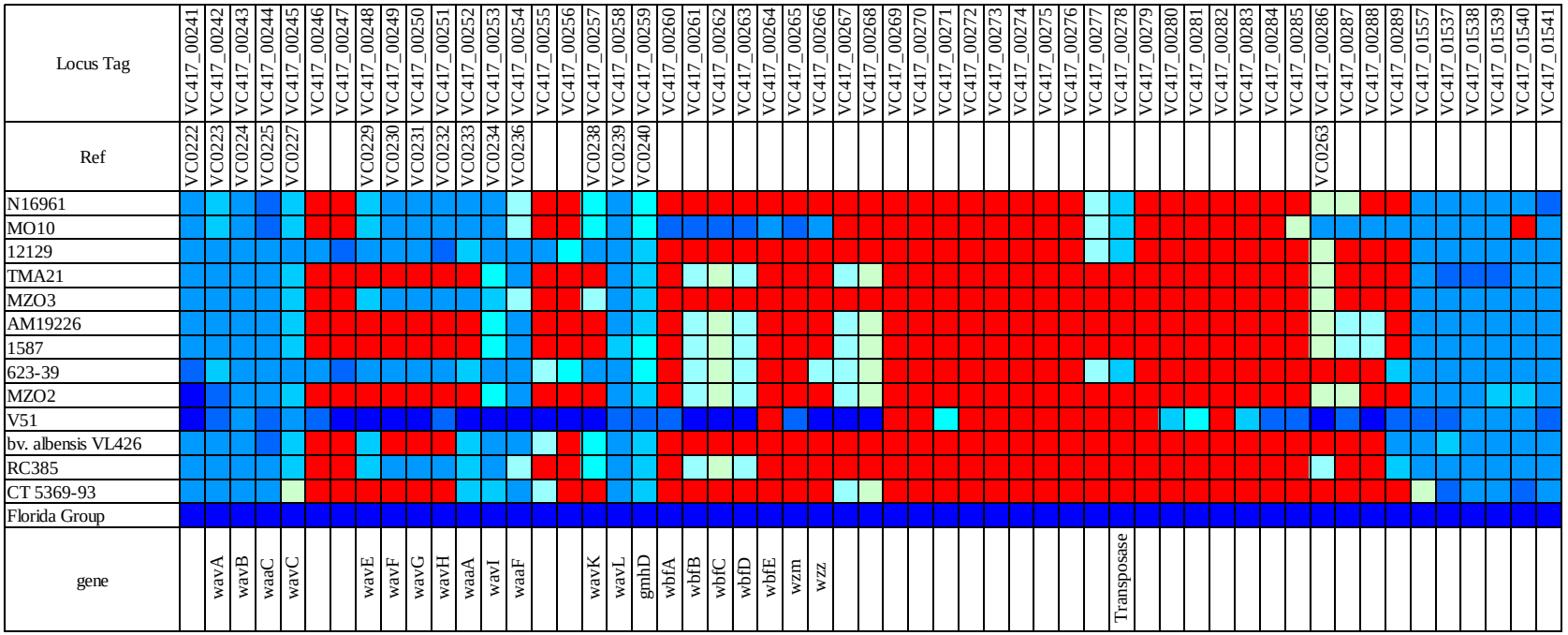



Figure 3.5B. Reciprocal BLASTN analysis of LPS coding region with *V. cholerae* FL Group as a reference.

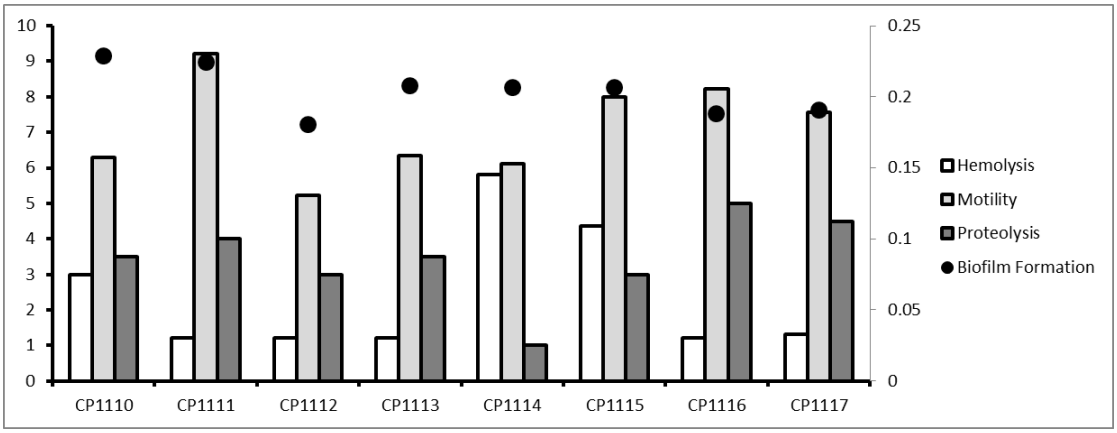



Figure 3.6. Virulence factor expression. The X-axis shows the *V. cholerae* FL Group isolates. The Y-axis on the left shows zones (mm) of hemolysis, proteolysis, and motility on sheep blood agar, milk agar, and motility agar, respectively. The Y-axis on the right shows biofilm formation in $OD_{500}$.

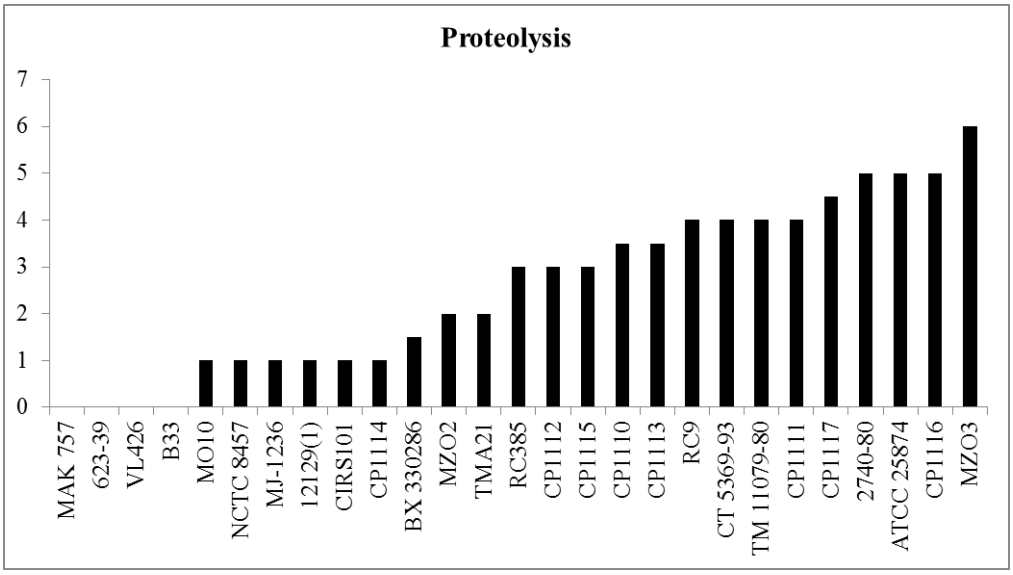



Fig. 3.7a. Virulence factor expression of selected *V. cholerae* isolates whose genomes have been sequenced. The X-axes show the isolates. The Y-axes shows zone of proteolysis (mm) on milk agar.

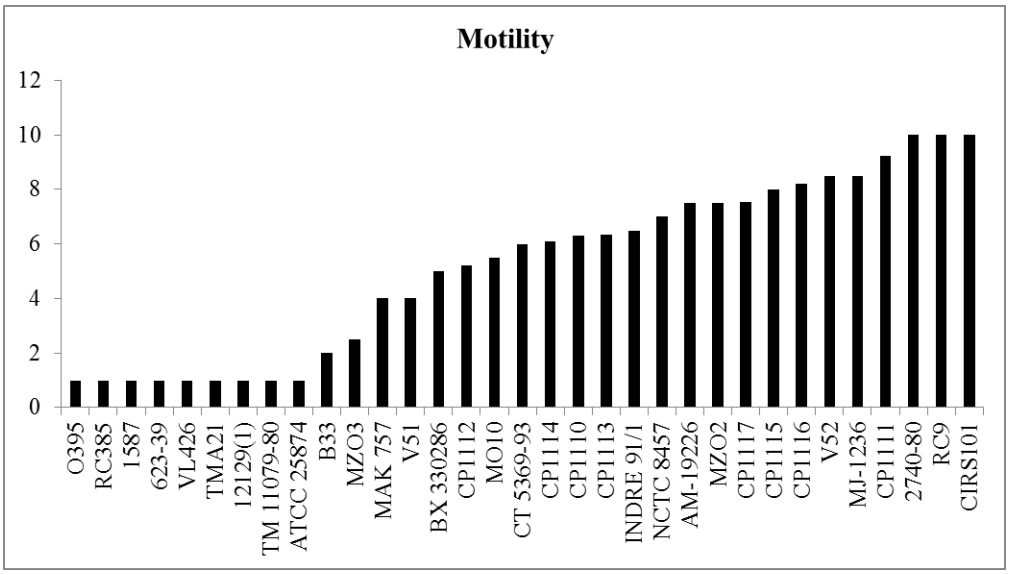
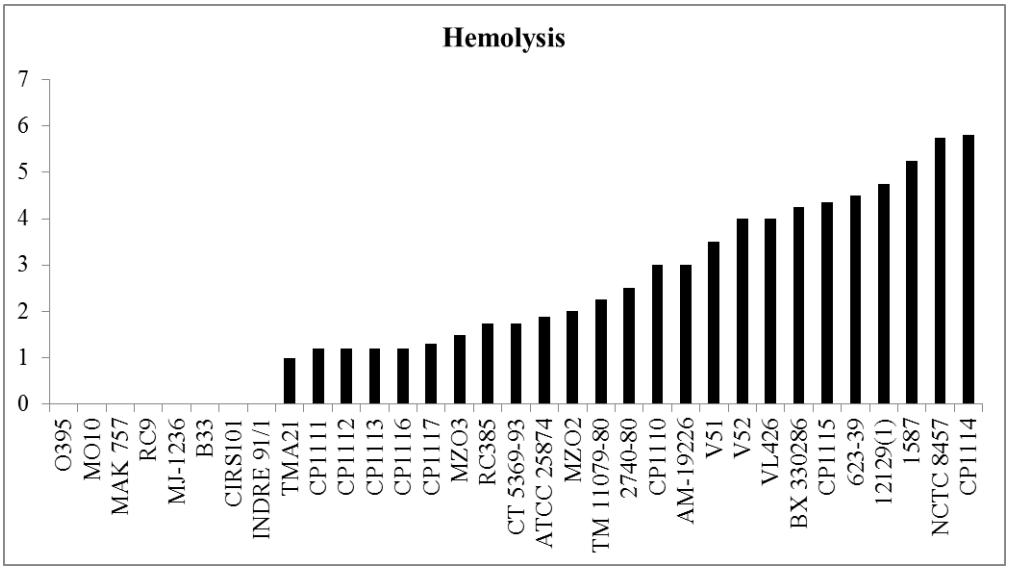



Fig. 3.7b. Virulence factor expression of selected *V. cholerae* isolates whose genomes have been sequenced. The X-axes show the isolates. The Y-axes shows zone of motility (mm) on motility agar.




Figures 3.7c. Virulence factor expression of selected *V. cholerae* isolates whose genomes have been sequenced. The X-axes show the isolates. The Y-axes show zone of hemolysis (mm) on sheep blood agar.

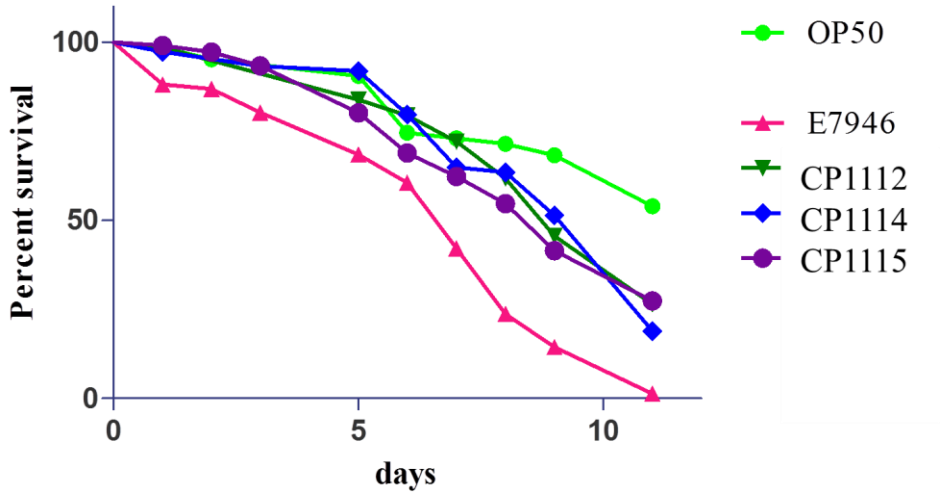



Figure 3.8. Survival curves of *C. elegans* challenged with *V. cholerae* CP1112, CP1114, CP1115, *V. cholerae* El Tor E7946, *Escherichia coli* OP50.

Table 3.1. Amino acid similarity of ORFs in *Vibrio* pathogenicity island 1 (VPI-1) between the *V. cholerae* FL Group and *V. cholerae* strains BX 330286, V51, N16961, and O395.

| *V. cholerae* FL Group | | BX 330286 | V51 | N16961 | O395 |
|---|---|---|---|---|---|
| Length | Annotation | % Similarity | | | |
| 162 | tmRNA-binding protein SmpB | 99.38 | 99.38 | 99.38 | 99.38 |
| 419 | Phage integrase | 97.63 | 100 | 97.63 | 97.63 |
| 53 | Phage integrase | 98.08 | 100 | 98.08 | 80.39 |
| 40 | Phage integrase | 60.53 | 100 | 60.53 | |
| 1591 | Accessory colonization factor AcfD precursor | 96.19 | 98.87 | 96.19 | 96.19 |
| 216 | Accessory colonization factor AcfA | 96.74 | 100 | 96.74 | |
| 303 | TagE protein | 99.01 | 99.34 | 99.01 | 99.01 |
| 254 | Accessory colonization factor AcfC | 98.81 | 98.81 | 98.81 | 99.35 |
| 627 | Accessory colonization factor AcfB | 98.88 | 96.96 | 98.72 | 98.72 |
| 255 | TCP pilin signal peptidase, TcpA processing | 96.39 | 98.39 | 96.39 | 96.39 |
| 277 | TCP pilus virulence regulatory protein ToxT, transcription activator | 99.64 | 98.19 | 99.64 | 99.64 |
| 339 | Toxin co-regulated pilus biosynthesis protein F, putative outer membrane channel for TcpA extrusion | 99.11 | 73.29 | 98.82 | 98.52 |
| 286 | Toxin co-regulated pilus biosynthesis protein E, anchors TcpT to membrane | 98.95 | 98.6 | 98.95 | 98.95 |
| 504 | Toxin co-regulated pilus biosynthesis protein T, putative ATP-binding translocase of TcpA | 99.2 | 99.33 | 99.01 | 99.2 |
| 153 | Toxin co-regulated pilus biosynthesis protein S | 100 | 98.68 | 100 | |
| 274 | Toxin co-regulated pilus biosynthesis protein D | 97.44 | 99.27 | 97.44 | 97.44 |
| 93 | Toxin co-regulated pilus biosynthesis protein R | 100 | 100 | 100 | |
| 490 | Toxin co-regulated pilus biosynthesis protein C, outer membrane protein | 99.39 | 99.18 | 99.39 | 99.8 |
| 127 | Toxin co-regulated pilus biosynthesis protein Q | 97.62 | 98.21 | 97.62 | |
| 431 | Toxin co-regulated pilus biosynthesis protein B | 94.88 | 97.44 | 96.74 | 97.44 |
| 225 | Toxin co-regulated pilin A | 82.14 | 82.14 | 82.14 | 98.66 |
| 101 | Toxin co-regulated pilus biosynthesis protein H, transcriptional activator of ToxT promoter | 97 | 97 | 97 | |
| 222 | Toxin co-regulated pilus biosynthesis protein P, transcriptional activator of ToxT promoter | 97.74 | 98.19 | 97.74 | |
| 621 | Toxin co-regulated pilus biosynthesis protein I, chemoreceptor, negative regulator of TcpA | 98.87 | 99.68 | 98.87 | 98.71 |
| 165 | Thiol peroxidase, Tpx-type (EC 1.11.1.15) | 99.39 | 99.39 | 99.39 | |
| 313 | Putative zinc metalloprotease | 97.44 | 98.72 | 97.44 | 98.4 |
| 418 | inner membrane protein, putative | 94 | 99.53 | 94 | 94.72 |
| 1107 | inner membrane protein, putative | 95.58 | 99.91 | 97.96 | 98.14 |
| 507 | Aldehyde dehydrogenase (EC 1.2.1.3) | 97.43 | 100 | 97.43 | 97.71 |
| 307 | Transposase | 97.22 | 100 | 97.22 | |

Table 3.2. ORFs with polymorphisms among the *V. cholerae* FL group. All ORFs shown represent all ORFs that have a polymorphism between at least two *V. cholerae* FL Group isolates. Strain 417 is used as a reference for a BLASTN comparison.

| VC417 | VC433 | VC447 | VC457 | VC469 | VC472 | VC1205 | Product |
|---|---|---|---|---|---|---|---|
| | | | | | | | RTX protein |
| | | | | | | | hypothetical protein |
| | | | | | | | Transcriptional regulator VpsR |
| | | | | | | | hypothetical protein |
| | | | | | | | FIG01201964: hypothetical protein |
| | | | | | | | Translation elongation factor Tu |
| | | | | | | | RTX toxins and related Ca2+-binding proteins |
| | | | | | | | Acetolactate synthase large subunit (EC 2.2.1.6) |
| | | | | | | | hypothetical protein |
| | | | | | | | hypothetical protein |
| | | | | | | | hypothetical protein |
| | | | | | | | ABC-type polar amino acid transport system2C ATPase component |
| | | | | | | | Lipoprotein VsaC |
| | | | | | | | putative transcriptional activator ToxR |
| | | | | | | | hypothetical protein |
| | | | | | | | RNA-binding protein |
| | | | | | | | Probable type IV pilus assembly FimV-related transmembrane protein |
| | | | | | | | protein of unknown function |
| | | | | | | | Variant SH3 domain protein |
| | | | | | | | conserved hypothetical protein |
| | | | | | | | FIG01201493: hypothetical protein |
| | | | | | | | hypothetical protein |
| | | | | | | | Var1 homologue |
| | | | | | | | Ornithine decarboxylase (EC 4.1.1.17) |
| | | | | | | | Transketolase (EC 2.2.1.1) |
| | | | | | | | Mll3428 protein |

# Chapter 4: *Vibrio* Pathogenicity Island 2 (VPI-2) Diversity


***Abstract***

*Vibrio* pathogenicity island 2 (VPI-2), VPI-2 variants and VPI-2-like elements, including novel variants, were detected in the whole genome sequences of members of the *Vibrionaceae*, including clinical and environmental *Vibrio cholerae*, *Vibrio mimicus*, *Vibrio orientalis*, and *Vibrio* sp. Ex25, a deep-sea hydrothermal vent *Vibrio*. The presence of previously described islands was confirmed in twenty-two *Vibrionaceae* strains chronologically and geographically distince. Regions homologous to that of VPI-2 sialic acid metabolism are present in the backbone of several *Vibrio* spp. Several variants encoded unique phage-like elements and type 3 secretion systems homologous to T3SS-2 in the *V. parahaemolyticus* pathogenicity island. Elements homologous to regions of VPI-2 in non-cholera vibrios were also found. Nucleotide diversity among homologous regions in clinical *V. cholerae* strains was low ($\pi$ = 0.001 to 0.005) which could be interpreted as selection for the canonical variant in the human host, but also found to encode variants of this island were environmental strains isolated from regions where cholera does not occur in pandemic or epidemic form. A multi-step evolutionary history of the elements is proposed to have led to emergence of a mosaic structure for VPI-2 variants, also shown at the nucleotide level as well. The presence of variants in both environment and clinical *V. cholerae* and non-cholera vibrios suggests a dual role for this island, in the environment and human disease. Characteristics of the VPI-2

variants suggest adaptations to different specific niches, rather than deletion outside the human host.

## ***Introduction***

Members of the *Vibrionaceae* are known to harbor genomic islands that allow colonization of novel environmental niches including marine organisms and human hosts. These islands, also known as fitness islands, confer unique traits on the carrier cell, e.g., novel metabolic pathways, xenobiotic degradation, antibiotic resistance, and virulence. *Vibrio cholerae*, the causative agent of cholera causes a severe secretory diarrhea and its genomes encodes islands associated with pathogenicity, such as CTXΦ, VPI-1 and -2, and VSP-I and –II. However, occurrence of these islands and their allelic variations in the environment (outside the human host) suggests a dual role in human infections and as yet undescribed function in the environment (Mukhopadhyay et al., 2001). Recently dual roles for the type three secretion system 2 (TTSS-2) of *V. parahaemolyticus*, which is highly similar in structure to that of *V. cholerae*, (Matz et al., 2011) and the shiga toxin (stx) of *Escherichia coli* have been reported (Steinberg and Levin, 2007).

*Vibrio* pathogenicity island 2 (VPI-2) is a mobile genetic element that is ca. 57 kb encodes phage-like integrase, sialic acid metabolism region, and type-I restriction modification system and is present in epidemic strains of *V. cholerae*. Encoded within the sialic acid metabolism region is *nanH*, a gene responsible for production of neuraminidase (also known as sialidase), a glycoside hydrolase that removes sialic acid from higher order gangliosides, unmasking the GM1 receptor for cholera toxin on human intestinal cells (Jermyn and Boyd, 2002). Genes juxtaposed to *nanH* (*nan-nag* gene

cluster) code for utilization of host-derived sialic acid as a carbon source (Moustafa et al., 2004; Schneider and Parker, 1982). Thus, VPI-2 is associated with pathogenesis and sialic acid metabolism of toxigenic *V. cholerae*. However, 10 of the VPI-2 of 11 *V. cholerae* serogroup O139 isolates examined in the study did not encode *nan-nag* genes or type-I restriction modification system (Jermyn and Boyd, 2005). Both deletions have been linked to reduced fitness and offered as an explanation for this serogroup to no longer be the dominant cause of epidemic cholera in the Indian subcontinent (Jermyn and Boyd, 2005).

Several variants of VPI-2 with major insertions and deletions have been identified (Murphy and Boyd, 2008). In this study, four *V. cholerae* strains encoded two variants of a type three secretion system (*V. cholerae* NRT36S, AM 19226, 623-39, 1587, and V51), two had large deletions (*V. cholerae* MO10 and MAK 757), and one encoded a novel phage-like element (*V. cholerae* V51). The objectives of this study were to evaluate the presence of VPI-2 variants and regions homologous to variants in *Vibrionaceae* genome sequences, determine polymorphisms in homologous ORFs, and apply the data to infer an evolutionary history of the variants.

## ***Methods***

### Strains

A total of 25 genomes, both draft and complete, of members of the *Vibrionaceae* encoding elements considered part of VPI-2 were evaluated for completeness and allelic variants of VPI-2 and VPI-2-like elements. The *V. cholerae*, *Vibrio mimicus*, *Vibrio*

*orientalis*, and *Vibrio* sp. Ex25 strains, isolated from clinical and environmental are listed in Table 4.1.

### Comparative Genomics

Because completeness and quality of nucleotide sequences varied among strains, genome to genome comparison was performed using three different approaches. Nucleotide sequences as whole contigs were directly aligned using the MUMmer program (Kurtz et al., 2004). ORFs of a given pair of genomes were identified and reciprocally compared with each other, using BLASTN (comparison based on ORFs). Finally, a bioinformatic pipeline was developed to identify homologous regions of a given ORF. Initially, a homologous region of an ORF of completely sequenced chromosomes or high quality contigs was identified using BLASTN. This region was then expanded in both directions by 2,000 bp each. Query ORF sequence and target homologous region were aligned using global pairwise alignment and the resultant matched region was extracted and retained as homolog (comparison based on similarity) (Myers and Miller, 1988). Orthologs and paralogs were differentiated by reciprocal comparison. In most cases, comparisons based on ORFs and similarity yielded the same orthologs, but similarity was preferable in the case of draft sequences of low quality, in which sequencing errors, albeit rare, hampered accurate identification of ORFs. By this analysis, contiguous ORFs demonstrating homology with regions of VPI-2 in the canonical variant or previously published variants were considered to be VPI-2 islands or VPI-2-like elements.

Multiple sequence alignments were constructed by aligning homologous nucleotide sequences using ClustalW2. Phylogenetic analysis was conducted using

MEGA 4 program for reconstruction of neighbor-joining trees using the Kimura-2 parameter (K2P) and Jukes-Cantor (JC69) nucleotide substitution models with 100 bootstrap iterations (Tamura et al., 2007). In cases where trees demonstrated congruency, the neighbor-joining tree using the K2P nucleotide substitution model is displayed in the text.

## ***Results and Discussion***

### Gene Structure and Attachment Loci of VPI-2 Varaints

Strains used in this study (Table 4.1) were found to encode several variants of VPI-2 (Figure 4.1). *V. cholerae* MJ-1236 encoded a previously undescribed variant of VPI-2 (Figure 4.1) and the non-cholera vibrios, *Vibrio* sp. Ex25, *V. mimicus* and *V. orientalis* encoded an element homologous to a large region of canonical VPI-2. Several other non-cholera vibrios isolated from both human clinical cases and the environment encoded regions within the genomic backbone similar to the sialic acid metabolism region of the canonical VPI-2. All except three VPI-2 variants were inserted at homologous tRNA-serine loci. The variant found in *Vibrio* sp. Ex25 was encoded in the superintegron region of the large chromosome of its genome (Figure 4.1). Interestingly, this tRNA-serine locus shows 100% nucleotide sequence similarity across the entire sequence with homologous tRNA-serine in all *V. cholerae* genomes, as well as in the *V. parahaemolyticus*, *V. harveyi*, *V. coralliilyticus*, *Vibrio* sp. Ex25, and *Shewanella loihica* genomes.

Based on bidirectional BLASTN pairwise comparisons, variable nucleotide sequence similarity was observed between variants (Figure 4.2). VPI-2 islands of *V.*

*cholerae* O395, RC9, BX 330286, 2740-80, V52, CIRS101, and B33 were similar to canonical VPI-2 of *V. cholerae* N16961, encoding the complete type-1 restriction modification system, sialic acid metabolism, and nearly complete Mu-like phage regions, all with $\geq$ 99% nucleotide similarity across homologous ORFs (Figure 4.2). *V. cholerae* RC9 and N16961, O1 El Tor strains, and CIRS101 and B33, the latter O1 El Tor hybrid strains, were isolated from clinical cases during the 7$^{th}$ pandemic and all are members of the same clonal complex (Chun et al., 2009). *V. cholerae* BX 330286 and 2740-80, environmental strains isolated from Australia in 1986 and from the US Gulf Coast in 1980, respectively, and demonstrating high similarity to pre-7$^{th}$ pandemic El Tor serogroup O1 strains, encode VPI-2 with similarity in gene content and nucleotide sequence with the canonical VPI-2 (Figures 4.1 and 4.2).

Interestingly, clinical isolates *V. cholerae* O395 O1 Classical (isolated in India in 1965), *V. cholerae* NCTC 8457 O1 El Tor (Saudi Arabia in 1910), and *V. cholerae* V52 serogroup O37 (Sudan 1968) isolated preceding the 7$^{th}$ pandemic, demonstrated similarity both in gene content and nucleotide sequence with canonical VPI-2, suggesting that the structure of the island is not strictly biotype or pandemic-specific as other elements such as VSP-I and II (Grim et al., 2009; Taviani et al., 2009) and CTXΦ (Figure 4.2). However, *V. cholerae* O395 VPI-2 carried nucleotide polymorphisms, compared to the 7$^{th}$ pandemic strain (99% sequence similarity, with highest divergence between ORFs of 99.7% in VC1801 for a hypothetical protein), but not with as much divergence as VPI-1 (98.7% sequence similarity, with highest divergence between ORFs of 77.5% in *tcpA*). However, the very small polymorphisms in VPI-2 of this strain most likely have arisen from genetic drift. Interestingly, VPI-2 of *V. cholerae* N16961 demonstrated a higher

average pairwise sequence similarity with *V. cholerae* BX 330286, an environmental strain and progeny of a hypothetical ancestor of the 7$^{th}$ pandemic strains, than *V. cholerae* O395 (99.99% and 99.96% similarity, respectively) supporting the hypothesis that pandemic strains arise in the natural environment.

*V. cholerae* INDRE 91/1, a biotype El Tor strain of the 7$^{th}$ pandemic and the first clinical isolate of a 7$^{th}$ pandemic outbreak in Mexico, revealed deletions in ORFs VC1761, an hypothetical protein proximal to the helicase and VC1801 to VC1803, three hypothetical proteins downstream of the Mu-like phage region, as well as VC1795, a putative transcriptional regulator (Figure 4.2).

*V. cholerae* MJ-1236, a 7$^{th}$ pandemic clinical isolate hybrid strain showed high sequence similarity (> 99%) with canonical VPI-2, but only 89% sequence similarity with the helicase (VC1760) proximal to the phage integrase (Figure 4.2). *V. cholerae* MJ-1236 encodes a 19 kb phage-like region inserted at VC1761 (Figure 4.1). This phage-like element of the MJ-1236 VPI-2 is located at other loci in this genome with undescribed homologs. This phage encodes a type-1 restriction modification system that may be involved in host-addiction.

The VPI-2 present *V. cholerae* non-O1/non-O139 TM 11079-80 seroconverted to O1 via lateral gene transfer. It was originally isolated from sewage and showed divergence across homologous sequences (77 to 100% nucleotide similarity) with an average pairwise sequence identity of 97.8% similar to the average genome-wide similarity between this strain and the 7$^{th}$ pandemic strains (97.9 to 98%) Four hypothetical proteins (VC1788 and VC1801 to VC1803) are deleted in the VPI-2 of this strain.

VPI-2 of *V. cholerae* M66-2 and MAK 757, two O1 El Tor strains isolated during a cholera outbreak in Celebes in 1937, demonstrated high sequence similarity with the canonical VPI-2 variant. However, the island in *V. cholerae* MAK 757 has complete deletion of the Mu-like phage, as well as regions VC1805 to VC1810, and *V. cholerae* M66-2 has a deletion of the Mu-like phage region, but encodes ORFs VC1805 to VC1810. Deletion of these regions may have occurred when the lineage evolved from the progenitor of the PG-1 clade, since other lineages have retained this region.

*V. cholerae* 12129(1) and TMA21 both have a VPI-2 similar to that previously described in *V. cholerae* AM-19226 (Murphy and Boyd, 2008) (Figure 4.1). *V. cholerae* 12129(1) and TMA21 both encode VPI-2 with type three secretion systems (T3SS) inserted in the region where the type-I restriction modification system is located in canonical VPI-2 variants (VC1760 to VC1772) (Figure 4.1). T3SS of *V. cholerae* 12129(1) and TMA21 encode 34 ORFs, all of which are 93 to 100% similar by bidirectional BLASTN analysis to ORFs of the T3SS of *V. cholerae* AM-19226 and V51. However, a 10 kb deletion in the region was detected in the T3SS of *V. cholerae* 12129(1) when compared with *V. cholerae* AM-19226. Direct repeats could not be detected in the flanking regions of these T3SS suggesting that homing between the T3SS and the remainder of the VPI-2 has occurred and that it is a stable element within this island. *V. cholerae* 12129(1), TMA21, V51, and AM-19226 contain 32 ORFs homologous to those of T3SS in *V. parahaemolyticus* while *V. cholerae* 1587 and 623-39 have 16 and 17 ORFs, respectively, homologous to this region (Figure 4.2). Sequence similarity between these ORFs in *V. parahaemolyticus* showed higher sequence similarity to *V. cholerae* 12129(1), V51, AM-19226, TMA21 (82%) than to *V. cholerae* 1587 and

623-39 (67%). For the purpose of the analysis reported here and based on average pairwise identity, VPI-2 of *V. cholerae* 12129(1), TMA21, V51, and AM-19226 comprise group 1 and *V. cholerae* 1587 and 623-39 group 2.

*V. cholerae* V51 the phage integrase of VPI-2 is inserted proximal to a 50 kb phage-like element, previously identified as GI-60 (Chun et al., 2009), that is itself inserted at the tRNA-serine locus (Figure 4.1). This phage encodes several hypothetical proteins and phage assembly proteins and its function is not deducible from its sequence. A conserved hypothetical protein (VCV51_1184) from this sequence shows 68% nucleotide sequence similarity to orf2 of *Shigella flexneri* bacteriophage V. However, other sequences are not homologous to any in the NCBI Genbank database. This phage may have transduced the VPI-2 element into the *V. cholerae* V51 genome, or it may have integrated at the tRNA-serine locus after the VPI-2 was integrated at this locus.

*Vibrio* sp. Ex25 encodes a phage integrase that is not homologous to VC1758 of VPI-2, but shows low amino acid sequence similarity (58%) to phage integrase VC0516 of the *Vibrio* Seventh Pandemic Island- II (VSP-II) *V. cholerae* N16961 by BLASTX. This strain also encodes the canonical type-1 restriction modification system and all ORFs between VC1758 and VC1772, with 86% nucleotide sequence identity. However, ORFs VC1773 to VC1810 are deleted and replaced with a 12 kb region that includes six hypothetical proteins and a phage integrase (Figure 4.1). This 12 kb insert does not have significant match to any sequences in the NCBI Genbank database, excluding a 103 bp region with 78% sequence similarity to a putative transcriptional regulator of *V. parahaemolyticus*. Interestingly, this VPI-2-like element is inserted into the superintegron region of *Vibrio* sp. Ex25, rather than a tRNA-serine locus and this may be

the result of due to the inability of this highly diverged phage integrase to integrate at this tRNA locus. The element is flanked by sequences with similarity to *V. parahaemolyticus* superintegron repeats (91% nucleotide identity) a closely related organism, and ORFs homologous to superintegron cassettes of other vibrios are found both upstream and downstream of this element. These data suggest the type-1 restriction modification system found in VPI-2 may derive from a phage-like element similar to the one found in *Vibrio* sp. Ex25, a deep-sea bacterium.

G+C content varied among the variants and regions of the VPI-2 strains (Table 4.2), ranging from 37% to 47%, indicative of a mosaic structure (Table 4.2). Further, variants encoding T3SS, the regions ranged between 37 and 38% to 43 and 47 G+C% content. Within the T3SS, G+C content of the ORFs were between 29.4% and 45.9% with difference in G+C content of up to 13% between some neighboring ORFs suggesting a mosaic structure for T3SS itself. The sialic acid metabolism region of the canonical variants showed 43% G+C%, content and the homologous region in *V. orientalis* was a 45 G+C%, suggesting sialic acid metabolism has an ancestral association with the type-1 restriction modification system and Mu-like phage regions (43 to 44% and 42 to 41% G+C content, respectively, across all canonical variants).

### Sialic Acid Metabolism ORFs in the Backbone of non-*V. cholerae* Genomes

Several *Vibrio* spp. were found to encode ORFs homologous to the sialic acid metabolism region of VPI-2. All of these regions were associated neither with mobile elements nor integrated at a tRNA locus, suggesting they are anchored in the backbone of these organisms. Interestingly, this region was found in the backbone of *V. mimicus* MB-

451, but not in *V. mimicus* 223. Based on high sequence similarity between sialic metabolism genes anchored in the backbone of *V. mimicus* MB-451 and VPI-2 elements, it is probable that origin of this region in VPI-2 most likely derived from an ancestor of *V. mimicus*/*V. cholerae* group, the region in VPI-2 most likely in the backbone of a non-cholera *Vibrio*, having been mobilized and recombined with elements of VPI-2.

**Nucleotide Diversity**

Results of pairwise BLASTN analysis of VPI-2 variants suggest high nucleotide diversity within this island and from nucleotide diversity ($\pi$) estimated between all variants and groups of variants in each of the three major regions of VPI-2, as well as nucleotide diversity of a concatenated alignment of three ORFs encoded in the backbone of all genomes used in this study; *mdh*, *groEL*, and *toxR* (Table 4.3). Diversity in regions of the VPI-2 variants between groups of strains, all clinical *V. cholerae* strains for example were compared, and further analysis was done comparing the results to nucleotide diversity of the three backbone ORFs.

Interestingly, nucleotide differences within type-1 restriction modification system, sialic acid metabolism, and Mu-like phage regions were smaller than the backbone ORFs for all of the clinical and environmental *V. cholerae* and non-cholera vibrios (Table 4.3). Within the T3SS, nucleotide differences were significantly than the background ORFs, even when the homologous T3SS region of *V. parahaemolyticus* was omitted from the analysis (Table 4.3).

**Evolution of VPI-2 Variants**

All three major regions were independent based on homologs of the regions independent of each other in strains of different species. Based on G+C% content phage integrase, type-1 restriction modification system, and sialic acid metabolism regions (GC% 44 to 43), an ancient association of these regions in an environmental setting or within an unknown donor organism is hypothesized. Phylogenetic analyses of the type-1 restriction modification system, the sialic acid metabolism regions, and both concatenated yielded congruent results (Figures 4.3 to 4.5) supporting co-integration into the *V. cholerae* genome. These data are consistent for both nucleotide substitution matrices used in phylogenetic tree reconstruction. It is further suggested that the Mu-like phage integrated with this element into an ancestral clinical strain after these two regions had been linked as it is only found in canonical strains while the type-1 restriction modification system and sialic acid metabolism regions are found in most variants in this study.

The canonical variant of clinical strains and one environmental strain are highly conserved with few polymorphisms within the 7$^{th}$ pandemic clade and near neighbors. These data coupled with previous knowledge of the evolution of the species and the suggestion that 7$^{th}$ pandemic *V. cholerae* El Tor strains and *V. cholerae* Classical strains evolved from separate lineages (Chun et al., 2009) suggests that a VPI-2 of a Classical strain was laterally transferred to an environmental pre-7$^{th}$ pandemic strain which gave rise to the clinical 7$^{th}$ pandemic strains or that an environmental progenitor of the 7$^{th}$ pandemic clade and the Classical strain received VPI-2 from a common source. This is further supported by an average pairwise similarity analysis of VPI-2 between *V.*

*cholerae* N16961, BX 330286, and O395 which shows that the island in N16961 is more similar to BX 330286 than to O395 (99.99% and 99.96% similar, respectively), while the genome-wide pairwise similarity between these three strains is lower (99.95% and 99.7% similar, respectively). The slight divergence in these The major deletions of the Mu-like phage that are observed in *V. cholerae* MAK 757 and M66-2, both closely related, are inferred to have occurred in this clonal complex after divergence from its phylogenetic near-neighbors.

Based on the highly divergent sequence of the sialic acid metabolism region of *V. mimicus* MB-451 and *V. orientalis* and the phylogenetic analyses of this region, it is apparent that horizontal transfer of this region between *V. mimicus* and *V. cholerae* did not occur recently (Figure 4.4). Most likely, the region evolved with the genomes, as part of the pool of accessory genes for each species, after evolution from an ancestral *Vibrio*.

Phylogenetic analyses of the conserved ORFs of the T3SS infer a closer evolutionary relationship between T3SS of *V. cholerae* 12129(1), V51, AM-19226, TMA21, and *V. parahaemolyticus* than *V. cholerae* 1587 and 623-39 (Fig 4.6). These data are consistent for both nucleotide substitution matrices used in phylogenetic tree reconstruction. The T3SS element of *V. cholerae* 12129(1), TMA21, V51, and AM-19226 is concluded to be more closely linked evolutionarily to T3SS of *V. parahaemolyticus* than to the T3SS of *V. cholerae* 1587 and 623-39, with a T3SS more distantly related to that of *V. parahaemolyticus*.

Although, *V. cholerae* V51 has a T3SS phylogenetically more closely related to those of *V. cholerae* 12129(1), TMA21, and AM-19226, it has a sialic acid metabolism region phylogenetically more closely related to those of *V. cholerae* 1587 and 623-39

(Figure 4.4 and 4.6). Two scenarios may account for this incongruency; *V. cholerae* V51 could have acquired an island homologous to those in 12129(1), TMA21, and AM-19226 followed by convergent evolution with the sialic acid metabolism regions of 1587 and 623-39, or V51 encoded a VPI-2 island similar to 1587 and 623-39 followed by replacement via recombination of the T3SS with one more similar to that found in 12129(1), TMA21, and AM-19226.

## ***Conclusions***

Results of this study demonstrate mosaic architecture for VPI-2, marked by presence and absence of regions, and evidence of horizontal gene transfer between variants. This mosaic is represented at individual ORF and nucleotide levels, demonstrated by variable G+C% content and nucleotide polymorphisms between variants. These data suggest independent elements formed type-1 restriction modification system, sialic acid metabolism, and Mu-like phage regions, while individual ORFs of different origin recombined to form an ancestral VPI-2 which later diversified via genetic drift. These observations are supported by presence of homologous sialic acid metabolism regions not associated with mobility genes in the genomes of two *Vibrionaceae* (*V. mimicus* MB-451 and *V. orientalis* CIP891) and the presence of a type-1 restriction modification system in a deep-sea *Vibrio* (*Vibrio* sp. Ex25) with high similarity to a homologous region of the canonical variant.

The presence of the canonical variant in environmental and clinical strains suggests it plays a significant role in *V. cholerae* pathogenicity as an accessory to cholera toxin, but may not be an essential element in the mechanism of cholera infection, and

furthermore, may also benefit *V. cholerae* outside of the human host. This hypothesis is based on persistence of an element in a genome depending on its utility to that organism (Lawrence, 1999) and further supported by the extremely low-rate of mutation in this island in *V. cholerae* BX 330286, compared to those of clinical strains. Lower rates of polymorphism in VPI-2 in clinical strains between different lineages ($7^{th}$ pandemic El Tor strains and Classical) suggest that this variant may be optimized at the nucleotide level for the human intestine, while variants with a similar gene content but higher levels of polymorphism may be optimized for their respective ecological niches outside of the human intestine as has been demonstrated for VPI-1 (Reguera and Kolter, 2005).

The presence of a T3SS homologous to that of T3SS-2 in the *V. parahaemolyticus* pathogenicity island suggests a common origin of this region in the two species. Interestingly, several strains encode a T3SS phylogenetically more closely related to that of *V. parahaemolyticus* than to each other, suggesting horizontal gene transfer of this region has occurred between the two species. The role of secretion systems in host-pathogen interactions has been demonstrated and the observation that the T6SS of *V. cholerae* is cytotoxic for *Dictyostelium* amoebae (Pukatzki et al., 2006), a predaceous organism found in aquatic environments, suggests T3SS of VPI-2 may play a similar environmental role. Clinical and environmental strains were found in both groups with high levels of polymorphism relative to other VPI-2 regions and the *V. cholerae* backbone, suggesting a "trench warfare" model of host-pathogen interactions for the high levels of polymorphism, preventing the host from adapting to an ubiquitous allele, as has been demonstrated with T3SS of *Pseudomonas syringae* (Guttman et al., 2006; Ma et al.,

2006). Further, hypervariability was observed within this region as evidence of two horizontal transfer events were observed between *V. cholerae* 12129(1) and TMA21.

The presence of a homologous type-1 restriction modification system in a VPI-2 like element in the genome of *Vibrio* sp. Ex25, a deep-sea *Vibrio*, and a homologous sialic acid metabolism region in the backbone of *V. orientalis* suggests there is more than a single function for this region in niches other than the human intestine. Sialic acid catabolism has been shown to enhance colonization of *V. cholerae* in the mouse intestine (Almagro-Moreno and Boyd, 2009a) and to be confined primarily to pathogens and commensals of mammals (Almagro-Moreno and Boyd, 2009b). However, *V. orientalis* was originally isolated from a shrimp. Interestingly, sialic acids are found in some marine organisms (Saito et al., 2001) suggesting this gene cluster found in *V. orientalis* may be involved in scavenging this sialic these organisms. Thus, the sialic acid cluster of *V. orientalis* represents a distantly related homolog of the VPI-2 sialic acid metabolism cluster that is involved in non-human host interactions in the environment. Interestingly, evidence of horizontal transfer of ORFs from this cluster has been reported (de Koning et al., 2000).

New variants are described and compared with the canonical *V. cholerae* N16961 O1 El Tor VPI-2. Novel variants in *V. cholerae* MJ-1236, an O1 El Tor hybrid strain isolated from a clinical case in Bangladesh in 1994, *Vibrio* sp. Ex25, a deep-sea *Vibrio* isolated from an hydrothermal vent are described. Results of this study demonstrate that flexibility is a characteristic of VPI-2, and show that VPI-2 and VPI-2-like elements are present in clinical and environmental *V. cholerae* and non-cholera vibrios.

In conclusion, VPI-2 is a fitness island that can be optimized for different niches, one of these niches being the human intestine with other microniches existing in the aquatic environment. However, optimization for one niche does not rule out utility in another as is seen by the canonical variant in environmental strains. This variant is optimized for functioning in the human intestine but remains stable in the environment, while the T3SS variants may be optimized for interaction with aquatic organisms or defense against predation.

| Strain | Serogroup | Biotype | Geographical origin | Source of isolation | Year of isolation | Accession |
|---|---|---|---|---|---|---|
| *V. cholerae* B33 | O1 Ogawa | El Tor | Beira, Mozambique | clinical | 2004 | ACHZ00000000 |
| *V. cholerae* CIRS101 | O1 Inaba | El Tor | Dhaka, Bangladesh | clinical | 2002 | ACVW00000000 |
| *V. cholerae* 623-39 | non-O1/O139 | | Bangladesh | environmental | 2002 | NZ_AAWG00000000 |
| *V. cholerae* AM-19226 | O39 | | Bangladesh | clinical | 2001 | NZ_AATY01000000 |
| *V. cholerae* MJ-1236 | O1 Inaba | El Tor | Matlab, Bangladesh | clinical | 1994 | CP001485/CP001486 |
| *V. cholerae* BX 330286 | O1 Inaba | El Tor | Australia | environmental | 1986 | ACIA00000000 |
| *V. cholerae* 1587 | O12 | | Lima, Peru | clinical | 1994 | NZ_AAUR01000000 |
| *V. cholerae* MO10 | O139 | | Madras,India | clinical | 1992 | AAKF03000000 |
| *V. cholerae* V51 | O141 | | USA | clinical | 1987 | NZ_AAKI02000000 |
| *V. cholerae* RC9 | O1 Ogawa | El Tor | Kenya | clinical | 1985 | ACHX00000000 |
| *V. cholerae* 12129(1) | O1 Inaba | El Tor | Australia | environmental | 1985 | ACFQ01000000 |
| *V. cholerae* TMA21 | non-O1/O139 | | Brazil | environmental | 1982 | ACHY00000000 |
| *V. cholerae* 2740-80 | O1 Inaba | El Tor | US Gulf Coast | environmental | 1980 | NZ_AAUT01000000 |
| *V. cholerae* TM 11079-80 | O1 Ogawa | El Tor | Brazil | environmental | 1980 | ACHW00000000 |
| *V. cholerae* N16961 | O1 Inaba | El Tor | Bangladesh | clinical | 1975 | NC_002505/NC_002506 |
| *V. cholerae* V52 | O37 | | Sudan | clinical | 1968 | AAKJ02000000 |
| *V. cholerae* O395 | O1 Ogawa | classical | India | clinical | 1965 | NC_009456/NC_009457 |
| *V. cholerae* MAK757 | O1 Ogawa | El Tor | Celebes Islands | clinical | 1937 | NZ_AAUS00000000 |
| *V. cholerae* M66-2 | O1 | | Indonesia | clinical | 1937 | NC_012578 |
| *V. cholerae* NCTC 8457 | O1 Inaba | El Tor | Saudi Arabia | clinical | 1910 | NZ_AAWD01000000 |
| *V. mimicus* MB-451 | | | Bangladesh | clinical | | NZ_ADAF00000000 |
| *V. orientalis* CIP 102891 | | | China | environmental | | NZ_ACZV00000000 |
| *Vibrio* sp. Ex25 | | | deep sea vent, East Pacific Rise | environmental | | NC_013456/NC_013457 |
| *V. parahaemolyticus* RIMD 2210633 [a] | | | Osaka, Japan | clinical | 1996 | NC_004603/NC_004605 |

a = used for comparison purposes and not considered to encode VPI-2

Table 4.1. Strains used in this study. NCBI Genbank accession numbers are listed in the right-hand column.

| Organsim | Genome | Total VPI-2 | V51 Phage | Phage Integrase | MJ-1236 Phage | T3SS | T1RMS | Sialic Acid Metabolism | Ex25 Phage | Mu-like Phage |
|---|---|---|---|---|---|---|---|---|---|---|
| *V. cholerae* N16961 | 47 | 43 | - | 44 | - | - | 43 | 43 | - | 42 |
| *V. cholerae* RC9 | 47 | 43 | - | 44 | - | - | 44 | 43 | - | 41 |
| *V. cholerae* MJ-1236 | 47 | 42 | - | 44 | 40 | - | 44 | 43 | - | 41 |
| *V. cholerae* B33 | 47 | 43 | - | 44 | - | - | 44 | 43 | - | 41 |
| *V. cholerae* CIRS 101 | 47 | 43 | - | 44 | - | - | 43 | 43 | - | . |
| *V. cholerae* INDRE 91/1 | 47 | 43 | - | 44 | - | - | 43 | 43 | - | 41 |
| *V. cholerae* MO10 | 47 | 42 | - | 44 | - | - | - | - | - | |
| *V. cholerae* BX 330286 | 47 | 43 | - | 44 | - | - | 44 | 43 | - | 41 |
| *V. cholerae* 8457 | 47 | 43 | - | 44 | - | - | 43 | 43 | - | 42 |
| *V. cholerae* 274080 | 47 | 43 | - | 44 | - | - | 43 | 43 | - | 41 |
| *V. cholerae* M66-2 | 47 | 42 | - | 44 | - | - | 43 | 43 | - | - |
| *V. cholerae* MAK757 | 47 | 43 | - | 44 | - | - | 43 | 43 | - | - |
| *V. cholerae* O395 | 47 | 43 | - | 44 | - | - | 44 | 43 | - | 41 |
| *V. cholerae* V52 | 47 | 43 | - | 44 | - | - | 43 | 43 | - | 42 |
| *V. cholerae* TM-11079 | 47 | 43 | - | 44 | - | - | 43 | 43 | - | - |
| *V. cholerae* 12129(1) | 47 | 40 | - | 44 | - | 38 | - | 43 | - | - |
| *V. cholerae* TMA21 | 47 | 40 | - | 44 | - | 38 | - | 43 | - | - |
| *V. cholerae* AM-19226 | 47 | 40 | - | 44 | - | 38 | - | 43 | - | - |
| *V. cholerae* 1587 | 47 | 41 | - | 43 | - | 37 | - | 43 | - | - |
| *V. cholerae* 623-39 | 47 | 41 | - | 43 | - | 37 | - | 43 | - | - |
| *V. cholerae* V51 | 47 | 43 | 47 | 43 | - | 38 | - | 43 | - | - |
| *V. mimicus* MB451 | 46 | 44 | - | - | - | - | - | 44 | - | - |
| *Vibrio* sp. Ex25 | 44 | 42 | - | 43 | - | - | 42 | - | 42 | - |
| *V. orientalis* ORI891 | 44 | 45 | - | - | - | - | - | 45 | - | - |
| *V. parahaemolyticus* | 45 | - | - | - | - | 39 | - | - | - | - |
| Mean | | 42 | 47.0 | 43.8 | 40.0 | 37.9 | 43.3 | 43.1 | 42.0 | 41.3 |

Table 4.2. Average G+C% content of the VPI-2 regions.

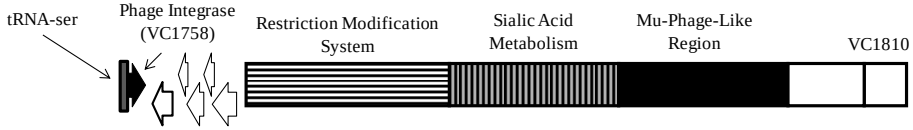
| Region | n | size | Total | | |
|---|---|---|---|---|---|
| | | | S | Θ | π |
| Type-1 restriction modification system | 15 | 16858 | 2763 | 0.054927 | 0.0298 (0.0356 [a]) |
| *V. cholerae* only | 14 | | 1049 | 0.020588 | 0.00985 (0.0029) |
| clinical *V. cholerae* | 11 | | 4 | 0.000086 | 0.00007 (0.0014) |
| 7$^{th}$ pandemic *V. cholerae* | 6 | | 0 | 0 | 0 (0) |
| environmental *V. cholerae* and *Vibrio* sp. Ex25 | 4 | | 2810 | 0.095212 | 0.0963 (0.1302) |
| T3SS [b] | 7 | 5156 | 1897 | 0.157664 | 0.18908 |
| *V. cholerae* only | 6 | | 1624 | 0.144738 | 0.1766 (0.00810) |
| clinical *V. cholerae* and *V. parahaemolyticus* | 4 | | 1883 | 0.209141 | 0.219745 |
| Sialic acid metabolism | 19 | 9754 | 555 | 0.017878 | 0.0133 (0.00691) |
| clinical strains (*V. cholerae* only) | 13 | | 330 | 0.011971 | 0.00897 (0.00566) |
| 7$^{th}$ pandemic *V. cholerae* | 6 | | 0 | 0 | 0 (0) |
| environmental *V. cholerae* | 6 | | 594 | 0.026693 | 0.0242 (0.0111) |
| Mu-like Phage Region | 12 | 4668 | 73 | 0.00522 | 0.002627 (0.00337) |
| clinical *V. cholerae* | 9 | | 0 | 0 | 0 (0.00164) |
| 7$^{th}$ pandemic *V. cholerae* | 6 | | 0 | 0 | 0 (0) |
| environmental *V. cholerae* | 3 | | 87 | 0.012441 | 0.0124 (0.00868) |

a = calculations for concatenated and *mdh* , *groEL* , and *toxR* alignment (3506 bp)
b = T3SS-2 of *V. parahaemolyticus* was included in this analysis
c = n < 4

Table 4.3. Nucleotide diversity of VPI-2 regions between different groups of VPI-2 encoding vibrios.

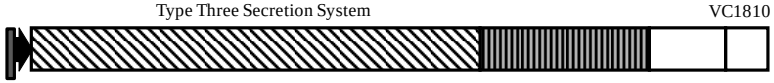
Figure 4.1. Schematic representation of VPI-2 variants found in this study. Homologous regions are coded in each variant by shade and pattern.

(1) Canonical variant:
*V. cholerae* N16861, RC9, B33, CIRS101, BX 330286, INDRE 91/1, 2740-80, NCTC 8457, O395, V52

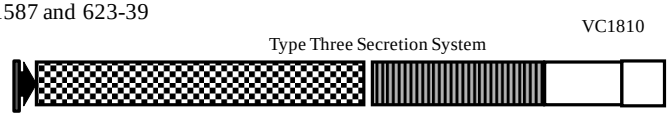



(2) *V. cholerae* MJ-1236


(3) *V. cholerae* AM-19226, 12129(1), and TMA21



(4) *V. cholerae* 1587 and 623-39




(5) *V. cholerae* V51


(6) *V. cholerae* MAK757

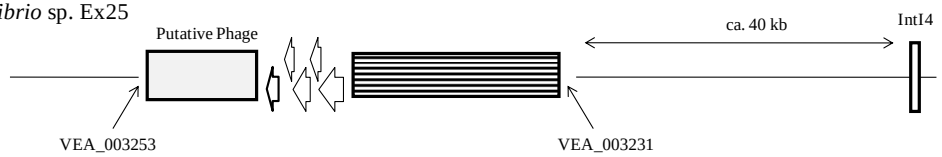

Figure 4.1 cont'd

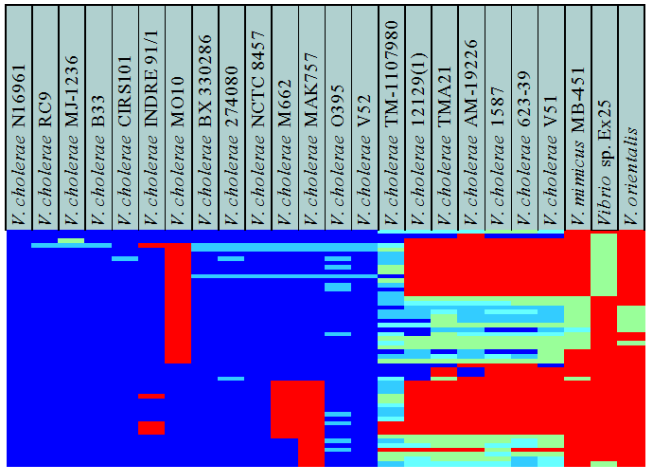
(7) *V. cholerae* MO10


(8) *V. mimicus* MB-451

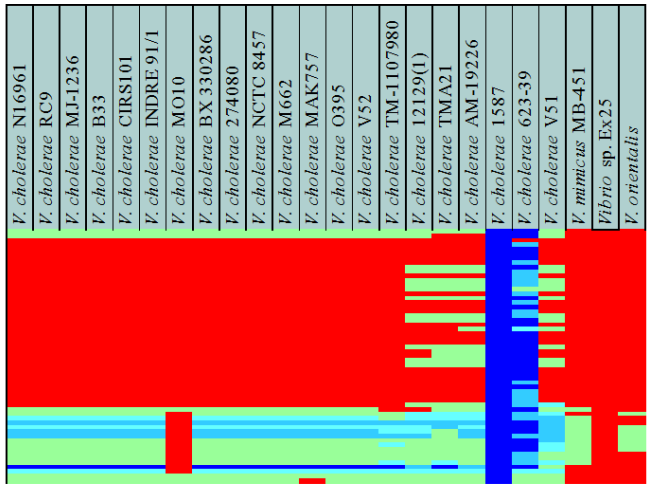
(9) *V. orientalis*

(10) *Vibrio* sp. Ex25

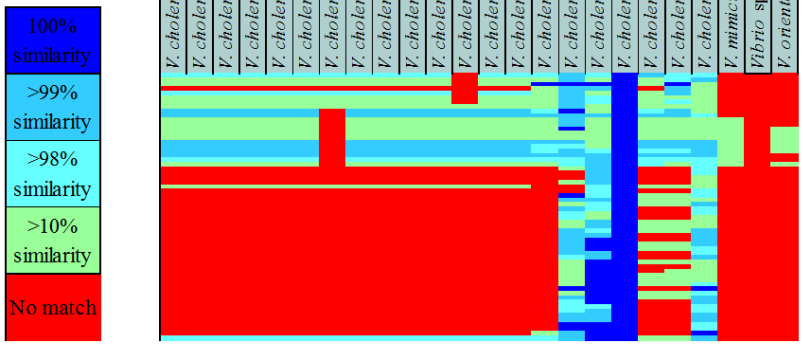




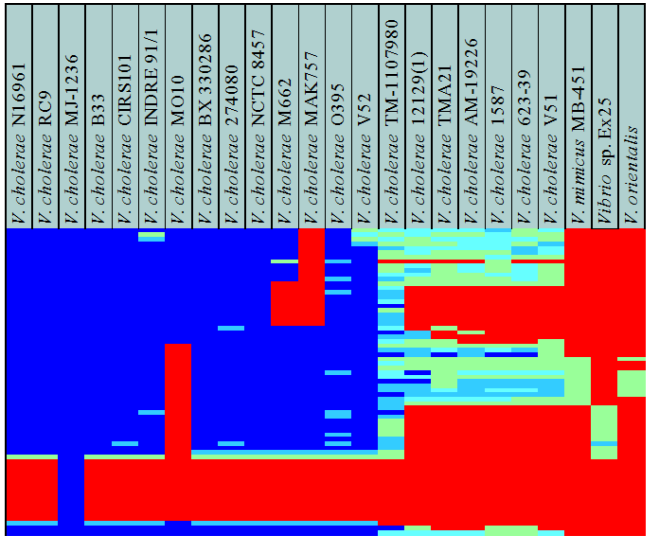

Canonical variant



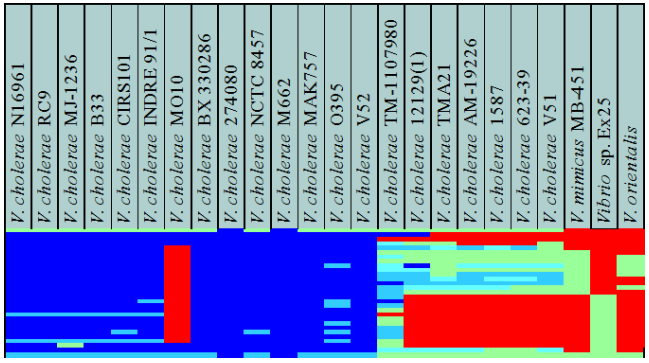

*V. cholerae* 1587 and 623-39 variant



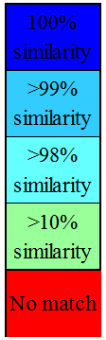

*V. cholerae* AM-19226, V51, TMA21 and 12129(1) variant




*V. cholerae* MJ-1236




*V. cholerae* M662





*Vibrio* sp. Ex25

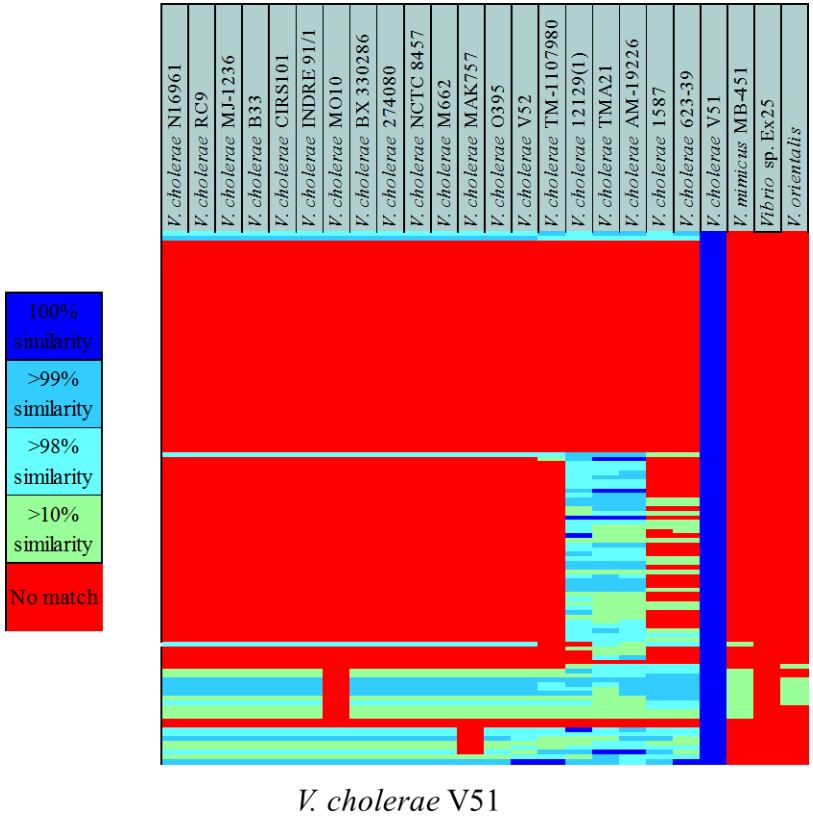



Figure 4.2. Heat map showing results of BLASTN comparison of representative variants against all strains used in this study.

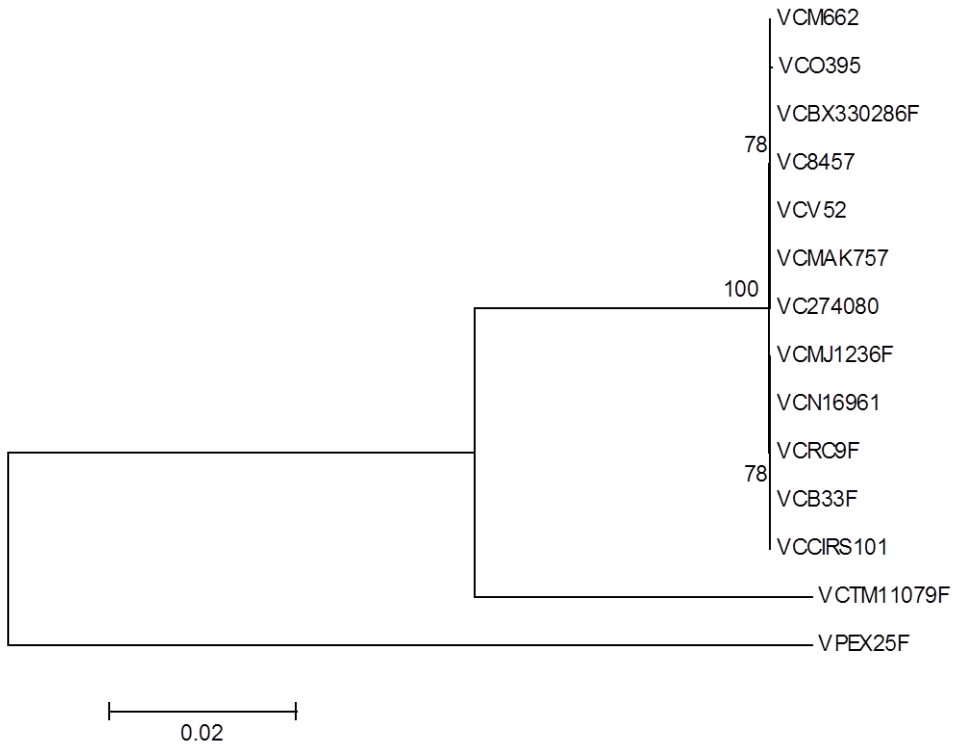



Figure 4.3. Neighbor-Joining tree of the type-1 restriction modification system.

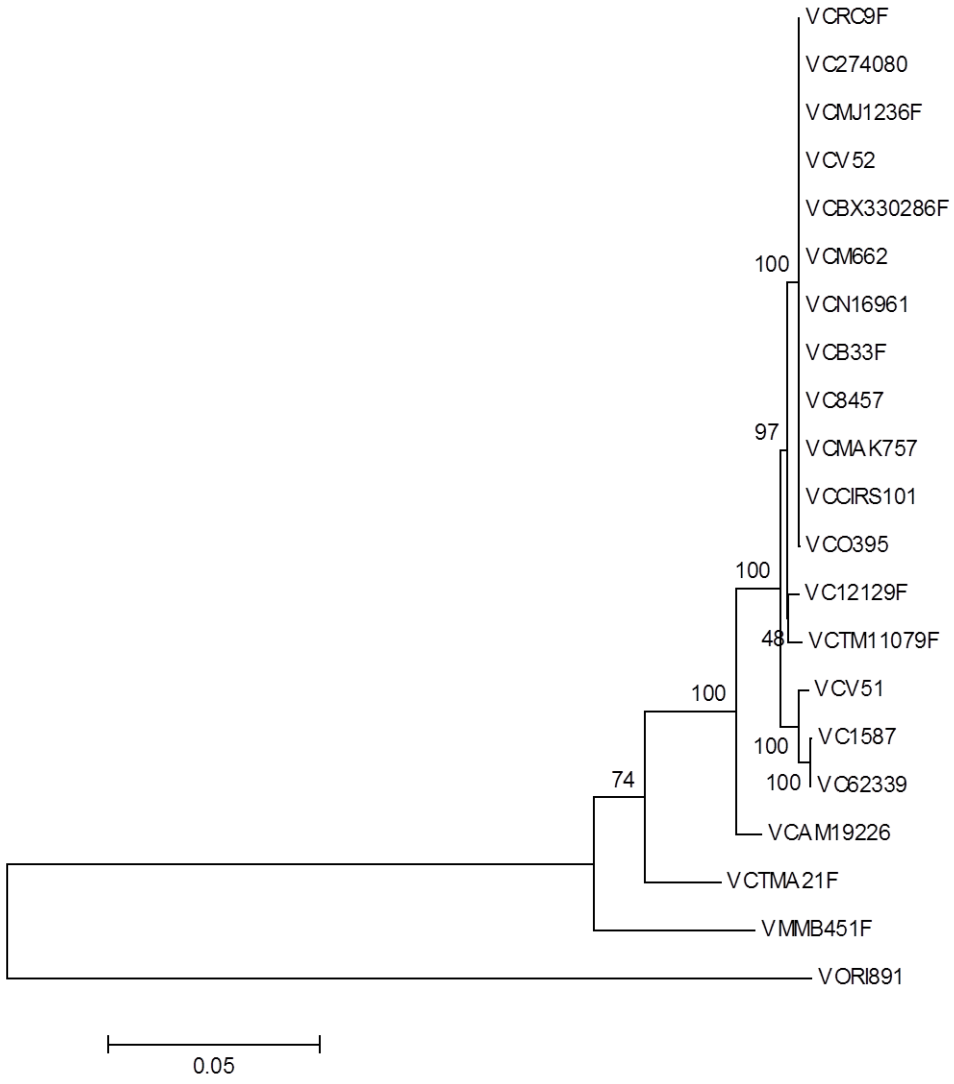



Figure 4.4. Neighbor-Joining tree of the sialic acid metabolism region.

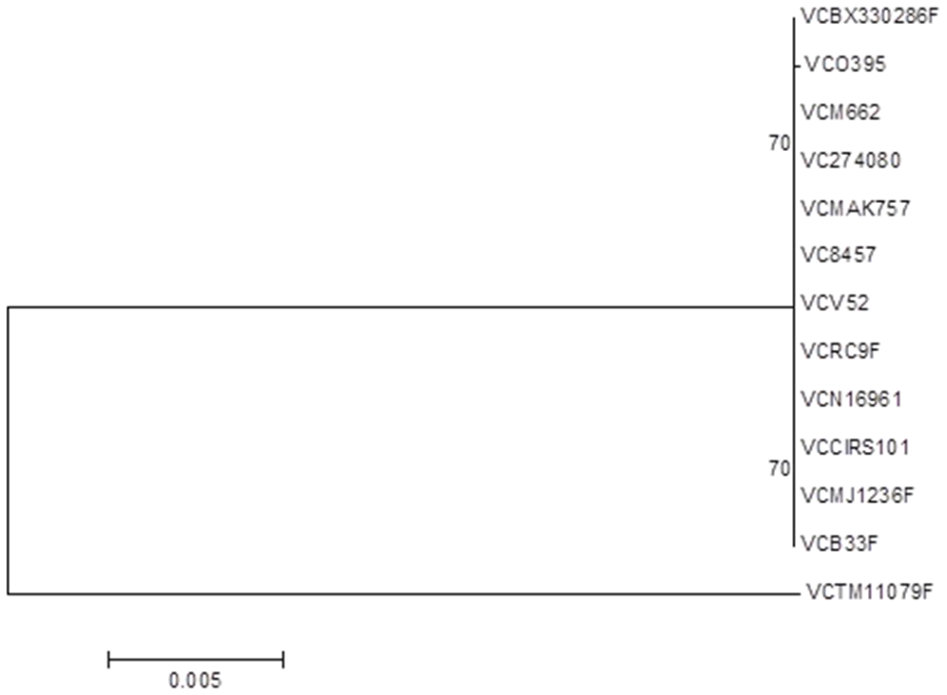



Figure 4.5. Neighbor-Joining tree of the concatenated type-1 restriction modification system and sialic acid metabolism regions.

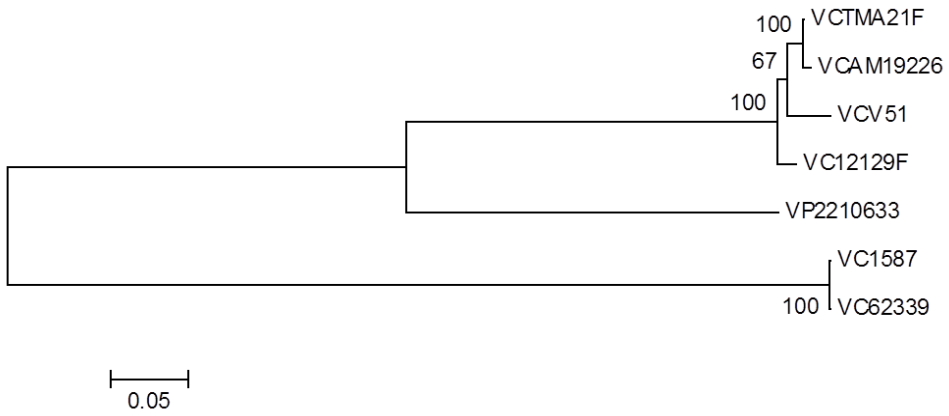



Figure 4.6. Neighbor-Joining tree of the type three secretion system.

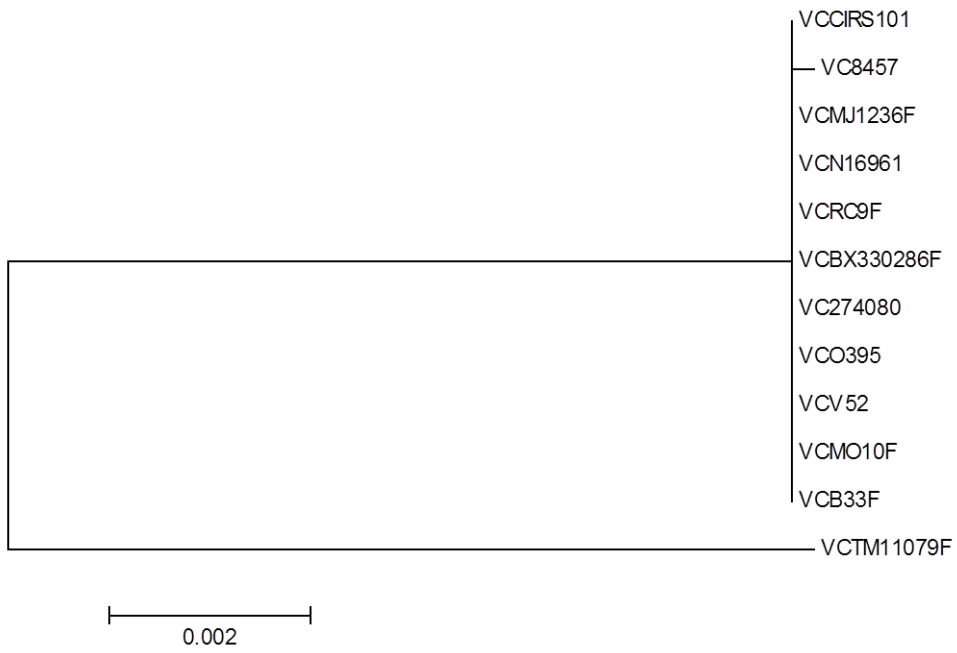



Figure 4.7. . Neighbor-Joining tree of the Mu-like phage region.

# Chapter 5: Comparative Genomics Reveals Evidence of Two Novel *Vibrio* Species Closely Related to *V. cholerae*


***Abstract***

In recent years genome sequencing has been used to characterize new bacterial species, a method of analysis available as a result of improved methodology and reduced cost. Included in a constantly expanding list of *Vibrio* species are several that have been reclassified as novel members of the *Vibrionaceae*. The description of two putative new *Vibrio* species, *Vibrio* sp. RC341 and *Vibrio* sp. RC586 for which we propose the names *V. metecus* and *V. parilis*, respectively, previously characterized as non-toxigenic environmental variants of *V. cholerae* is presented in this study. Based on results of whole-genome average nucleotide identity (ANI), average amino acid identity (AAI), *rpoB* similarity, MLSA, and phylogenetic analysis, the new species are concluded to be phylogenetically closely related to *V. cholerae* and *V. mimicus*. *Vibrio* sp. RC341 and *Vibrio* sp. RC586 demonstrate features characteristic of *V. cholerae* and *V. mimicus*, respectively, on differential and selective media, but their genomes show a 12 to 15% divergence (88 to 85% ANI and 92 to 91% AAI) compared to the sequences of *V. cholerae* and *V. mimicus* genomes (ANI <95% and AAI <96% indicative of separate species). *Vibrio* sp. RC341 and *Vibrio* sp. RC586 share 2104 ORFs (59%) and 2058 ORFs (56%) with the published core genome of *V. cholerae* and 2956 (82%) and 3048 ORFs (84%) with *V. mimicus* MB-451, respectively. The novel species share 2926 ORFs with each other (81% *Vibrio* sp. RC341 and 81% *Vibrio* sp. RC586). Virulence-

associated factors and genomic islands of *V. cholerae* and *V. mimicus,* including VSP-I and II, were found in these environmental *Vibrio* spp. Results of this analysis demonstrate these two environmental vibrios, previously characterized as variant *V. cholerae* strains, are new species which have evolved from ancestral lineages of the *V. cholerae* and *V. mimicus* clade. The presence of conserved integration loci for genomic islands as well as evidence of horizontal gene transfer between these two new species, *V. cholerae,* and *V. mimicus* suggests genomic islands and virulence factors are transferred between these species.

## ***Introduction***

The genus *Vibrio* comprises a diverse group of gamma-proteobacteria autochthonous to the marine, estuarine, and freshwater environment. These bacteria play a role in nutrient cycling, degrade hydrocarbons, and can be devastating pathogens for fish, shellfish, and mammals as well as humans (Pacha, et al. 1969; Kushmaro, et al., 2001). From 1981 to 2009, the number of validly described species within the genus increased from 21 to 100 (Thompson et al., 2004). The most notorious, *V. cholerae,* is the etiological agent of the severe diarrheal disease cholera, endemic in southeast Asia for at least 1,000 years and the cause of seven pandemics since 1817. Shown to be autochthonous to the aquatic environment globally, more than 200 serogroups of *V. cholerae* have been described. Epidemics of cholera are caused by *V. cholerae* O1 and O139, with *V. cholerae* non-O1/non-O139 strains associated with sporadic cholera cases and extraintestinal infections (Huq et al., 1983; Nair et al., 1988). Cholera infections have been ascribed to the presence and expression of virulence genes, e.g., *ctxA, tcpA,*

*tcpP*, and *toxT* (Davis et al., 1981; Shinoda et al., 2004), which are also harbored by toxigenic strains of *V. mimicus*, a phylogenetic near-neighbor of *V. cholerae*. Genomic analyses of *V. cholerae* and *V. mimicus* demonstrated significant similarity, suggesting horizontal exchange of virulence factors, such as CTXΦ and VPIs-1 and -2 (Boyd et al., 2000). Based on results of phylogenetic analyses reported by Thompson et al.(2006), *V. cholerae* and *V. mimicus* should be assigned to separate genera, a taxonomic assignment not yet resolved.

The aims of this study were to describe the genomes of two *Vibrio* strains previously characterized as variant *V. cholerae* by culture-based and molecular methods (Choopun, 2004; Zo, 2005), and compare them to closely related *Vibrio* genomes. Results of this study suggest these two strains represent novel species and demonstrate evidence of horizontal gene transfer with their near-neighbors, *V. cholerae* and *V. mimicus*. We present here the genomic characterization of two new *Vibrio* species, *Vibrio* sp. RC341 (for which we propose the name *Vibrio metecus*) and *Vibrio* sp. RC586 (for which we propose the name *Vibrio parilis*), that share a close phylogenetic and genomic relationship with *V. cholerae* and *V. mimicus*, but are distinct species, based on comparative genomics, average nucleotide identity (ANI), average amino acid identity (AAI), multi-locus sequence analysis (MLSA), and phylogenetic analysis. Also, we present results of a comparative genomic analysis of these two novel species with 22 *V. cholerae*, two *V. mimicus* and one each of *V. vulnificus* and *V. parahaemolyticus* (Table 5.1). The new *Vibrio* species are characterized as *Vibrio* sp. RC341 and *Vibrio* sp. RC586, sharing genes and mobile genetic elements with *V. cholerae* and *V. mimicus*. These data suggest that *Vibrio* sp. RC341 and *Vibrio* sp. RC586 may act as reservoirs of

mobile genetic elements, including virulence islands, for *V. cholerae* and *V. mimicus*, Horizontal gene transfer among these bacteria enables colonization of new niches in the environment, as well as conferring virulence in the human host.

## ***Methods***

### Genome Sequencing

Draft sequences were obtained from a blend of Sanger and 454 sequences and involved paired end Sanger sequencing on 8kb plasmid libraries to 5X coverage, 20X coverage of 454 data, and optional paired end Sanger sequencing on 35kb fosmid libraries to 1-2X coverage (depending on repeat complexity). To finish the genomes, a collection of custom software and targeted reaction types were used. In addition to targeted sequencing strategies, Solexa data in an untargeted strategy were used to improve low quality regions and to assist gap closure. Repeat resolution was performed using in house custom software (Han and Chain, 2006). Targeted finishing reactions included transposon bombs (Goryshin and Reznikoff, 1998), primer walks on clones, primer walks on PCR products, and adapter PCR reactions. Gene-finding and annotation were achieved using an automated annotation server (Aziz et al., 2008). The genomes of these organisms have been deposited in the NCBI Genbank database (accession nos. NZ_ACZT00000000 and NZ_ADBD00000000).

**Comparative Genomics**

Genome to genome comparison was performed using three approaches, since completeness and quality of nucleotide sequences varied from strain to strain in the set examined in this study. Firstly, nucleotide sequences, as whole contigs were directly aligned using the MUMmer program (Kurtz et al., 2004). Secondly, ORFs of a given pair of genomes were reciprocally compared each other, using the BLASTN, BLASTP and TBLASTX programs (ORF-dependent comparison). Thirdly, a bioinformatic pipeline was developed to identify homologous regions of a given query ORF. Initially, a segment on a target contig homologous to a query ORF was identified using the BLASTN program. This potentially homologous region was expanded in both directions by 2,000 bp, after which, nucleotide sequences of the query ORF and selected target homologous region were aligned using a pairwise global alignment algorithm (Myers and Miller, 1988). The resultant matched region in the subject contig was extracted and saved as a homolog (ORF-independent comparison). Orthologs and paralogs were differentiated by reciprocal comparison. In most cases, both ORF-dependent and -independent comparisons yielded the same orthologs, though the ORF-independent method performed better for draft sequences of low quality, in which sequencing errors, albeit rare, hampered identification of correct ORFs.

To determine average nucleotide (ANI) and average amino acid identities (AAI) for the purpose of assigning genetic distances between strains and strains to species groups, a recripocal best match BLASTN analysis was performed for each genome. The average similarity between genomes was measured as the average nucleotide identity (ANI) and average amino acid identity (AAI) of all conserved protein-coding genes,

following the methods of Konstantinidis and Tiedje (2005). By this method, AAI>95% and ANI>94% with >85% of protein-coding genes conserved between the pair of genomes, is judged to correspond to strains of the same species, whereas AAI<95% and ANI <94% and <85% conservation of protein-coding genes indicate different species. Dinucleotide relative abundances were determined for each genome used in this analysis. Genomic dissimilarities between genomes were determined following the methods of Karlin et al. (1997). A multi-locus sequence analysis (MLSA) was determined following standard methods for the *Vibrionaceae* (Thompson et al., 2009). Data for the MLSA were reported as percent similarity between concatenated homologous ORFs for the genomes which encoded these ORFs. These criteria were applied to results of the analyses employed in this study.

**Identification and Annotation of Genomic Islands**

Putative genomic islands (GIs) were defined as a continuous array of five or more ORFs discontinuously distributed among genomes of test strains following the methods of Chun et al (2009). Correct transfer or insertion of GIs was differentiated from deletion events by comparing genome-based phylogenetic trees and complete matrices of pairwise orthologous genes between test strains. Identified GIs were designated, and annotated using the BLASTP search of its member ORFs against the Genbank nr database. Arrays of continuous unique ORFs annotated as encoding phage-related elements and/or transposases were also identified as putative genomic islands. Genomic islets were identified as regions less than 5 ORFs and flanked by genomic island insertion loci Chun

et al. 2009). Putative genomic islands were also investigated using the web-based application IslandViewer (Langille and Brinkman, 2009).

### Phylogenomic Analyses Employing Genome Sequences

A set of orthologues for each ORF of *V. cholerae* N16961 was obtained for different sets of strains, and individually aligned using the CLUSTALW2 program (Larkin et al., 2007). The resultant multiple alignments were concatenated to generate genome scale alignments that were subsequently used to reconstruct the neighbor-joining phylogenetic tree (Saitou and Nei, 1987). The evolutionary model of Kimura was used to generate the distance matrix (Kimura, 1980). The MEGA program was used for phylogenetic analysis (Kumar et al., 2008).

## ***Results and Discussion***

### Strains

The two strains analyzed in this study, *Vibrio* sp. RC341 and *Vibrio* sp. RC586, were isolated from water samples from the Chesapeake Bay, MD in 1999. *Vibrio* sp. RC341 and *Vibrio* sp. RC586 were presumptively classified as variant *V. cholerae* (Choopun, 2004; Zo, 2005), based on amplification of the 16S-23S intergenic spacer unit (Chun et al., 1999) and similarity to the 16S ribosomal RNA of *V. cholerae*. *Vibrio* sp. RC341 appears yellow *V. cholerae*-like cells and *Vibrio* sp. RC586 appears as green *V. mimicus*-like cells on TCBS agar. Both strains were typeable with *V. cholerae* antisera, *Vibrio* sp.

RC586 as serogroup O133 and *Vibrio* sp. RC341 as serogroup O153 (Choopun, 2004; Zo, 2005).

**General Genome Overview**

The genomes of *Vibrio* sp. RC341 and *Vibrio* sp. RC586 span 28 and 16 contigs, respectively, and putatively encode 3574 and 3592 ORFs totaling 4,008,705 bp and 4,082,591 bp, respectively. *Vibrio* sp. RC341 encodes 91 RNAs, 71 of which are tRNAs. *Vibrio* sp. RC586 encodes 115 RNAs, 91 of which are tRNAs. The %GC content of each genome is ca. 46%, while the %GC content of *V. cholerae* strains is 47%. *Vibrio* sp. RC341 encodes 681 hypothetical proteins (19% of total ORFs) and *Vibrio* sp. RC586 encodes 719 hypothetical proteins (19.6% of total ORFs) determined by subsystem annotation. Twenty-four of these hypothetical proteins of *Vibrio* sp. RC586 and 48 of *Vibrio* sp. RC341 showed no homology to any of the sequences in the NCBI database.

Both genomes putatively encode two chromosomes, determined by comparing both chromosomes of *V. cholerae* N16961 to draft genome sequences of *Vibrio* sp. RC341 and *Vibrio* sp. RC586 using the MUMmer program (Kurtz et al., 2004) (see additional figures 5.1 and 5.2). The smaller chromosome of *Vibrio* sp. RC586 putatively encodes 1035 predicted ORFs, totaling approximately 1,155,676 bp. By this method, 951 ORFs were detected in *Vibrio* sp. RC341 totaling 987,354 bp. The smaller size of the second chromosome of *Vibrio* sp. RC341 can be attributed to low-quality coverage of this genome or uncaptured gaps. Both putative small chromosomes of the two species encode a superintegron region homologous to that of *V. cholerae*. The superintegron region of *Vibrio* sp. RC586 is ca. 93.6 kb, putatively encodes 96 ORFs, 66

(69%) of which are hypothetical proteins and the superintegron region of *Vibrio* sp. RC341 is ca. 68.6 kb, putatively encodes 66 ORFs, only 17 (26%) of which are hypothetical proteins. Interestingly, the superintegron of *Vibrio* sp. RC341 encodes several membrane bound proteins suggesting their role in the interaction with the extracellular environment.

**Genome Comparisons**

The genomes of *Vibrio* sp. RC341 and *Vibrio* sp. RC586 were compared with each other and to 22 *V. cholerae*, two *V. mimicus*, one *V. vulnificus* and one *V. parahaemolyticus* genome sequences by pairwise reciprocal BLAST analysis. *Vibrio* sp. RC341 and *Vibrio* sp. RC586 share 2104 non-duplicated ORFs (58% of the *Vibrio* sp. RC341 protein-coding genome) and 2058 non-duplicated ORFs (57% of the *Vibrio* sp. RC586 protein-coding genome) with 22 *V. cholerae* strains. Chun et al. (2009) determined that the current *V. cholerae* core contains 2432 ORFs, indicating a dramatic difference in number of core genes between *Vibrio* sp. RC341/RC586 and *V. cholerae* core genomes. *Vibrio* sp. RC341 shares 2613 ORFs with *V. cholerae* N16961 (73% of *V.* sp. RC341), and *Vibrio* sp. RC586 shares 2581 ORFs with *V. cholerae* N16961 (71% of *Vibrio* sp. RC586) (Figure 5.1). *Vibrio* sp. RC341 shares 2956 ORFs with *V. mimicus* MB-451 (82% of *Vibrio* sp. RC341), and *Vibrio* sp. RC586 shares 3048 ORFs with *V. mimicus* MB-451 (84% of *Vibrio* sp. RC586) (Figure 5.1). *Vibrio* sp. RC341 and *Vibrio* sp. RC586 share 2926 ORFs with each other (81% of ORFs in both genomes) (Figure 5.1).

To determine average nucleotide identity (ANI) and average amino acid identity (AAI) between each genome, the average pairwise similarity between ORFs conserved

between the compared genomes was calculated, following methods of Konstantinidis and Tiedje (2006) and Konstantinidis et al. (2005). In this approach, two genomes with an ANI >95% and AAI >96% belong to the same species, while those with ANI and AAI below these thresholds, comprise separate species (Konstantinidis and Tiedje, 2006) and Konstantinidis et al. 2005). The ANI and AAI between *Vibrio* sp. RC586 and *Vibrio* sp. RC341 was 85 and 92%, respectively (see additional figures 5.3, 5.4, and 5.5). The ANIs between *Vibrio* sp. RC586 and individual *V. cholerae* ranged between 84 and 86%, while the ANI between *Vibrio* sp. RC341 and *V. cholerae* ranged between 85 and 86% (see additional figures 5.3, 5.4, and 5.5). The AAIs between *Vibrio* sp. RC341 and individual *V. cholerae* genomes and *Vibrio* sp. RC341 and *V. cholerae* were 92% in all comparisons (data not shown). The ANIs between *Vibrio* sp. RC586 and *V. mimicus* MB-451 and VM223 were 88% and 87%, respectively, and 86% for *Vibrio* sp. RC341 and both *V. mimicus* genomes (see additional figures 5.3, 5.4, and 5.5). The AAI between *Vibrio* sp. RC341 and *V. mimicus* strains MB-451 and VM223 was 92% in both comparisons, while the AAI between *Vibrio* sp. RC586 and both *V. mimicus* strains was 93% (data not shown).

The *V. cholerae* genomes had ANI >95% and AAI >96% and both *V. mimicus* strains a 98% ANI and AAI. The ANI for all *V. cholerae* and both *V. mimicus* strains was 86%. Based on these data, it is concluded that *Vibrio* sp. RC341 and *Vibrio* sp. RC586 are, indeed, separate species, genetically distinct from *V. mimicus* and *V. cholerae* and from each other. Strains of interspecies comparisons shared <95% ANI and <96% AAI with members of other species included in this study, the threshold for species demarcation (Konstantinidis and Tiedje, 2006) and Konstantinidis et al. 2005), as applied

to *Vibrio*, *Burkholderia*, *Escherichia*, *Salmonella*, and *Shewanella* spp. (Konstantinidis et al., 2006; Thompson et al., 2009; Vanlaere et al., 2009). When *Vibrio* sp. RC341 and *Vibrio* sp. RC586 were compared with the more distantly related *V. vulnificus* and *V. parahaemolyticus*, *Vibrio* sp. RC586 showed 72 and 73% ANI and 73 and 73% AAI, respectively and *Vibrio* sp. RC341 73 and 72% ANI and 73 and 73% AAI with *V. vulnificus* and *V. parahaemolyticus*, respectively (see additional figures 5.3, 5.4, and 5.5). Furthermore, comparative analysis of the *rpoB* sequence demonstrates that *Vibrio* sp. RC341 and *Vibrio* sp. RC586 have <97.7% sequence identity with the *rpoB* sequences of all *V. cholerae* and *V. mimicus* strains included in this study. In a comparative DNA-DNA hybridization and ANI analysis, Adékambi et al. (2008) demonstrated that *rpoB* <97.7% correlated with DNA-DNA hybridization <70% and ANI <95%, both being interpreted as demarcation thresholds for bacteria. All *V. cholerae* strains included in this study showed >99.5% *rpoB* sequence similarity with *V. cholerae* N16961 (data not shown). Based on a standard MLSA for the *Vibrionaceae* (Thompson et al., 2009), *Vibrio* sp. RC341 and *Vibrio* sp. RC586 both <95% pair-wise similarity with *V. cholerae*, *V. mimicus*, *V. vulnificus*, and *V. parahaemolyticus* strains. All *V. cholerae* strains and both *V. mimicus* strains used in this analysis demonstrated >95% similarity between concatenated genes of like-species (data not shown). Karlin's dissimilarity signatures were also calculated between these two genomes and the *Vibrio* genomes used in this study. *Vibrio* sp. RC586 shared >10 dissimilarity with all *V. cholerae* (11.5 to 16.2), *V. vulnificus* (19.6), and *V. parahaemolyticus* (41.6) genomes, and > 7 with both *V. mimicus* strains. *Vibrio* sp. RC341 >10 dissimilarity for all *V. cholerae* (10.2 to 14) except *V. cholerae* B33 (9.4) and TMA21 (9.8). *Vibrio* sp. RC341 shared >10 genome signature

dissimilarity with *V. parahaemolyticus* (40.2), *V. vulnificus* (16.3), and both *V. mimicus* (>14) genomes. *Vibrio* sp RC341 and RC586 shared a genomic dissimilarity of 8.7 with each other. Taken together these data indicate that *Vibrio* sp. RC341 and *Vibrio* sp. RC586 are new species with a high genomic relatedness to *V. cholerae* and *V. mimicus*.

**Evolution of *Vibrio* sp. RC341 and *Vibrio* sp. RC586 Lineages**

The phylogenies of *Vibrio* sp. RC341 and *Vibrio* sp. RC586 were inferred by constructing a supertree, using a 362,424 bp homologous alignment of *V. cholerae, V. mimicus*, and the new species (Figure 5.2). Based on the supertree analysis *Vibrio* sp. RC341 and *Vibrio* sp. RC586 are deeply rooted in ancestral nodes, suggesting ancient evolution of the two species. Results of this phylogentic analysis suggest the *Vibrio* sp. RC341 lineage evolved from a progenitor of the *V. cholerae* and *V. mimicus* lineages (Figure 5.2), a finding supported by strong bootstrap support and further evidenced by the evolutionary distance of *V. cholerae* and *V. mimicus* from *Vibrio* sp. RC341 (see additional figure 5.6). The two *V. mimicus* strains are interspersed among *V. cholerae*, with respect to evolutionary distance, suggesting that evolutionary distances of *V. cholerae* and *V. mimicus* are equidistant from *Vibrio* sp. RC341 (see additional figure 5.6).

The phylogeny of *Vibrio* sp. RC586 suggests it evolved from an ancestral member of the *V. mimicus* lineage after the lineage evolved from a progenitor of *V. mimicus*/*V. cholerae* (Figure 5.2). These iterations are supported by strong bootstrap support calculations. A close evolutionary relationship for *Vibrio* sp. RC586 and *V. mimicus* is also supported by shorter evolutionary distances between the *Vibrio* sp. RC586 and *V.*

*mimicus* strains (see additional figures 5.7 and 5.8). The evolutionary distance of all genomes used in this study from *V. cholerae* BX 330286, a putative progeny of the progenitor of the 7$^{th}$ pandemic clade (Chun et al., 2009; Haley et al., 2010), is shown in additional figure 5.9.

**Virulence Factors**

Both *Vibrio* sp. RC586 and *Vibrio* sp. RC341 genomes encode several virulence factors found in toxigenic and non-toxigenic *V. cholerae* and *V. mimicus*. These include the *toxR*/*toxS* virulence regulators, multiple hemolysins and lipases, VSP-I and II, and a type 6 secretion system. Both VSP islands are also present in pathogenic strains of the seventh pandemic clade (Dziejman et al., 2002). Although neither genome encodes CTXΦ phage, the major virulence factor encoding the cholera toxin (CT) that is responsible for the profuse secretory diarrhea caused by toxigenic *V. cholerae* and *V. mimicus*, both genomes do have homologous sequences of the chromosomal attachment site for this phage. Although these genomes do not encode TcpA, the outer membrane protein that CTXΦ attaches to during its infection cycle and ToxT, involved in CTXΦ replication and activation, they do encode several other mechanisms necessary for the complete CTXΦ life cycle and both CT production and translocation, including TolQRA, inner membrane proteins involved in CTXΦ attachment to the cell, XerCD tyrosine recombinases, which catalyze recombination between CTXΦ and the host genome, LexA, involved in CTXΦ expression, and EspD, involved in the secretion of the CTXΦ virion and CT translocation into the extracellular environment.

Neither *Vibrio* sp. RC341 nor *Vibrio* sp. RC586 encode VPI-1 or VPI-2, but *Vibrio* sp. RC341 encodes one copy of both VSP-I (VCJ_003466-VCJ_003480) and VSP-II (VCJ_000310 to VCJ_000324) and *Vibrio* sp. RC586 encodes one copy of VSP-I (VOA_002906-VOA_002918). However, neither of these strains encodes complete VSP islands, but rather variants of canonical VSP islands. Incomplete VSP islands have been frequently found in environmental *V. cholerae* and *V. mimicus* isolates (Grim et al., 2010; Taviani et al., 2010).

The *toxR*/*toxS* virulence regulators, hemolysins, lipases, and type 6 secretion system are present in all pathogenic and non-pathogenic strains of *V. cholerae* and both VSP islands are present in pathogenic strains of the seventh pandemic. Presence of these virulence factors in *V. cholerae* genomes sequenced to date, as well as their divergence consistent with the conserved core of *Vibrio* sp. RC341 and *Vibrio* sp. RC586, suggests that they comprise a portion of the backbone of many *Vibrio* species. Their widespread occurrence suggests the ability of all vibrios to be potential pathogens, but more likely, these factors have an important role in their ecology.

**Natural Competence**

Analysis of the 22 *V. cholerae* genomes that have been sequenced revealed the presence of type IV pili genes, involved in natural transformation of *Haemophilus* spp. and *Neisseria* spp. and other competent Bacteria (Barnhart et al., 1963; Wolfgang et al., 1998). *Vibrio* sp. RC341 and *Vibrio* sp. RC586 also encode this system. Moreover, both species encode all 33 ORFs described by Meibom et al. (2004; 2005) that comprise the chitin utilization program for induction of natural competence. The presence of these

systems in the two new species and in *V. cholerae* indicates natural competence is widely employed by vibrios to incorporate novel DNA into their genomes and, thereby, enhance both adaption to new environments and in evolution. Furthermore, the well-established association of these bacteria with chitinous organisms and with high densities in biofilms (Pruzzo et al., 2008) supports the notion that natural competence and horizontal gene transfer are both highly expressed and common in vibrios.

### Genomic Islands and Integration Loci for Exogenous DNA

Analysis of 23 complete and draft *V. cholerae* genomes by Chun et al. (2009) showed 73 putative genomic islands to be present. By pairwise reciprocal comparison, the genomes of *Vibrio* sp. RC341 and *Vibrio* sp. RC586 are concluded to encode several of these genomic islands, as well as many of the insertion loci of *V. cholerae* genomic island (2009), indicating extensive horizontal transfer of genomic islands. *V. cholerae* insertion loci are not specific to individual genomic islands, but can act as integration sites for a variety of islands(2009). *Vibrio* sp. RC586 contains 33 putative GI insertion loci and *Vibrio* sp. RC341 contains 40 that are homologous to *V. cholerae*. In addition to having highly similar attachment sequences and insertion loci, as found in *V. cholerae*, most of the homologous tRNA sequences between *Vibrio* sp. RC341, *Vibrio* sp. RC586, and *V. cholerae* are identical (100%). However, three glutamine-tRNA and one aspartate-tRNA sequence of *Vibrio* sp. RC586 and four glutamine-tRNA and four aspartate-tRNA sequences of *Vibrio* sp. RC341 show between 99 and 97% similarity with homologous *V. cholerae* tRNA sequences. These sites serve as integration loci for many pathogenicity islands. Interestingly, all tRNA-Ser, the loci most commonly targeted by island encoded

integrases of mobile elements in *V. cholerae* (Boyd et al., 2009), were 100% similar between all strains. This high similarity of platforms serving to insert exogenous DNA suggests that the same or highly similar genomic islands are readily shared. GIs and islets with homologous *V. cholerae* insertion loci and putative function and annotations are described in Tables 5.2, 5.3.

*Vibrio* sp. RC586 putatively encodes eighteen genomic islands and islets that are also found in *V. cholerae* (Table 5.3). Of these, VSP-I, islet -2 and GIs-2, -4, -33, -34, -35, -41, -62, -64, -73, and *Vibrio* sp. RC586-GI-1 are located on the large chromosome and islets-3 and 4, and GIs-9, -10, -20, -and -61 are located on the small chromosome (see additional Table 5.2). The VSP-I island is located at the homologous insertion locus for VSP-I (VOA_002906-VOA_002918) in *V. cholerae* strains, but is a variant of the canonical island having a deletion in VC0175 (deoxycytidylate deaminase-related protein) and 90% sequence similarity to the canonical island.

*Vibrio* sp. RC586 also encodes five sequences with homology to the CTXΦ attachment site, with four of them being tandemly arranged on the putative large chromosome (VOA_000105-VOA_000126). At these loci are four elements with high similarity (82 and 81% AAI) to the RS1Φ phage-like elements (rstA1 and rstB1) of *V. cholerae* SCE264 and 97 to 100% nucleotide identity to the RS1Φ-like elements in *V. cholerae* TMA21, TM11079-80, VL426, and 623-39, reported by Chun et al. (2009) to be GI-33 (Figure 5.3). RS1Φ is a satellite phage related to CTXΦ and assists in integration and replication of the CTXΦ (Davis and Waldor, 2003; Faruque et al., 2002). However, these *V. cholerae* strains were either CTXΦ-negative or encode a CTXΦ on the other chromosome, while encoding sequences with high similarity to *rstA*, and *rstB* of RS1Φ,

RS1-type sequences (Mukhopadhyay et al., 2001). Immediately upstream of the rstA1-like sequence is an hypothetical protein and immediately downstream of this rstB1-like sequence is an hypothetical protein with 52% identity with that of *Colwellia psychrerythraea* 34H, and a sequence with 99% similarity to an end-repeat (ER) region and an intergenic region (ig) of CTXΦ (Figure 5.3). This region may represent a novel phage containing ORFs with similarity to the RS1Φ satellite phage and ER and ig-1 regions with high similarity to CTXΦ. Absence of an integrase in this region suggests it may integrate into the genome via XerCD tyrosine recombinases, as does CTXΦ. All genomic islands shared by *V. cholerae* and *Vibrio* sp. RC586 are listed in Table 5.3.

*Vibrio* sp. RC341 putatively encodes 14 genomic islands and islets that are also found in *V. cholerae* (Table 5.2). VSP-I and -II and GIs-1 to 4, 33, and islets-1 to 5 are located on the large chromosome, while GI-9 and 10 are located on the small chromosome. These GIs were described by Chun et al. (2009) and two are single copies of VSP-I (VCJ_003466 to VCJ_003480) and VSP-II (VCJ_000310 to VCJ_000324). Neither of the VSP islands was present in their entirety, compared to 7$^{th}$ pandemic *V. cholerae* strains. Similar to the VSP-I variant in *Vibrio* sp. RC586, the variant in *Vibrio* sp. RC341 has a deletion of VC0175. Also, ORFs VCJ_003468 to VCJ_003470 are annotated as phage integrase, transposase, and phage integrase, respectively. The homologous ORFs of this VSP-I variant have a 92% sequence similarity to the canonical VSP-I island. Interestingly, VSP-II variant of *Vibrio* sp. RC341 contains a 10 kb putative phage encoding a type 1 restriction modification system, has a %GC of ca. 38%, and is located at the GI-56 insertion locus (tRNA-Met) (Figure 5.4). This phage shares significant similarity with *V. vulnificus* YJ016 phage (94% query coverage and 98%

sequence similarity). Several variants of VSP-II are encoded in multiple strains of *V. cholerae* (Taviani et al., 2010). However, the variant encoded in *Vibrio* sp. RC341 is, to date, unique.

Interestingly, *Vibrio* sp. RC341 encodes *V. cholerae* GI-33, a ca. 2615 bp region, (VCJ_001870 to VCJ_001874) similar to RS1Φ-like phage in *Vibrio* sp. RC586, *V. cholerae* strains VL426, SCE264, TMA21, TM11079-80, and 623-39, showing 93 to 96% nucleotide sequence similarity across 67 to 79% of the phage (Figure 5.3). This region in *Vibrio* sp. RC341 encodes only the rstA1 and rstB1 and the 3' hypothetical protein flanked by CTXΦ-like end repeats and an intergenic region, inserted at the homologous CTXΦ attachment site on chromosome I (Figure 5.3). Analysis of this and similar phages inserting at this locus suggests an extremely high diversity of vibriophages in both structure and sequence in the environment.

**Horizontal Gene Transfer of Genomic Islands**

Homologous genomic islands typically showed higher ANI between strains than the conserved backbone regions of these genomes, an indication of recent transfer of these islands among the same and different species. All GIs shared by *Vibrio* sp. RC586 and *V. cholerae* strains were 87 to 100% ANI%, with the exception of two GIs with 77% (GI-9) and 82% (GI-62) ANI. All GIs among *Vibrio* sp. RC341 and *V. cholerae* had 87 to 99% ANI, excluding three GIs with 81 to 82% (GIs-3, 9, and 2), and two with and 85% (GI-1, *Vibrio* sp. RC341 islets 1 and 2) ANI.

Phylogenetic analysis using homologous ORFs of the genomic islands yielded evidence of recent lateral transfer of VSP-I, and GIs-2, 41, and 61 among *V. cholerae* and

*Vibrio* sp. RC586. In all cases, phylogenies inferred by the ORFs were incongruent with species phylogeny, suggesting the elements were transferred after the species diverged (see additional Figures 5.10 to 5.14). Using the same methods, we found evidence of recent lateral transfer of VSP-I, GI-4, and islet-3, between *V. cholerae* and *Vibrio* sp. RC341. In all cases, phylogenies inferred by the ORFs were incongruent with species phylogeny (see additional Figures 5.12, 5.13, and 5.15). Our data suggests that VSP-I was transferred from a progenitor of the *V. cholerae* biotype albensis (*V. cholerae* VL426) lineage to *Vibrio* sp. RC586 and *Vibrio* sp. RC341. We also found evidence of horizontal transfer of *V. cholerae* GI-2 from *V. cholerae* to *Vibrio* sp. RC586, as well as *Vibrio* sp. RC341 Islet-3 and *V. cholerae* GI-4 from *Vibrio* sp. RC341 to *V. cholerae* strains.

VSP-II, islets-2, -4, -5, and GIs-1, -2, -3, -9, -10, all present in at least one *V. cholerae* genome and in *Vibrio* sp. RC341, showed no evidence of horizontal gene transfer. Most likely there are many undescribed variants of these elements, in both structure and nucleotide sequence, yet to be found in the natural environment, with certain variants more frequently transferred among strains of the same species. Coevolution of the island and host genome over time no doubt occurs. In any case, based on the data reported here *V. cholerae* is not alone in propagating these elements. They surely cycle among different but closely related species in the environment.

**Unique Genomic Islands**

*Vibrio* sp. RC586 encodes five unique genomic islands and islets not yet reported for *V. cholerae*. *Vibrio* sp. RC586 GI-2 and islet-5 encode phage-like elements.

Interestingly, islet-5 is annotated as probable coat protein A precursor, with similarity to bacteriophage f237 ORF5 of *V. campbellii* and zona occludens toxin (zot), with high similarity to *V. parahaemolyticus* and *V. harveyi* zot (VOA_001598-VOA_001600). This phage-like element is inserted at the homologous locus for *V. cholerae* O1 Classical CTXΦ insertion (VCA0569-VCA0570). *Vibrio* sp. RC586 GI-4 encodes sequences homologous to the Tn7 transposition *tns*ABCDE, a transposon known to integrate into phylogenetically diverse organisms and form genomic islands. *Vibrio* sp. RC586 GIs-1, -3, -4, and islets-1 through 6 all share homologous insertion loci with previously described *V. cholerae* GIs.

*Vibrio* sp. RC341 encodes six putative unique genomic islands not reported before. *Vibrio* sp. RC341 GIs-1, 2, 3, 4, and 7 all encode phage-like/related elements. *Vibrio* sp. RC341 GI-4 and 7 both encode several transposases and a sequence with homology to an insertion-like sequence in the *V. parahaemolyticus* insertion sequence element ISV-3L. *Vibrio* sp. RC341 GI-6 (VCJ_002614 to VCJ002618), ca. 4962 bp region of hypothetical proteins and transposases, is inserted at the homologous locus for *V. cholerae* O1 Classical CTXΦ, a locus shown to harbor a variety of GIs and phages (Chun et al., 2009).

## ***Conclusions***

The genomes of two new *Vibrio* species previously characterized as variant *V. cholerae,* have been sequenced and their sequences used to describe their interesting and important features. The genomes of both species reveal significant nucleotide sequence divergence (12 to 15%) from each other and from *V. cholerae* and *V. mimicus* genomes,

supporting the conclusion that both represent unique species not described before. Moreover, genes conserved among *V. cholerae, V. mimicus,* and the two new species varied sufficiently to suggest ancient speciation via genetic drift of the ancestral core genomic backbone. Furthermore, results of our analyses suggest *Vibrio* sp. RC341 to have evolved from a progenitor of *V. cholerae* and *V. mimicus,* whereas *Vibrio* sp. RC586 is concluded to have evolved from an early *V. mimicus* clade. Although the ANI of all genomes analyzed in this study demonstrates divergence, putative genomic islands were found to cross species boundaries, often at an higher ANI than the conserved backbone. These data, coupled with phylogenetic analyses, point to lateral transfer of the islands and phages among *V. cholerae, V. mimicus, Vibrio* sp. RC341, and *Vibrio* sp. RC586 in the natural environment. Furthermore, homologous GI insertion loci were present in both new species and in the case of *V. cholerae,* these insertion loci were not GI-specific. The pool of DNA laterally transferred between and among members of the *Vibrionaceae* strongly suggests that near-neighbors of *V. cholerae* act as reservoirs of mobile genetic elements and virulence in the environment and that *V. cholerae* is not alone in propagating these elements therein. Results of this study also demonstrate a widespread allelic variation in these elements and evidence of evolution of mobile genetic elements, including pathogenicity islands, through a multistep mosaic recombination with other elements, including phage. The ability of vibrios to incorporate exogenous DNA at several loci that encode a large combination of GIs, thereby, allows optimization of the genome for success in a specific niche or wider ecology in the natural environment.

| Species | Strain | Serogroup/ Serotype | Biotype | Geographical origin | Source of isolation | Year of isolation | Accession |
|---|---|---|---|---|---|---|---|
| *Vibrio* sp. RC341 | RC341 | O153 | | Chesapeake Bay | Water | 1999 | NZ_ACZT00000000 |
| *Vibrio* sp. RC586 | RC586 | O133 | | Chesapeake Bay | Water | 1999 | NZ_ADBD00000000 |
| *V. cholerae* | N16961 | O1 Inaba | El Tor | Bangladesh | Clinical | 1975 | NC_002505/NC_002506 |
| *V. cholerae* | RC9 | O1 Ogawa | El Tor | Kenya | Clinical | 1985 | NZ_ACHX00000000 |
| *V. cholerae* | MJ-1236 | O1 Inaba | El Tor | Matlab, Bangladesh | Clinical | 1994 | NC_012668/NC_012667 |
| *V. cholerae* | B33 | O1 Ogawa | El Tor | Beira, Mozambique | Clinical | 2004 | NZ_ACHZ00000000 |
| *V. cholerae* | MO10 | O139 | | Madras,India | Clinical | 1992 | NZ_AAKF00000000 |
| *V. cholerae* | 2740-80 | O1 Inaba | El Tor | US Gulf Coast | Water | 1980 | NZ_AAUT01000000 |
| *V. cholerae* | BX 330286 | O1 Inaba | El Tor | Australia | Water | 1986 | NZ_ACIA00000000 |
| *V. cholerae* | MAK757 | O1 Ogawa | El Tor | Celebes Islands | Clinical | 1937 | NZ_AAUS00000000 |
| *V. cholerae* | NCTC 8457 | O1 Inaba | El Tor | Saudi Arabia | Clinical | 1910 | NZ_AAWD01000000 |
| *V. cholerae* | O395 | O1 Ogawa | Classical | India | Clinical | 1965 | NC_009456/NC_009457 |
| *V. cholerae* | V52 | O37 | | Sudan | Clinical | 1968 | NZ_AAKJ02000000 |
| *V. cholerae* | 12129(1) | O1 Inaba | El Tor | Australia | Water | 1985 | NZ_ACFQ00000000 |
| *V. cholerae* | TM 11079-80 | O1 Ogawa | El Tor | Brazil | Sewage | 1980 | NZ_ACHW00000000 |
| *V. cholerae* | VL426 | non-O1/O139 | albensis | Maidstone, Kent, UK | Water | Unknown | NZ_ACHV00000000 |
| *V. cholerae* | TMA21 | non-O1/O139 | | Brazil | Seawater | 1982 | NZ_ACHY00000000 |
| *V. cholerae* | 1587 | O12 | | Lima, Peru | Clinical | 1994 | NZ_AAUR01000000 |
| *V. cholerae* | RC385 | O135 | | Chesapeake Bay | Plankton | 1998 | NZ_AAKH02000000 |
| *V. cholerae* | MZO-2 | O14 | | Bangladesh | Clinical | 2001 | NZ_AAWF01000000 |
| *V. cholerae* | V51 | O141 | | USA | Clinical | 1987 | NZ_AAKI02000000 |
| *V. cholerae* | MZO-3 | O37 | | Bangladesh | Clinical | 2001 | NZ_AAUU01000000 |
| *V. cholerae* | AM-19226 | O39 | | Bangladesh | Clinical | 2001 | NZ_AATY01000000 |
| *V. cholerae* | 623-39 | non-O1/O139 | | Bangladesh | Water | 2002 | NZ_AAWG00000000 |
| *V. mimicus* | VM223 | | | Sao Paulo, Brazil | Bivalve | | NZ_ADAJ00000000 |
| *V. mimicus* | MB-451 | | | Matlab, Bangladesh | Clinical | | NZ_ADAF00000000 |
| *V. parahaemolyticus* | RIMD2210633 | O3:K6 | | Kansai, Japan | Clinical | 1996 | NC_004603/NC_004605 |
| *V. vulnificus* | YJ016 | | 1 | Taiwan | Clinical | | NC_005139/NC_005140 |

Table 5.1. *Vibrio* strains used in the comparative genomics utilized in this study. NCBI Genbank accession numbers are listed in the right column.

| *Vibrio* sp. RC341 | | | | | | |
|---|---|---|---|---|---|---|
| | GI | Insertion Site | | Hosts | Direction of Transfer | Description/Functional Annotation |
| Chromosome I | *Vibrio* sp. RC341 GI-1 | VCJ_000612-VCJ_000633 [e] | VC0378-VC0379 [f] | 25 [g] | N/A | site-specific recombinase, phage integrase family, methyl-accepting chemotaxis protein, hypothetical proteins, putative phage-related replication initiation protein |
| | VSP-I [a, b, d] | VCJ_003465-VCJ_003481 | VC0186-VC0174 | 1-5, 20, 23 | *V. cholerae* → RC341 | VSP-I |
| | VSP-II [b, d] | VCJ_000309-VCJ_rna011 | VC0489-VC0517 | 1-5, 17,18,21, 25 | N/A | VSP-II; phage integrase, type I restriction modification system |
| | *Vibrio* sp. RC341 Islet-1 [b] | VCJ_000450-VCJ_000452 | VC0080-VC0081 | 21,24, 25 | N/A | methyl-accepting chemotaxis protein/sensory box protein |
| | *Vibrio* sp. RC341 GI-2 [c] | VCJ_000471-VCJ_rna020 | VC0060 | None | N/A | integrase for prophage CP-933T, putative replication initiation protein, phage-related, C protein, methyl-accepting chemotaxis protein |
| | *Vibrio* sp. RC341 GI-3 [c, d] | VCJ_001039-VCJ_001102 | VC1425 | None | N/A | transposases, chitinase, ompA precursor, phage-related protein, antirestriction protein |
| | *Vibrio* sp. RC341 Islet-2 [b] | VCJ_001117-VCJ_001119 | VC1407-VC1411 | 1-17, 19, 21 | N/A | multidrug efflux pump operon, TetR (AcrR) family |
| | Vc GI-1 [b] | VCJ_001123-VCJ_001134 | VC1391-VC1406 | 1-16, 19 | N/A | motility and chemotaxis |
| | *Vibrio* sp. RC341 Islet-3 [a, b] | VCJ_001197-VCJ_001202 | VC1328-VC1332 | 12, 14, 15, 18 | RC341 → *V. cholerae* | multidrug resistance efflux pump |
| | *Vibrio* sp. RC341 GI-4 [c, d] | VCJ_001439-VCJ_01452 | VC2622-VC2623 | None | N/A | putative regulatory prophage protein, phage recombinase, putative, type I restriction-modification system |
| | *Vibrio* sp. RC341 Islet-4 [b] | VCJ_001536-VCJ_001538 | VC2714-VC2715 | 10, 13, 15, 16, 18, 19 | N/A | GGDEF domain protein |
| | Vc GI-33 [d] | VCJ_001869-VCJ_001875 | VC1445-VC1480 | 15, 20-23 | N/A | RSΦ-like element |
| | *Vibrio* sp. RC341 Islet-5 [b] | VCJ_001947-VCJ_001950 | VC1560-VC1563 | 1-13, 16-19, 23 | N/A | Beta-lactamase-related protein, LysR-family transcriptional regulator VC1561 |
| | Vc GI-2 [b] | VCJ_001965-VCJ_001969 | VC1587-VC1579 | 1-9, 11-12, 18, 20, 23 | N/A | superoxide stress response |
| | Vc GI-3 [b] | VCJ_002126-VCJ_002130 | VC1746-VC1754 | 1-9, 11-12, 18-19, 21 | N/A | paraquat-inducible protein A, small-conductance mechanosensitive channel |
| | *Vibrio* sp. RC341 GI-5 [c] | VCJ_002133-VCJ_002141 | VC1757-VC1810 | None | N/A | anaerobic C4-dicarboxylate transporter dcuC, tripeptide aminopeptidase, alpha-aspartyl dipeptidase peptidase E, serine/threonine protein kinase related protein, adenylate cyclase, transcriptional regulator |
| | Vc GI-4 [a, b, d] | VCJ_002147-VCJ_002157 | VC1816-VC1828 | 1-12,15, 17-19, 23, 24 | RC341 → *V. cholerae* | carbohydrates (PTS system) |
| Chromosome II | Vc GI-9 [b] | VCJ_003140-VCJ_003132 | VCA 0849-VCA 0859 | All exc. 18, 19-20 | N/A | response regulator, autolysin sensor kinase, ABC-type transport system |
| | Vc GI-10 [b] | VCJ_003152-VCJ_003207 | VCA 0870-VCA 0885 | 7, 10, 14, 17, 18, 20, 22, 23 | N/A | phage integrase, streptococcal hemagglutinin protein, secretion activator protein, hypothetical proteins, type II restriction modification system |
| | *Vibrio* sp. RC341 GI-6 [c] | VCJ_002613-VCJ_002619 | VCA 0569-VCA 0570 | None | N/A | transposases, hypothetical proteins |
| | *Vibrio* sp. RC341 GI-7 [c, d] | VCJ_002860-VCJ_002887 | VCA 0235-VCA 0236 | None | N/A | phage-related protein, transposases, type I restriction modification system, hypothetical protein, DNA helicase related protein |
| | *Vibrio* sp. RC341 Islet-6 [c] | VCJ_002907-VCJ_002910 | VCA 0197-VCA 0204 | None | N/A | transposase, hypothetical protein |

a evidence of horizontal transfer as determined by incongruent phylogeny (see additional file 1: Figs. 10 to 15)

b genomic islands found in clinically isolated strains of *V. cholerae* or *V. mimicus* as determined by Chun et al. [15]

c genomic islands not yet found in other strains

d confirmed by IslandViewer [41]

e insertion loci in *Vibrio* sp RC341 (either flanking genes or island borders if island runs off both ends of a contig)

f homologous flanking ORFs in *V. cholerae* N16961 as determined by Chun et al. [15]

g strains used in this study encoding the same islands (see additional file: Table 3)

Table 5.2. Putative genomic islands of *Vibrio* sp. RC341.

| | | | | | | |
|---|---|---|---|---|---|---|
| Chromosome I | Vc GI-33 [d] | VOA_000105-VOA_000126 [e] | VC1445-VC1481 [f] | 15, 20-23 [g] | N/A | RS1Φ-like element |
| | *Vibrio* sp. RC586 Islet-1 | VOA_000163-VOA_000165 | VC1519-VC1520 | None | N/A | hypothetical protein |
| | Vc GI-64 | VOA_000199-VOA_000212 | VC1560-VC1561 | 18, 20 | N/A | lux operon |
| | *Vibrio* sp. RC586 GI-1 | VOA_000212-VOA_000216 | VC1560-VC1562 | 18 | N/A | hypothetical and conserved proteins with similarity to *V. cholerae* RC385 and *V. vulnificus* |
| | Vc GI-2 [a, b] | VOA_000234-VOA_000239 | VC1582-VC1587 | 1-9, 11-12, 18, 20, 23 | *V. cholerae* → RC586 | superoxide stress response |
| | *Vibrio* sp. RC586 GI-2 [c, d] | VOA_000549-VOA_000568 | VC1067-VC1068 | None | N/A | phage-related proteins, hypothetical proteins, error-prone repair proteins |
| | Vc GI-41 [a, d] | VOA_000782-VOA_000810 | VC0806-VC0847 | 18, 22 | N/A | phage integrase;hypothetical proteins, HCP, VgrG protein |
| | *Vibrio* sp. RC586 GI-3 [c] | VOA_002149-VOA_002155 | VC1757-VC1810 | None | N/A | transcriptional regulator, Anaerobic C4-dicarboxylate transporter dcuC, Tripeptide aminopeptidase, alpha-aspartyl dipeptidase peptidase E, diguanylate cyclase/phosphodiesterase domain 2 |
| | *Vibrio* sp. RC586 GI-4 [c] | VOA_001387-VOA_001395 | VC0487-VC0488 | None | N/A | transposon Tn7 transposition proteins tnsABCDE, retron-type reverse transcriptase, ompA precursor |
| | Vc GI-62 | VOA_001969-VOA_001973 | VC1407-VC1414 | 20 | N/A | type I restriction-modification system |
| | Vc GI-73 [d] | VOA_002071-VOA_002077 | VC1668-VC1669 | 18 | N/A | chalcone synthase (EC 2.3.1.74), cytotoxic necrotizing factor 1, ISSod2, transposase |
| | Vc GI-4 [b] | VOA_002158-VOA_002165 | VC1816-VC1828 | All exc. 13-14, 16, 20-22 | N/A | carbohydrates (PTS system) |
| | Vc GI-35 [b] | VOA_002247-VOA_002255 | VC1910-VC1911 | 13-16, 21-22, 24-25 | N/A | MoxR-like ATPase;Outer membrane receptor protein;ABC-type oligopeptide transport system;Oligopeptide transport system |
| | *Vibrio* sp. RC586 Islet-2 [b] | VOA_rna036-VOA_002388 | VC2041-VC2042 | All exc. 8, 18-19 | N/A | histone deacetylase/AcuC/AphA family protein |
| | Vc GI-34 | VOA_003551-VOA_003558 | VC2714-VC2715 | 13-14, 21-22, 24-25 | N/A | ribose ABC transport system, Beta-hexosaminidase |
| | VSP-I [a, b] | VOA_002906-VOA_002918 | VC0174-VC0186 | 1-5, 20, 23 | *V. cholerae* → RC586 | VSP-I |
| | Vc GI-61 [a, b] | VOA_001088-VOA_001091 | VCA0369-VCA1067 | 18-19 | N/A | Outer membrane lipoprotein-sorting protein;Predicted exporter of the RND superfamily;Transcriptional activator, LuxR/UhpA family of regulators |
| | *Vibrio* sp. RC586 Islet-3 | VOA_001152-VOA_001156 | VCA0018-VCA0022 | 18, 21 | N/A | putative membrane protein, hypothetical protein |
| | *Vibrio* sp. RC586 Islet-4 | VOA_001364-VOA_001367 | VCA0197-VCA0204 | 22 | N/A | chromosome segregation ATPase |
| | *Vibrio* sp. RC586 Islet-5 [c, d] | VOA_001597-VOA_001601 | VCA0569-VCA0570 | None | N/A | probable coat protein A precursor, zona occludens toxin |
| | Vc GI-20 [b] | VOA_001636-VOA_001668 | VCA0608-VCA0614 | 7 | N/A | DNA methylase, SSU ribosomal protein S2p (SAe), bacteriophage tail sheath protein, phage baseplate assembly protein, transposases, lyzozyme |
| | *Vibrio* sp. RC586 Islet-6 [b] | VOA_001820-VOA_001822 | VCA0789-VCA0797 | 24-25 | N/A | periplasmic divalent cation tolerance protein cutA |
| | Vc GI-9 [b] | VOA_000884-VOA_000892 | VCA0849-VCA0859 | 18-20, 22 | N/A | response regulator;Autolysin sensor kinase;ABC-type transport system |
| | Vc GI-10 [b, d] | VOA_000902-VOA_000957 | VCA0870-VCA0885 | 14, 15, 18, 21, 23 | N/A | phage integrase, streptococcal hemagglutinin protein, secretion activator protein, hypothetical proteins, type II restriction modification system |

a evidence of horizontal transfer as determined by incongruent phylogeny (see additional file 1: Figs. 10 to 15)

b genomic islands found in clinically isolated strains of *V. cholerae* or *V. mimicus* as determined by Chun et al. [15]

c genomic islands not yet found in other strains

d confirmed by IslandViewer [41]

e insertion loci in *Vibrio* sp RC586 (either flanking genes or island borders if island runs off both ends of a contig)

f homologous flanking ORFs in *V. cholerae* N16961 as determined by Chun et al. [15]

g strains used in this study encoding the same islands (see additional file: Table 3)

Table 5.3. Putative genomic islands of *Vibrio* sp. RC586.

| Strain | Strain No. |
|---|---|
| *V. cholerae* N16961 | 1 |
| *V. cholerae* RC9 | 2 |
| *V. cholerae* MJ-1236 | 3 |
| *V. cholerae* B33 | 4 |
| *V. cholerae* MO10B | 5 |
| *V. cholerae* BX 330286 | 6 |
| *V. cholerae* NCTC 8457 | 7 |
| *V. cholerae* 2740-80 | 8 |
| *V. cholerae* MAK 757 | 9 |
| *V. cholerae* 12129(1) | 10 |
| *V. cholerae* O395 | 11 |
| *V. cholerae* V52 | 12 |
| *V. cholerae* AM 19226 | 13 |
| *V. cholerae* 1587 | 14 |
| *V. cholerae* 623-39 | 15 |
| *V. cholerae* MZO-2 | 16 |
| *V. cholerae* MZO-3 | 17 |
| *V. cholerae* RC385 | 18 |
| *V. cholerae* V51 | 19 |
| *V. cholerae* VL426 | 20 |
| *V. cholerae* TMA21 | 21 |
| *V. cholerae* TM11079-80 | 22 |
| *Vibrio* sp. RC341/RC586 | 23 |
| *V. mimicus* MB451 | 24 |
| *V. mimicus* 223 | 25 |

Table 5.4. Strain no. corresponds to “Host” column in Tables 2 and 3 (above).

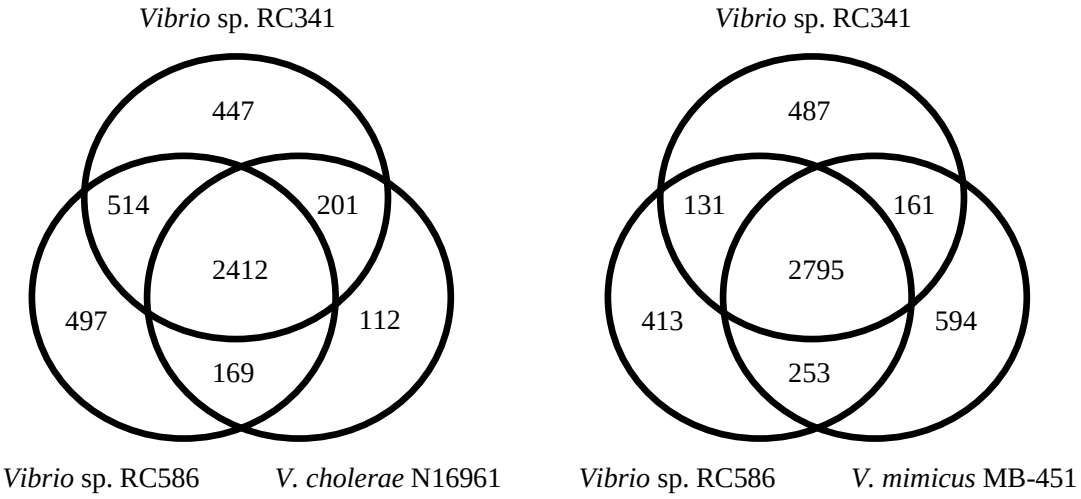



Figure 5.1. Venn diagrams showing ORFs shared by *Vibrio* sp. RC341, *Vibrio* sp. RC586, *V. cholerae* N16961, and *V. mimicus* MB-451. The number in the middle shows the conserved number of ORFs shared by the three strains. The numbers show that there are ORFs unique to that strain or that there are ORFs shared.

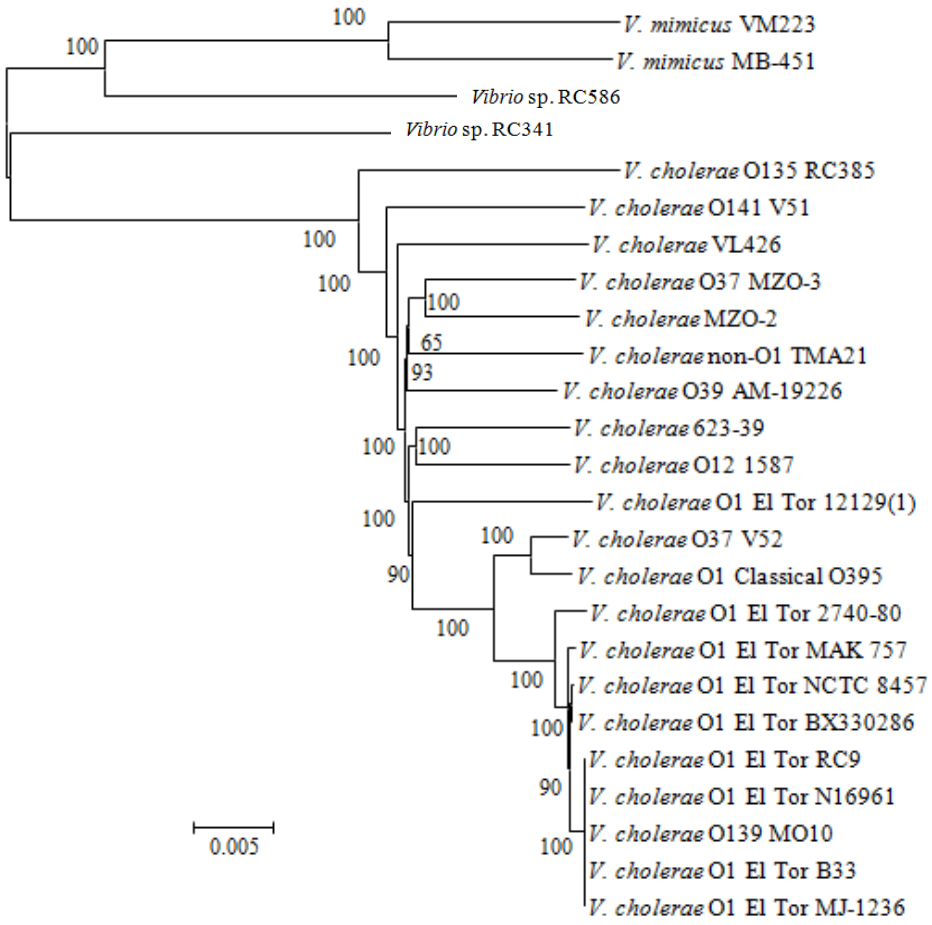



Figure 5.2. Neighbor-joining tree based on 362,424 bp alignment of homologous sequences unsing the Kimura-2 parameter for nucleotide substitution. The bootstrap supports, as percentage, are indicated at the branching points. Bar represents 0.005 substitutions per site.

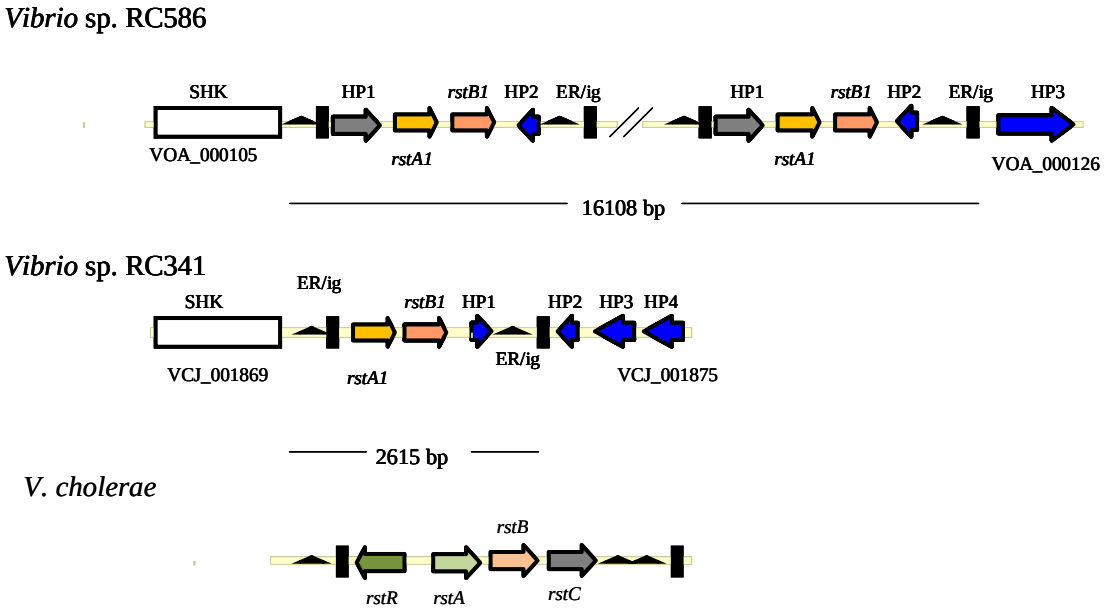



Figure 5.3. RS1Φ-like elements located at CTXΦ attachment sites on the large chromosomes of *Vibrio* sp. RC586 and *Vibrio* sp. RC341 and the canonical RS1Φ of *V. cholerae*. SHK = sensor histidine kinase, HP = hypothetical protein, ER = end repeat, ig = intergenic region.

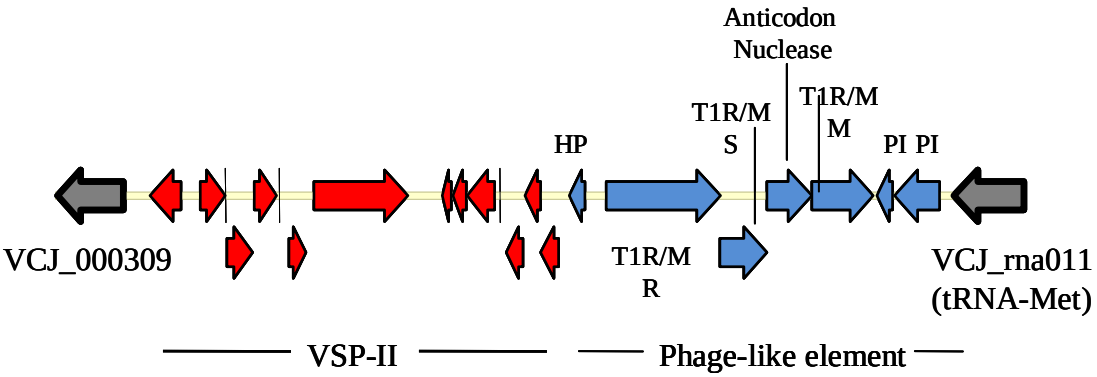
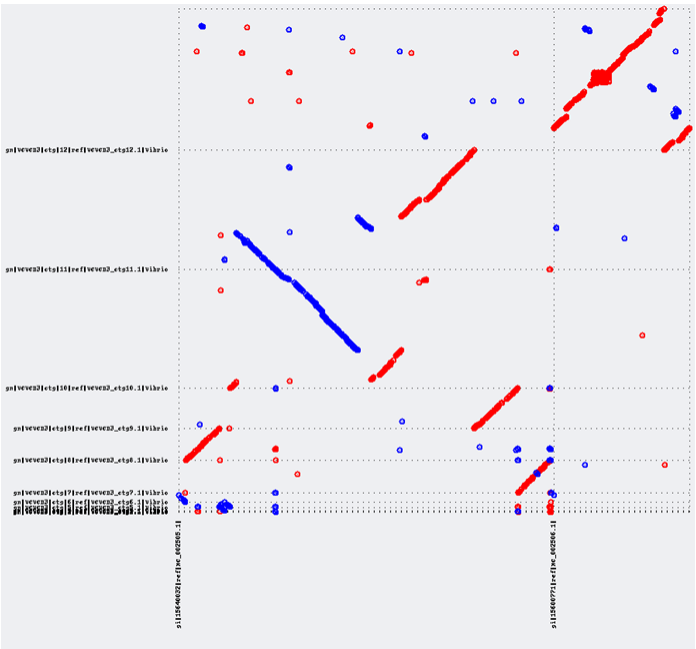
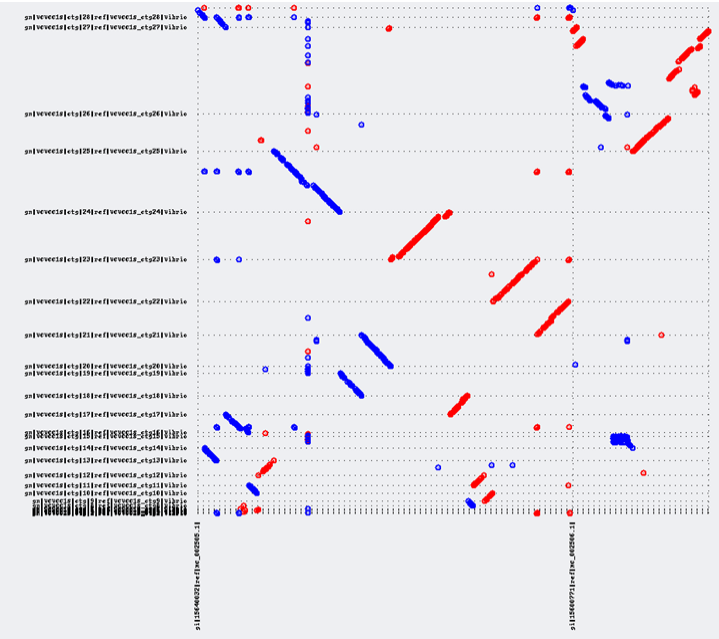


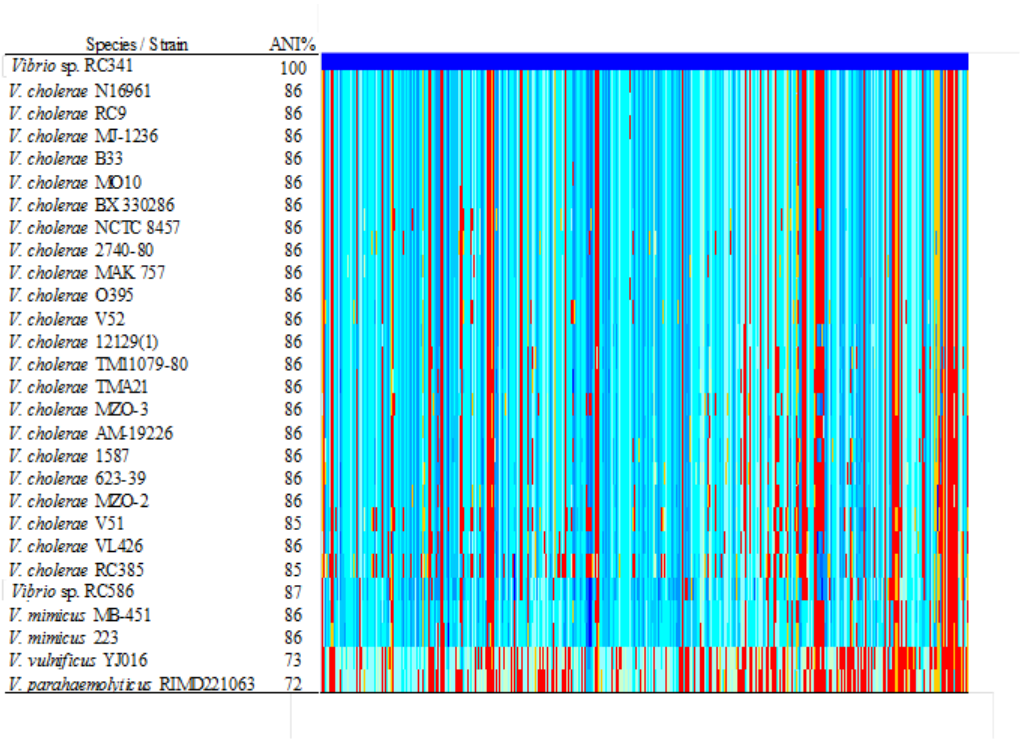

Figure 5.4. Novel VSP-II variant found in *Vibrio* sp. RC341. Red arrows represent VSP-II ORFs and blues arrows represent the novel phage-like region in the 3' region of the sequence. Grey arrows represent the adjacent flanking sequences. T1R/M = type I restriction modification system. PI = phage integrase.



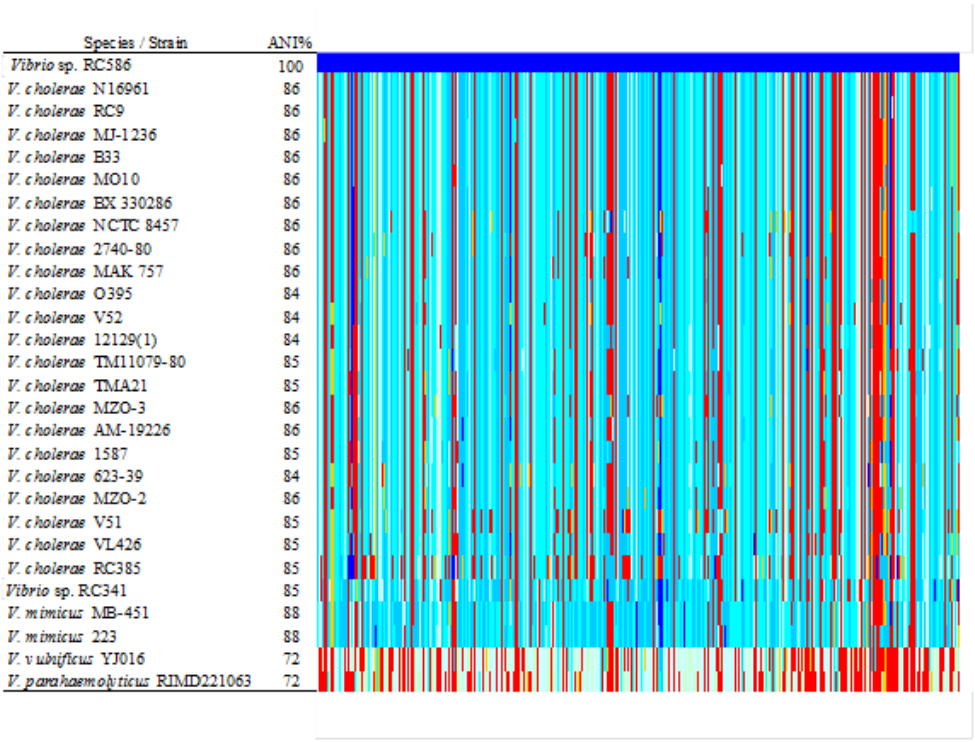
Additional Figure 5.1 MUMmer plot of *Vibrio* sp. RC586 as query and *V. cholerae* N16961 as reference. *Vibrio* sp. RC586 contigs are on Y-axis and *V. cholerae* N16961 chromosomes are on X-axis. *V. cholerae* N16961 chromosome I begins at XY-intercept and chromosome II is located on the right section of the X-axis.



Additional Figure 5.2 MUMmer plot of *Vibrio* sp. RC341 as query and *V. cholerae* N16961 as reference. *Vibrio* sp. RC341 contigs are on Y-axis and *V. cholerae* N16961 chromosomes are on X-axis. *V. cholerae* N16961 chromosome I begins at XY-intercept and chromosome II is located on the right section of the X-axis.

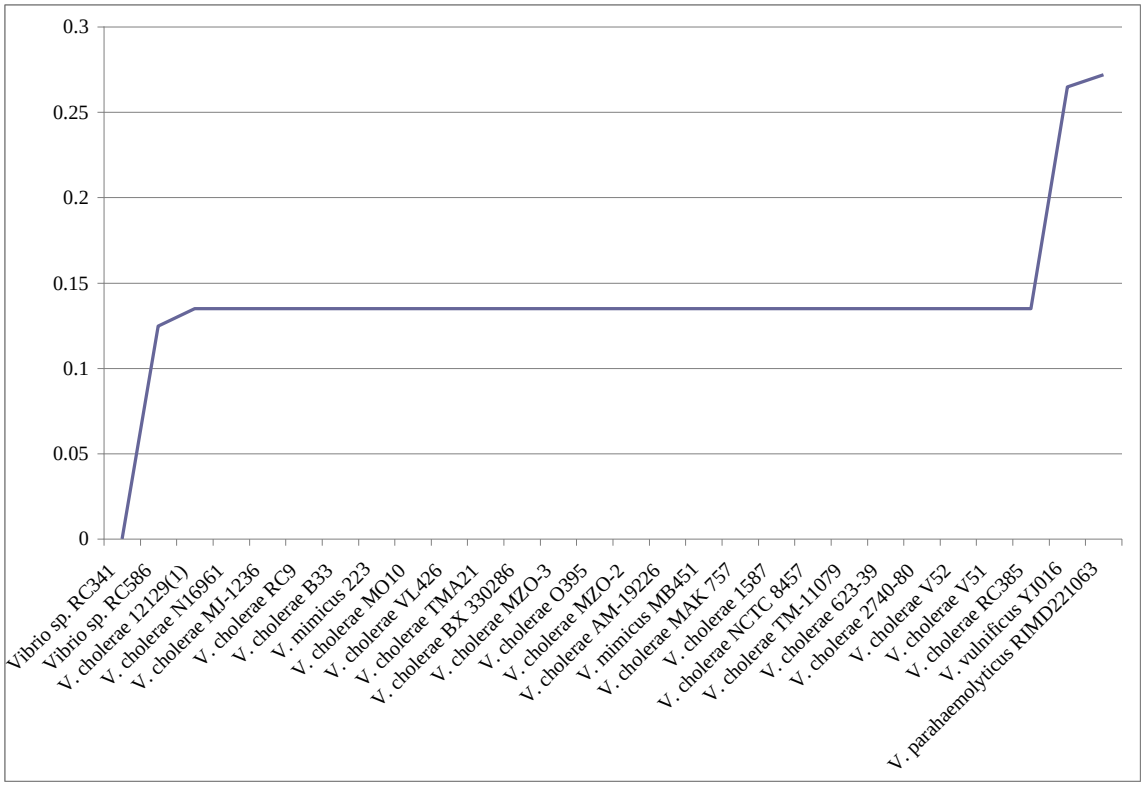



Additional Figure 5.3. Average nucleotide identity (ANI%) between *Vibrio* sp. RC341 and *Vibrio* genomes used in this study.

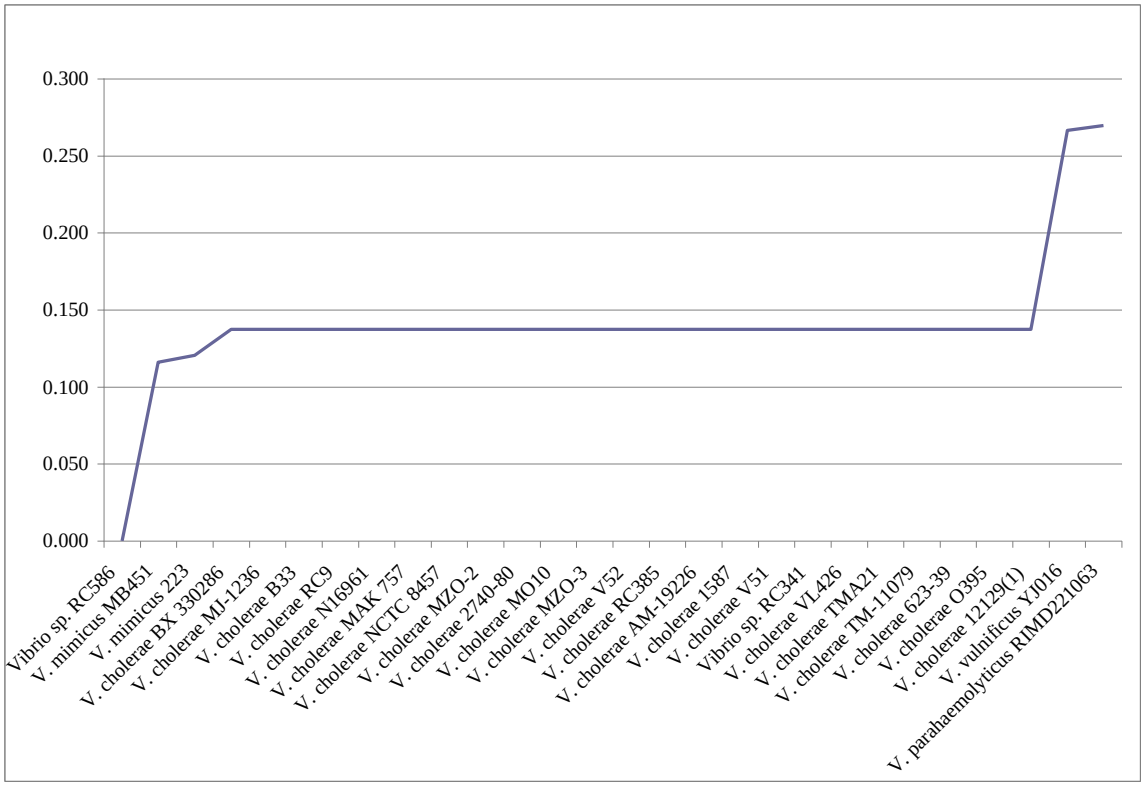



Additional Figure 5.4. Average nucleotide identity (ANI%) between *Vibrio* sp. RC586 and *Vibrio* genomes used in this study.

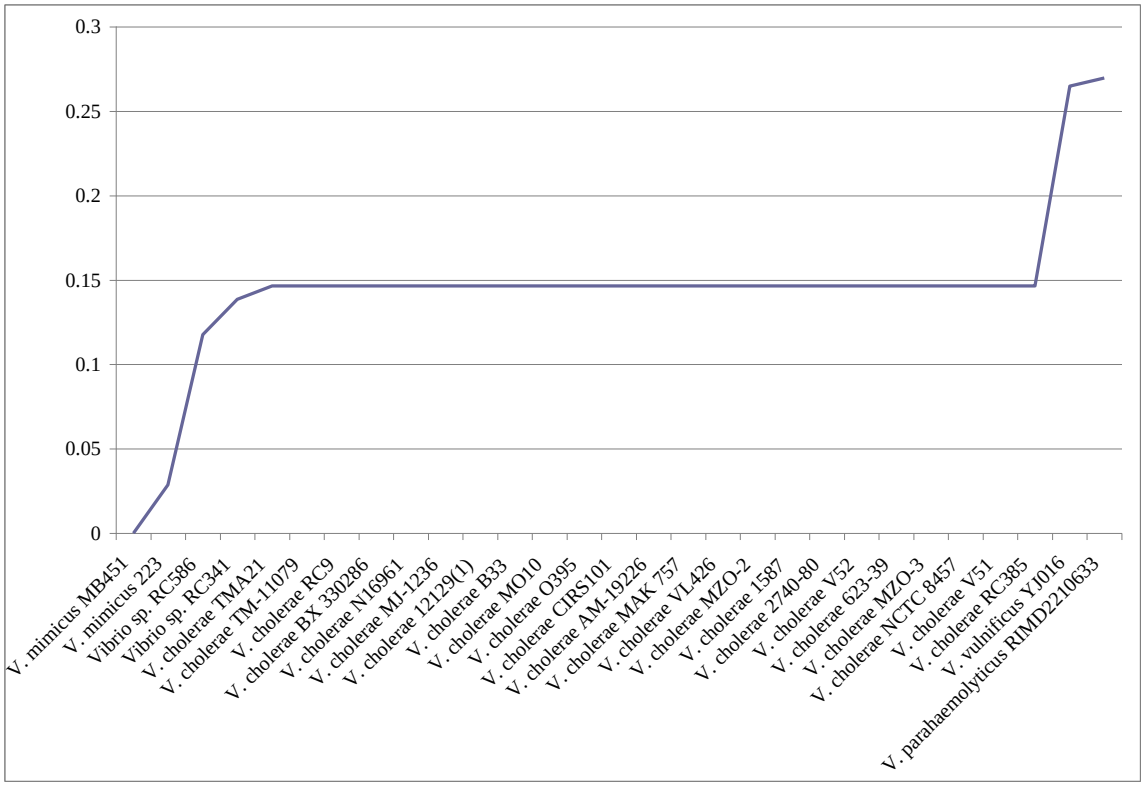

Additional Figure 5.5. BLAST atlas key for additional Figures 3 and 4.



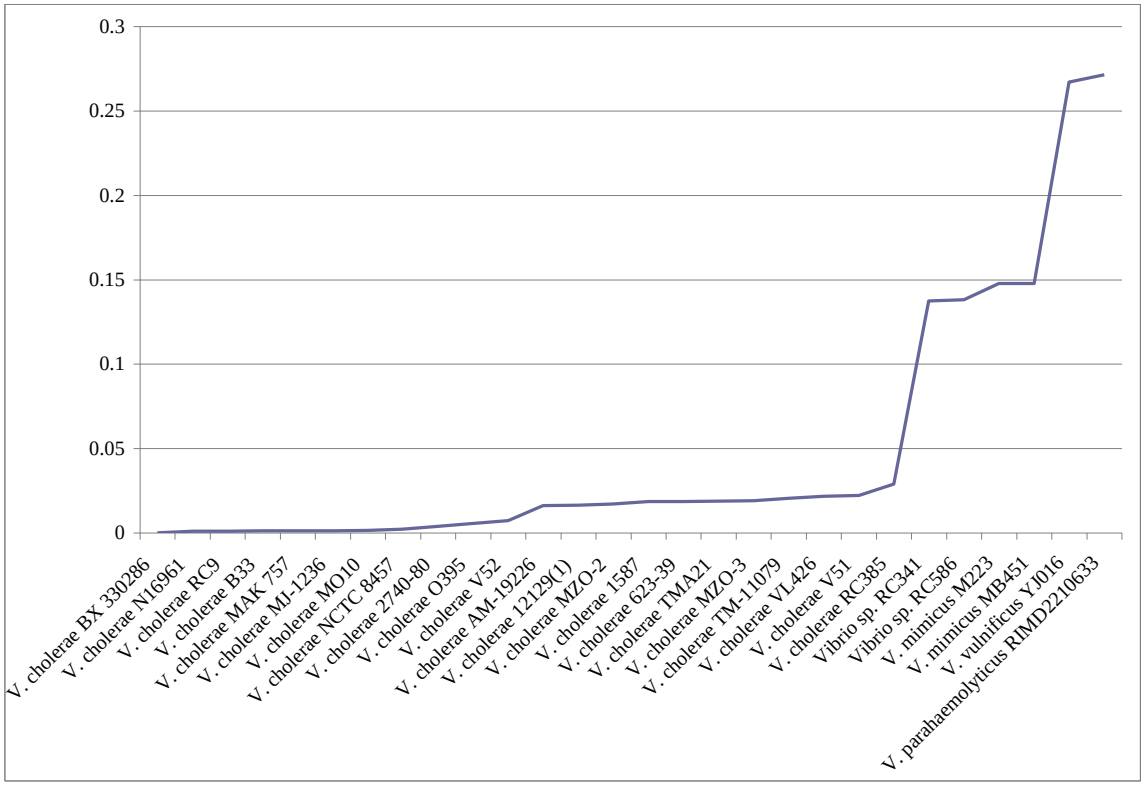

Additional Figure 5.6. Evolutionary distance of strains used in this study from *Vibrio* sp. RC341 as determined by ANI between *Vibrio* sp. RC341 and all strains used in this study.



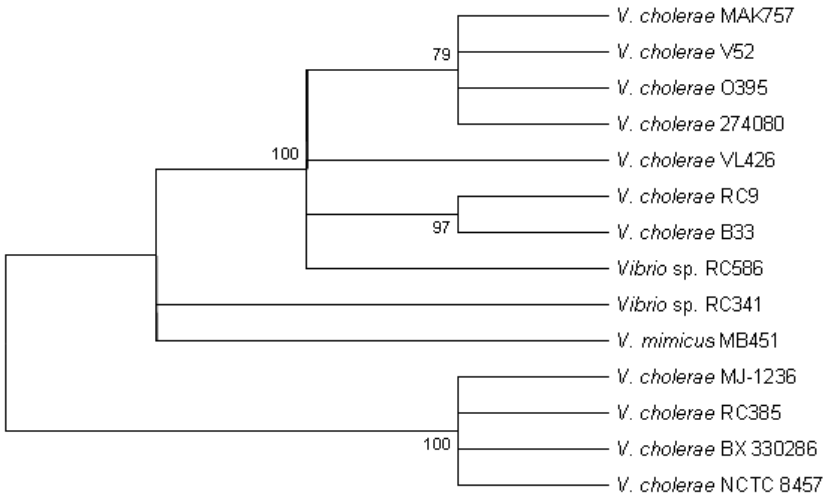

Additional Figure 5.7. Evolutionary distance of strains used in this study from *Vibrio* sp. RC586 as determined by ANI between *Vibrio* sp. RC586 and all strains used in this study.



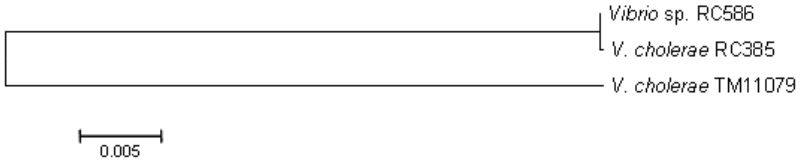

Additional Figure 5.8. Evolutionary distance of *Vibrio* sp. RC586 and *Vibrio* sp. RC341 from *V. mimicus* MB451 as determined by ANI between *V. mimicus* MB451 and all strains used in this study.



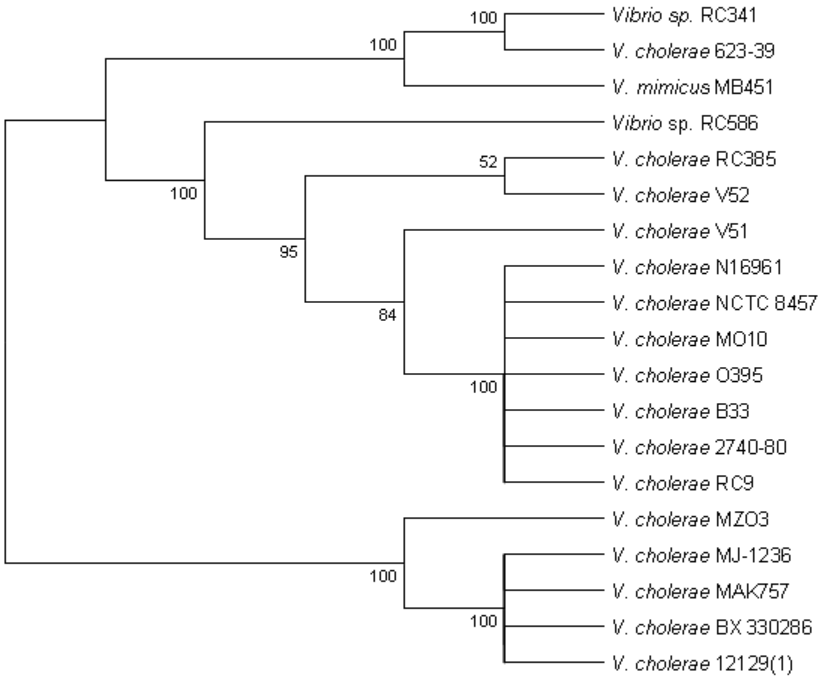

Additional Figure 5.9. Evolutionary distance of *Vibrio* sp. RC586 and *Vibrio* sp. RC341 from strains *V. cholerae* BX 330286 as determined by ANI between *V. cholerae* BX 330286 and all strains used in this study.



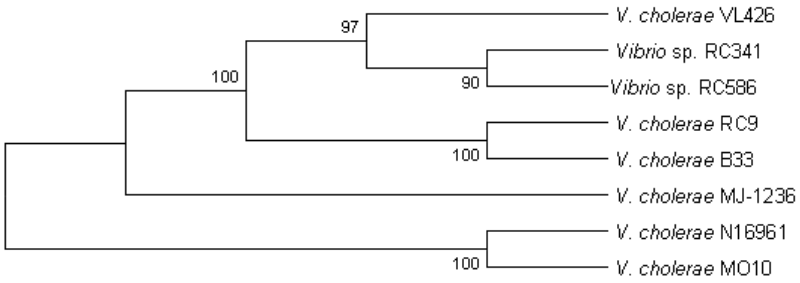

Additional Figure 5.10. Phylogeny of the genomic island GI-2 as determined by reconstructing a neighbor-joining tree using the Kimura-2 parameter as a nucleotide substitution model.



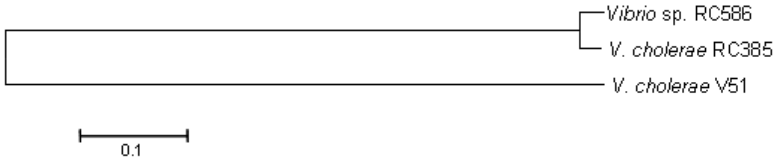

Additional Figure 5.11. Phylogeny of the genomic island GI-41 as determined by reconstructing a neighbor-joining tree using the Kimura-2 parameter as a nucleotide substitution model.



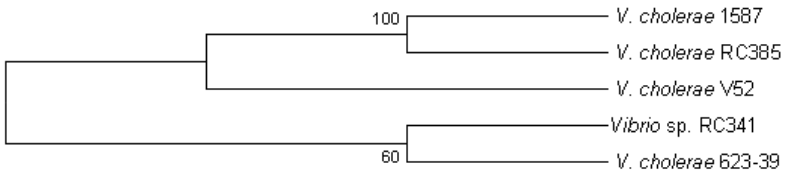

Additional Figure 5.12. Phylogeny of the genomic island GI-4 as determined by reconstructing a neighbor-joining tree using the Kimura-2 parameter as a nucleotide substitution model.




Additional Figure 5.13. Phylogeny of the genomic island VSP-I as determined by reconstructing a neighbor-joining tree using the Kimura-2 parameter as a nucleotide substitution model.




Additional Figure 5.14. Phylogeny of the genomic island GI-61 as determined by reconstructing a neighbor-joining tree using the Kimura-2 parameter as a nucleotide substitution model.




Additional Figure 5.15. Phylogeny of *Vibrio* sp. RC341 Islet-3 as determined by reconstructing a neighbor-joining tree using the Kimura-2 parameter as a nucleotide substitution model.

*Reproduction License Agreement*

# Chapter 6: Seasonality of *Vibrio metecus* in the Chesapeake Bay, MD


***Abstract***

*Vibio metecus,* a recently described bacterial species closely related to *Vibrio cholerae,* has been reported to be autochthonous to the Chesapeake Bay, the coastal waters of Massachusetts, and the Bay of Bengal. Analysis of phenotypic data indicates this organism is indistinguishable from *V. cholerae*. However, genomic analyses demonstrate genome-wide divergence between the two species. Although it has been isolated from the aquatic environment, its ecology remains to be elucidated. Here we report results of a two year study of the occurrence and distribution of *V. metecus* in the Chesapeake Bay using direct PCR of water, plankton, sediment, and oyster samples collected from the Chester River and Tangier Sound. *V. metecus* were most numerous during winter, yet it could be detected throughout the year, even during the warm summer months. Results of regression analyses demonstrated that when water temperature and salinity decreased, *V. metecus* densities and frequencies of detection increased, suggesting seasonal fluctuation in these parameters in the Chesapeake Bay influences the ecological dynamics of this organism. It is concluded that the ecology of *V. metecus* differs from that of *V. cholerae* which typically thrives at water temperatures at or near 15°C.

## ***Introduction***

Genomic analyses have demonstrated significant diversity among species of the *Vibrionaceae*, and within the *Vibrio cholerae* species (Chun et al., 2009; Thompson et al., 2009). Results of phylogenetic analysis demonstrated *Vibrio metecus* to be a novel species, phenotypically identical to *V. cholerae* in key features, but genetically distinct (Haley et al., 2010; Turnsek et al. 2010; Boucher et al., 2011). Mobile elements and virulence factors are shared by both species (Haley et al., 2010; Boucher et al., 2011). A number of clinical isolates, described as *V. cholerae* by conventional polymerase chain reaction, were subsequently shown to be *V. metecus* using specific primers. Human infections from which *V. metecus* was isolated have been increasing with time, suggesting this organism is an emerging pathogen (Turnsek et al., 2010). According to Boucher et al. (2011), like *V. cholerae* and other members of the *Vibrionaceae, V. metecus* occurs in coastal environments worldwide.

It has been established that members of the *Vibrionaceae* are seasonal in occurrence and their distribution in the aquatic environment is that are associated with physical, chemical, and biological factors, including water temperature, salinity, and zooplankton distribution (Kaneko and Colwell, 1974; Louis et al. 2003; Lizárraga-Partida et al. 2009; Turner et al. 2009; de Magney et al., 2009; Banaker et al. 2011; Schuster et al. 201; Shair et al. 2012). This newly recognized species of the *Vibrionaceae* demonstrates unique seasonal trends. To establish the ecology of this species, water, sediment, oyster, and plankton samples were collected over a 30 month period from two stations in the Chesapeake Bay, The Chester River near Annapolis, MD, and Tangier Sound, on the Eastern Shore of Maryland at the border of Virginia and Maryland (Figure

6.1). The unique seasonal signature of *V. metecus* in the Chesapeake Bay is presented here.

***Materials and Methods***

Primer sequences specific to the *toxR* (cholera toxin transcriptional activator) gene of *V. metecus* RC341 (NCBI Genbank accession no. ACZT00000000.1) were developed to discriminate this organism from closely related species. As there is at present a single reference genome for this novel species, the *toxR* gene was selected as a target since it is highly conserved within the genomes of the species, yet highly diverged between genomes of different species. A homolog of the *toxR* gene in *Vibrio* sp. RC341 was obtained by aligning the *toxR* gene (VC0984) of a high quality annotated reference genome of *Vibrio cholerae* N16961 (NCBI reference sequence accession number NC_002505.1) to the draft genome sequence of *Vibrio metecus* RC341. A multiple sequence alignment of the homologous sequences of members of the *Vibrionaceae*, whose genomes had been sequenced and taxonomically validated by determining average genomic nucleotide identity, were aligned to determine candidate primer sequences specific to *V. metecus*. This was done to ensure taxonomic delineations were related to the entire genome rather than a subset of ORFs from a multilocus sequence analysis (MLSA) that would represent ca. $\leq$ 0.1% of the genome. Furthermore, as any nucleotide sequence is potentially subject to horizontal gene transfer, with the probability of this increasing as genomic sequence similarity between genomes increase, MLSA can substantially skew species delineations in cells which this has occurred.

Primers were tested *in silico* against the NCBI nucleotide (nr/nt), reference genomic sequence (refseq_genomic) and whole-genome shotgun contigs (WGS) database to evaluate specificity for *V. metecus*. For conduct PCR 1µl of the primers The primers VmetF 5'-TCATTCGCCTAGGCAGTAACGA -3' and VmetR 5'-TTGATGCCAGCAATCAGCAGC-3' were added in 20 µM concentrations to GoTaq PCR MasterMix and 5 µl of boiled cell DNA template. Primers were also tested against a set of *Vibrionaceae* genomes (listed in Figure 6.2) using polymerase chain reaction (PCR).

Water, sediment, plankton, and oyster samples were collected in the Chester River and Tangier Sound in the Chesapeake Bay every month at each site except June through August, when samples were collected twice per month at each site. Grab water samples were collected using a beta vertical water sampler (Wildlife Supply Company, Yulee, FL), which was submerged and released approximately 1 meter from the surface. Twenty liters of water was decanted into two sterile opaque polypropylene bottles. Approximately 20 to 30 oysters were collected by trawling and associated shell sediments used as sediment samples. Oysters were placed in a plastic bag and the sediment samples were collected into a sterile 500 ml opaque polypropylene bottle. The > 20 µm plankton fractions were collected by pouring 5 liters of water through a 20 µm mesh plankton net and adding 50 ml of the trapped material to a sterile centrifuge tube. Plankton-free water was collected in a sterile 1 L polypropylene bottle by capturing the filtrate of the 20 µm plankton mesh net. Triplicates of 900, 90, and 9 ml water samples were added to 100, 10, and 1 ml 10 X alkaline peptone water (APW) and incubated at 33°C overnight on a shaker at 25 rpm. Plankton samples were further concentrated to 25 ml and 2 ml of the

concentrate were added to 18 ml 1 X APW and incubated overnight at 33°C on shaker at 25 rpm. An equal volume of sterile phosphate buffered saline solution (PBS) was added to each sediment sample, and the mixture shaken to distribute the sample. Eleven ml of PBS/sediment mixture was added to 1 ml of 10 X APW and incubated overnight at 33°C on a shaker at 25 rpm. The oysters were rinsed in cold deionized water to remove the debris from the shells and opened with a sterile oyster knife. Approximately 250 ml oyster meat and hemolymph were transferred to a sterile blender with an equal amount of PBS and homogenized for 30 seconds. As was done with the water samples, triplicates of 10, 1, and 0.1 g were added to 25 ml 10 X APW and incubated overnight at 33°C on a shaker at 25 rpm. From each overnight incubation 1 ml was removed, centrifuged at 10,000 g, decanted, resuspended in 1 ml TE, resuspended, and boiled for 10 min at 1,000 rpm. Boiled cells were chilled on ice, diluted 10-fold in TE with 40 µl bovine serum albumin per ml added to remove interaction of inhibitors. PCR was done on the boiled and diluted cells to estimate presence and abundance of *V. metecus* in samples collected from the Chester River and Tangier sound over the study period. Presence or absence of *V. metecus* in each sample type was determined by PCR on boiled APW enrichments and concentrations of *V. metecus* in water and oysters were determined by PCR on dilutions of each and calculating the most probable number (MPN) based on patterns of *V. metecus* presence and absence in each dilution following methods described in Standard Methods for the Examination of Water and Wastewater ($17^{th}$ edition, 1989). Environmental parameters were determined by using a portable water meter (YSI, Yellow Springs, OH) and recording values at both the water surface (ca. 1/3 meter from the surface) and bottom (ca. 1/3 meter from the bottom).

Correlations between *V. metecus* densities and environmental parameters were determined using Spearman rank analyses. Multiple stepwise regressions were calculated for *V. metecus* with each environmental parameter as independent variables. A backward elimination method was used and P-values of each variable were determined. Variables with $P < 0.05$ were assumed to be significant and all other variables were removed from the model. The multiple regression models were completed when the addition or deletion of variables did not improve prediction of *V. metecus* concentrations.

***Results***

An average nucleotide identity matrix showed *Vibrio* sp. RC341 was the only member of the newly recognized species, *V. metecus*, whose genome has been fully sequenced (Figure 6.2), confirmed previously (Haley et al. 2010). The, *toxR* primer sequences were developed based on this genome as the target and other species as background. The *toxR* sequence was selected as a target since it typically demonstrates $\leq$ 2 % sequence divergence within the species and > 10 % sequence divergence between species of the *Vibrionaceae*. This gene has been used repeatedly as a target for species identification in the *Vibrionaceae*, demonstrating consistency. Primer sequences developed for the *V. metecus toxR* gene did not match any sequence in the NCBI databases queried (nr/nt, refseq_genomic, and WGS) and did not result in amplification for any of the bacterial strains (listed in Figure 6.2) evaluated by PCR excluding *Vibrio* sp. RC341 (*V. metecus*).

*V. metecus*-specific PCR was done on 285 isolates recovered from plankton and water samples collected in the Chesapeake Bay between 1998 and 1999. The isolates

were confirmed as *V. cholerae* by 16S-23S rRNA intergenic spacer unit PCR (Chun et al., 1999). Our results showed that of the 285 isolates, 39 (13.6%) were *V. metecus* by *toxR* PCR. These *V. metecus* isolates formed yellow, sucrose-fermenting colonies on TCBS agar and blue-green colonies on CHROMagar™ *Vibrio*, results consistent with those of previous studies stating *V. metecus* to be indistinguishable from *V. cholerae* on selective and differential media (Haley et al., 2009; Turnsek et al., 2010). Of the 39 *V. metecus* isolates, 32 were from plankton samples (18 from >64 µm and 14 from the 63-20 µm fractions) while seven were from water. The majority of these isolates were recovered from sites at the Smithsonian Environmental Research Center in Edgewater, MD (17 isolates) and Horn Point in Cambridge, MD (14 isolates), with a few from Kent Island in Stevensville, MD (4 isolates), the Baltimore inner harbor (2 isolates), and mid-Chesapeake Bay (2 isolates).

Forty-six and forty-three water, plankton (> 20 µm), oyster, sediment, and plankton-free water (< 20 µm filtrate) were collected from the Chester River and Tangier Sound, respectively, from January 2009 to December 2011. In the Chester River *V. metecus* was detected in at least one of the sample types during 22 of 46 (49%) samplings (Figure 6.3). Chester river water samples showed *V. metecus* densities ranging from < 1 to 155.7 MPN $L^{-1}$, with a mean of 5.4 MPN $L^{-1}$ and standard deviation of 23.7. The highest density was recorded in December, 2009. In the Chester River *V. metecus* was detected in a single oyster sample (February 2009) at a concentration of 36 MPN $L^{-1}$. *V. metecus* was detected in the > 20 µm plankton fraction on ten sampling trips and in the < 20 µm plankton-free fraction 13 times. *V. metecus* was not detected in sediment. Water temperature (surface), salinity (surface), water temperature (bottom), and conductivity

(bottom) were found to be significantly correlated with *V. metecus* by analysis using a multiple regression model (Table 6.1). In a binary logistic regression, salinity (surface and bottom), conductivity (surface) and total dissolved solids (bottom) were predictive for *V. metecus* in water and in all sample types combined, while salinity(surface and bottom) and conductivity (surface) were predictive of *V. metecus* in plankton-free water, and pH (surface) was predictive of *V. metecus* in the plankton fraction.

For Tangier Sound samples *V. metecus* was detected in at least one of the sample types in 13 of 43 (30%) samplings (Figure 6.4). Tangier Sound water samples showed, *V. metecus* densities ranging from < 1 to 155.7 MPN $L^{-1}$, with a mean of 14.5 MPN $L^{-1}$ and standard deviation of 45.8. Highest numbers (155.7 MPN $L^{-1}$) were recorded in December, 2009, and January, February, and March, 2010. In Tangier Sound *V. metecus* was detected in only two oyster samples (December, 2009 and March, 2010) and could be enumerated in December 2009 at 201.1 MPN $L^{-1}$. The *V. metecus*-positive March sample was not detected by MPN assay sample, but in concurrently processed samples for binary analysis (positive or negative for *V. metecus*). *V. metecus* was detected in the > 20 µm plankton fraction during six sample collections and in the < 20 µm plankton-free fraction 10 times, and twice in sediment. Water temperature (bottom), salinity (surface), and pH (surface) were significantly correlated with *V. metecus* in a multiple regression model (Table 6.1). In a binary logistic regression, salinity and conductivity (surface) and salinity and total dissolved solids (bottom) were predictive of *V. metecus* in all sample types combined. Salinity, conductivity, and total dissolved solids (bottom) and salinity and total dissolved solids (bottom) were predictive of *V. metecus* in water samples. Salinity, conductivity, total dissolved solids (surface) and salinity, conductivity, total

dissolved solids, and dissolved oxygen (benthos) were predictive of *V. metecus* in plankton-free water. Water temperature (surface and benthos) were predictive of *V. metecus* in the >20 µm plankton fraction.

## ***Discussion***

A significant number of isolates (13.6%) previously characterized as *V. cholerae* by amplification of the 16S-23S rRNA intergenic spacer unit PCR (Chun et al., 1999) were subsequently identified as *V. metecus* using species-specific PCR indicating that *V. cholerae* is indeed ubiquitous in the Chesapeake Bay, but earlier less sensitive methods of characterization may have overestimated total numbers. *V. metecus* isolates were indistinguishable from *V. cholerae* on TCBS and CHROMagar™ *Vibrio*, two media employed for presumptive differential identification of human pathogenic vibrios, and by standard biochemical profiles used to differentiate *V. cholerae* from other species (data not shown) (Huq et al., 2006). The results of this study suggest that taxonomic identification based on biochemical characteristics alone should be carefully interpreted.

Results of this study show the number of *V. metecus* in the Chesapeake Bay samples varies with season, with the highest numbers in the winter. The trend was consistent at both sampling locations (Figures 6.3 and 6.4). Although *V. metecus* is closely related to *V. cholerae*, a globally distributed bacterium and the causative agent of cholera, the results of this study demonstrate a different seasonal distribution from *V. cholerae* with respect to population size but present throughout the year, including the warmer summer months (Louis et al., 2003; Turner et al., 2009; de Magney et al., 2010). *V. cholerae* can be detected year-round, but its population size is maximum during the

warmer months of the year. *V. cholerae* populations have been linked to zooplankton blooms (Huq et al., 2005), whereas *V. metecus* detection is also not restricted to cold months but its population size is greatest during the winter at both sampling sites and frequently undetectable during warmer months of the year.

In this study, *V. metecus* was not found to be as closely associated with plankton as *V. cholerae*. Copepod densities peak in the spring and summer in the Chesapeake Bay (Louis et al., 2003), and the ability of *V. metecus* to grow at low temperatures certainly is a factor. The vast majority of culturable *V. metecus* collected between 1998 and 1999 were detected in association with 32 plankton samples and only seven water samples. VBNC *V. metecus* may account for this difference in the two sets of results and further work will resolve this apparent anomaly.

A negative correlation between water temperature and *V. metecus* densities was determined statistically, supporting the field observations (Table 6.3). A similar relationship between *V. metecus* and was also observed in which the density of the organism decreases as salinities increase. The data indicate *V. metecus* may prefer freshwater, but viable in saline water, and its presence may be associated with changes in freshwater fronts and changing tidal levels in the bay. Further, this trend of higher densities of *V. metecus* in winter months with presence detected in spring months could also be related to the high riverine discharge that occurs in the Chesapeake Bay due to cycles of high precipitation and snow melt during these months (Schubel and Pritchard, 1986). This annual trend in the Chesapeake Bay coupled with the strong negative association of *V. metecus* with salinity levels in the multiple regression models and binary

logistic regression analyses strongly suggest its movement in the bay may be due to these larger meteorological events.

From a public health perspective, *V. metecus* is an emerging pathogen that causes illness similar to *V. cholerae* non-O1/non-O139, suggesting previously diagnosed *V. cholerae* infections may have been *V. metecus* infections. Further, the ability of genomic islands, namely pathogenicity islands and the CTX phage to be transferred between *V. cholerae* and *V. mimicus* (Boyd et al., 2000), a more distantly related species, suggests the ability of highly pathogenic *V. metecus* cells to arise in the environment. A recent study by Boucher et al. (2011) demonstrated that *V. metecus* frequently exchanges genetic cassettes with co-occurring *V. cholerae* in the environment demonstrating significant genetic flow between them. Thus, *V. metecus* can be considered not only receiving genes from *V. cholerae*, also sharing genes with *V. cholerae, V. mimicus*, *V. parilis*, and other *V. metecus* strains in a local environment.

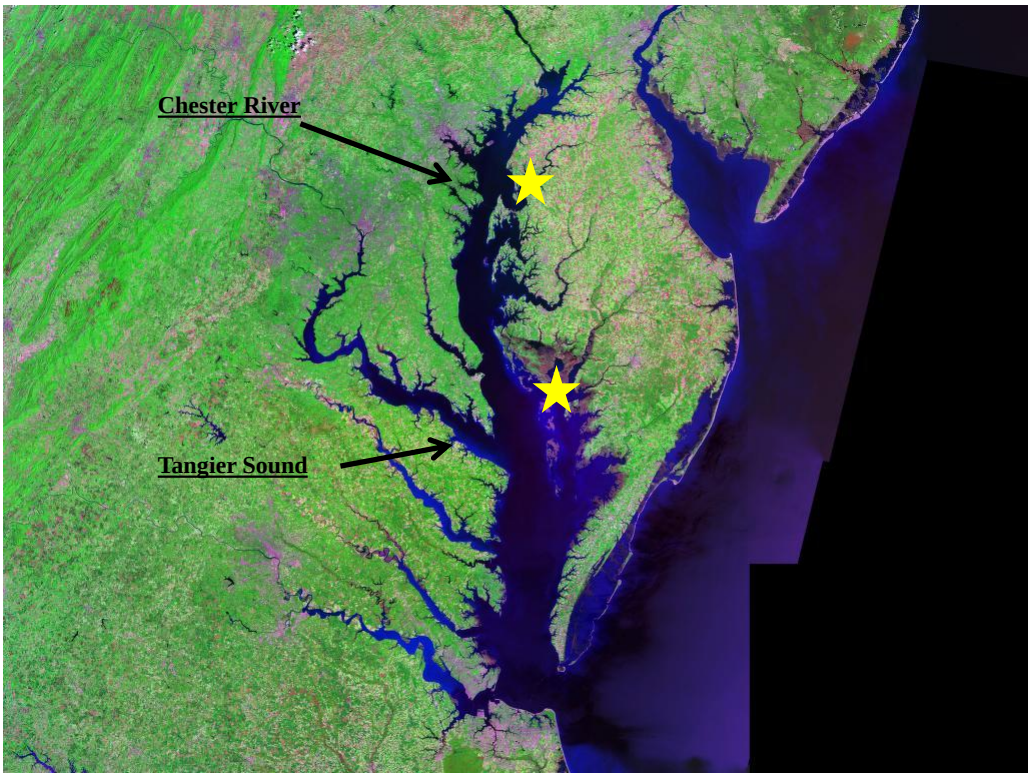



Figure 6.1. Location of sampling stations in the Chesapeake Bay.

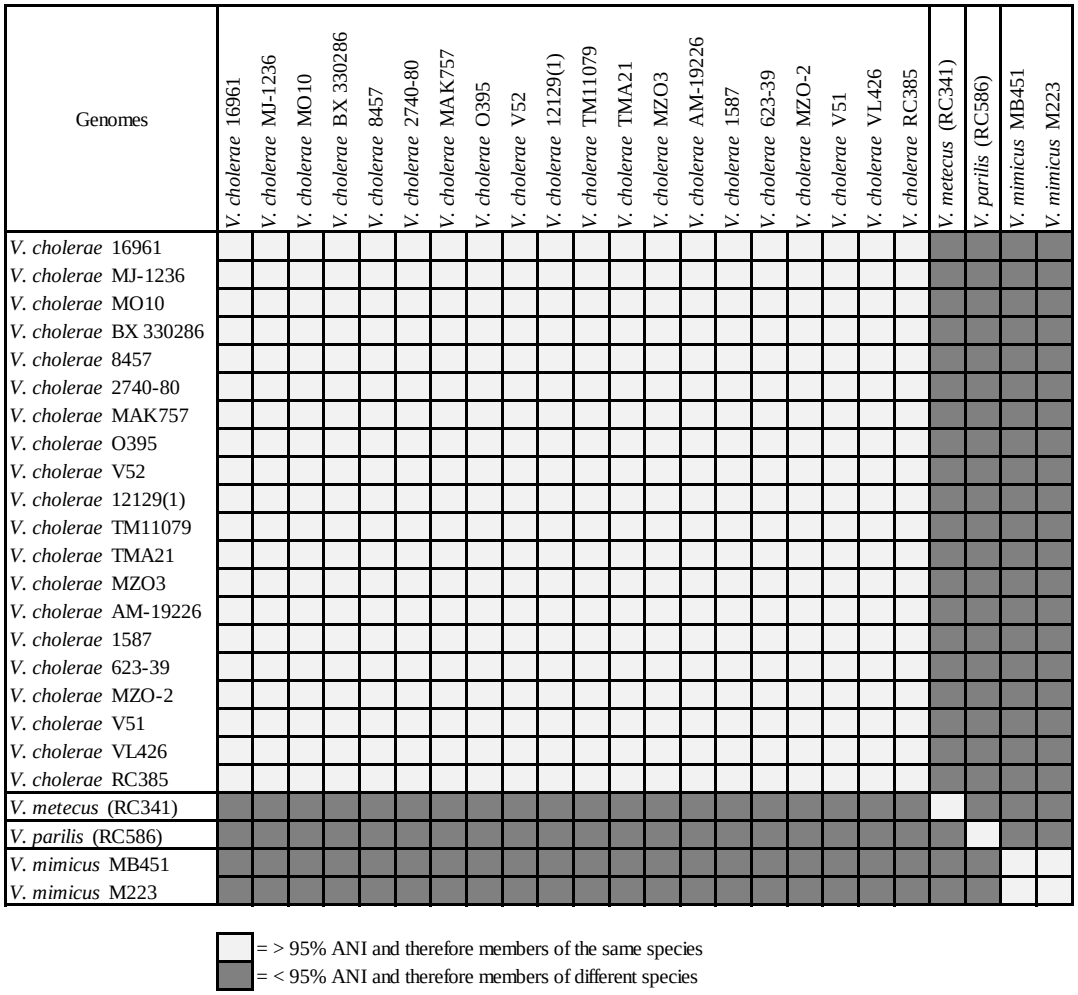



Figure 6.2. Average nucleotide identity matrix of *Vibrionaceae* genomes.

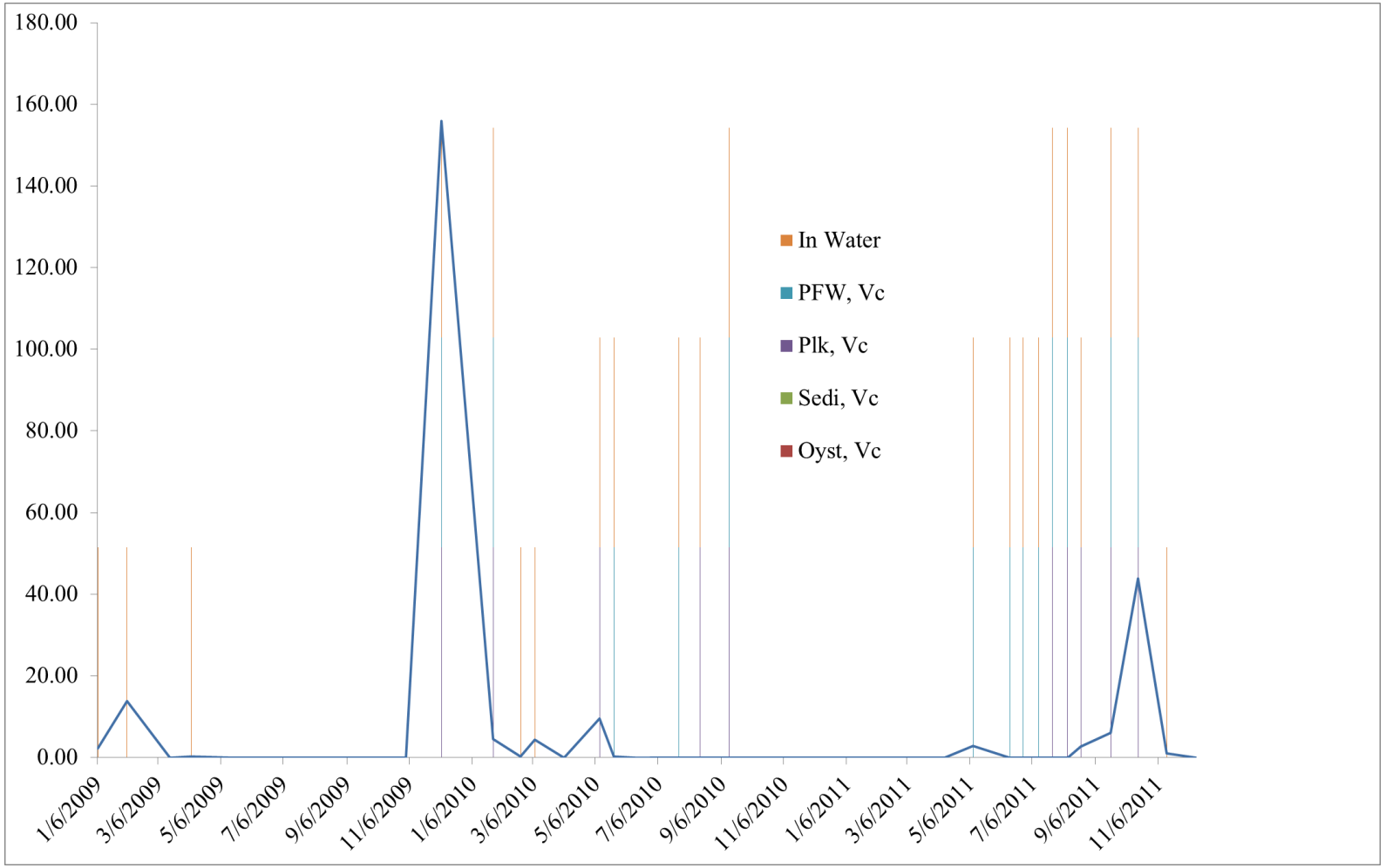

180.00
160.00
140.00
120.00
100.00
80.00
60.00
40.00
20.00
0.00
1/6/2009
3/6/2009
5/6/2009
7/6/2009
9/6/2009
11/6/2009
1/6/2010
3/6/2010
5/6/2010
7/6/2010
9/6/2010
11/6/2010
1/6/2011
3/6/2011
5/6/2011
7/6/2011
9/6/2011
11/6/2011
In Water
PFW, Vc
Plk, Vc
Sedi, Vc
Oyst, Vc

Figure 6.3. Densities and presence/absence of *V.metecus* in the Chester River. Lines show changes in *V. metecus* densities (MPN L-1) in water samples and vertical bars show *V. metecus* presence in sediment, water, oyster, plankton, and plankton-free water samples.

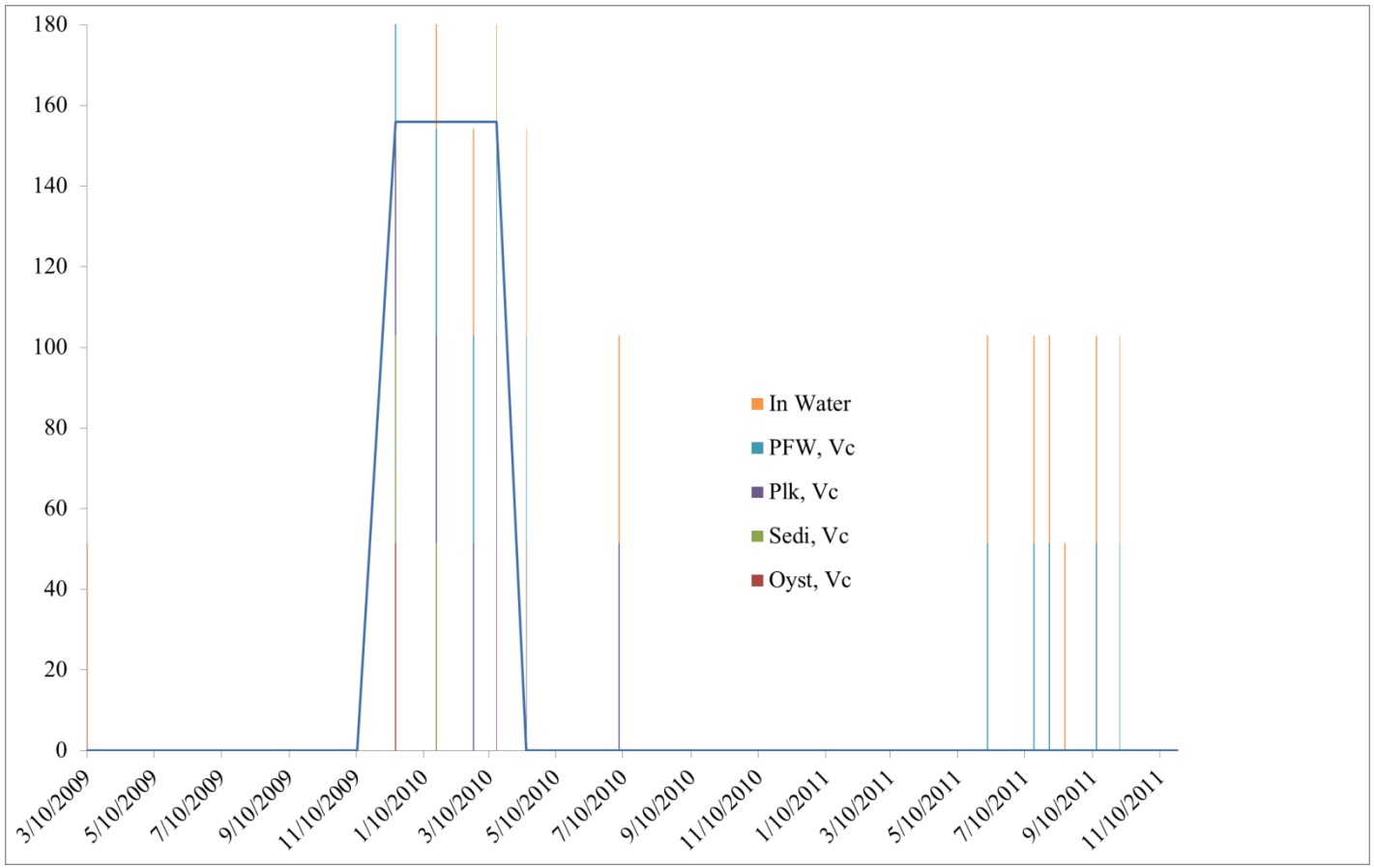

180
160
140
120
100
80
60
40
20
0
3/10/2009
5/10/2009
7/10/2009
9/10/2009
11/10/2009
1/10/2010
3/10/2010
5/10/2010
7/10/2010
9/10/2010
11/10/2010
1/10/2011
3/10/2011
5/10/2011
7/10/2011
9/10/2011
11/10/2011
In Water
PFW, Vc
Plk, Vc
Sedi, Vc
Oyst, Vc

Figure 6.4. Densities and presence/absence of *V.metecus* in Tangier Sound. Lines show changes in *V. metecus* densities (MPN L-1) in water samples and vertical bars show *V. metecus* presence in sediment, water, oyster, plankton, and plankton-free water samples.

Chester River

| Parameter | B | P |
|---|---|---|
| Water Temp (surface) | -0.379 | < 0.05 |
| Dissolved Oxygen (surface) | 0.325 | < 0.05 |
| Water Temp (benthos) | -0.366 | < 0.05 |

Tangier Sound

| Parameter | B | P |
|---|---|---|
| Water Temp (surface) | -0.413 | <0.01 |
| DO (surface) | 0.4 | <0.01 |
| Water Temp (benthos) | -0.413 | <0.01 |
| Dissolved Oxygen (benthos) | 0.452 | <0.01 |

Table 6.1. Spearman rank correlations between *V. metecus* and environmental parameters.

Chester River

| All sample types | B | p |
|---|---|---|
| Salinity (surface) | -0.282 | < 0.05 |
| Conductivity (surface) | -0.169 | < 0.05 |
| Salinity (benthos) | -0.272 | < 0.05 |
| TDS (benthos) | -0.186 | < 0.05 |

| PFW | B | p |
|---|---|---|
| Salinity (surface) | -0.288 | < 0.05 |
| Conductivity (surface) | -0.178 | < 0.05 |
| Salinity (benthos) | -0.298 | < 0.05 |

| PLK | B | p |
|---|---|---|
| pH (surface) | -1.499 | < 0.05 |

Tangier Sound

| All sample types | B | P |
|---|---|---|
| Salinity (surface) | -0.365 | <0.05 |
| TDS (surface) | -0.445 | <0.05 |
| Salinity (benthos) | -0.329 | <0.05 |
| TDS (benthos) | -0.479 | <0.05 |

| Water | B | P |
|---|---|---|
| Salinity (surface) | -0.481 | <0.05 |
| Conductivity (surface) | -0.174 | <0.05 |
| TDS (surface) | -0.599 | <0.05 |
| Salinity (benthos) | -0.427 | <0.05 |
| TDS (benthos) | -0.611 | <0.05 |

| PFW | B | P |
|---|---|---|
| Salinity (surface) | -0.581 | <0.05 |
| Conductivity (surface) | -0.209 | <0.05 |
| TDS (surface) | -0.724 | <0.05 |
| Salinity (benthos) | -0.505 | <0.05 |
| Conductivity (benthos) | -0.183 | <0.05 |
| TDS (benthos) | -0.675 | <0.05 |
| Dissolved Oxygen (benthos) | 0.259 | <0.05 |

| PLK | B | P |
|---|---|---|
| Water Temp (surface) | -0.107 | <0.05 |
| Water Temp (benthos) | -0.111 | <0.05 |

Table 6.2. Binary logistic regression describing association between environmental parameters and presence of *V. metecus* in each sample type. For the Chester River, the frequency of *V. metecus* detection in water was the same as frequency of detection in all sample types.

Chester River

| p (model) | R (model) | Parameter | B | p (var) |
|---|---|---|---|---|
| < 0.01 | 0.601 | Water Temp (surface) | -0.685 | < 0.01 |
| | | Salinity (surface) | -0.647 | < 0.01 |
| | | Water Temp (benthos) | 0.672 | < 0.01 |
| | | Conductivity (benthos) | 0.383 | < 0.01 |

Tangier Sound

| p (model) | R (model) | Parameter | B | p (var) |
|---|---|---|---|---|
| < 0.001 | 0.806 | Salinity (surface) | -0.638 | < 0.01 |
| | | pH (surface) | 0.745 | < 0.01 |
| | | Water Temp (benthos) | -0.046 | < 0.001 |
| | | TDS (benthos) | 0.472 | 0.075 |

Table 6.3. Multiple regression analyses of the association of environmental parameters with *V. metecus*.

# Chapter 7: *Vibrio cholerae* in a Historically Cholera-Free Country


***Abstract***

We report the autochthonous existence of *Vibrio cholerae* in coastal waters of Iceland, a geothermally active country where cholera is absent and has never been reported. Seawater, mussel, and macroalgae samples were collected close to and distant from sites where geothermal activity causes a significant increase in water temperature during low tides. *V. cholerae* was detected only at geothermal-influenced sites during low-tides. None of the *V. cholerae* isolates encoded cholera toxin (*ctxAB*) and all were non-O1/non-O139 serogroups. However, all isolates encoded other virulence factors that are associated with cholera as well as extra-intestinal *V. cholerae* infections. The virulence factors were functional at temperatures of coastal waters of Iceland, suggesting an ecological role. It is noteworthy that *V. cholerae* was isolated from samples collected at sites distant from anthropogenic influence, supporting the conclusion that *V. cholerae* is autochthonous to the aquatic environment of Iceland.


***Introduction***

*Vibrio cholerae*, a Gram-negative bacterium and the causative agent of cholera, has caused seven pandemics since 1816, as well as sporadic inter-epidemic outbreaks. Studies of *V. cholerae* in the environment have often focused on geographic regions of cholera endemicity, in order to elucidate links between its aquatic reservoir and clinical cases of cholera. This has led to the assumption that disease-causing strains of *V.*

*cholerae* are confined to geographical regions where cholera occurs annually or sporadically, that the human gastrointestinal tract is an essential environment for *V. cholerae* presence and dissemination, and that the primary role of virulence factors is infection of the human body. Here we report the presence of *V. cholerae* in Iceland, a geothermally active island in the subarctic north Atlantic with active marine and terrestrial hot-springs, where cholera has never been recorded, and *V. cholerae* has never been isolated from humans (Islam et al., 1993; Gunnardsdóttir, 2008; Haraldur Briem, Chief Epidemiologist, Infectious Disease Control, personal communication). Results demonstrated the presence along the coast of this cholera-free country of genetically diverse *V. cholerae* populations encoding virulence factors known to be integral in human disease, namely cholera. This study demonstrates that *V. cholerae* is autochthonous in a region of the world where cholera never occurs and that the human body is not an obligate environment for the presence and dispersal of this organism. Moreover, virulence factors shown to be conserved in these strains are concluded to be related to the ecology of this bacterium in its native aquatic habitat.

## ***Materials and Methods***

Water, mussel, and macroalgae samples were collected at stations along the coast of Iceland (Figure 7.1). All samples were collected within ca. 5 meters of a geothermal outlet or source (geothermal-influenced) or from a distance of ca. 1 km or greater from a geothermal source (non-geothermal-influenced). Water temperature and salinity were recorded at the time of sample collection. Sampling was not paired at some of the sampling locations because of inclement weather conditions at the sampling sites.

Iceland experiences extreme weather in some locations where sampling sites are located with access difficult. Paired samples were obtained when conditions permitted.

A water sample was collected at Berserkseyri from the pore space of sand collected in the tidal zone at low tide. The hot water source at this site is a natural hot spring, flowing ca. 10 to 30 meters, depending on the tide, where it meets the ocean. A hole was dug into the sand to a depth sufficient to collect pore water (ca. 15 cm).

In the peninsula of Vatnsnes (northern Iceland), samples were collected at the Skarðshver hot spring at two sites where the water temperatures were significantly different and located near the hot spring water flowing to the sea. Water, macroalgae, and mussel samples were collected in sterilized bottles and bags and transported on ice to the laboratory for processing within four hours of collection (Table 7.1).

One liter water samples were filtered through 0.22 µm nitrocellulose membranes. Each membrane was placed in 200 ml of 1% alkaline peptone water (APW) and incubated overnight at 37ºC with shaking at 100 rpm. The meat of each mussel was removed aseptically and 10 g were weighed and diluted in 100 ml of PBS + 2% NaCl, and homogenized in a stomacher (Seward, West Sussex, UK). The homogenate (10 ml) was transferred to a flask with 200 ml of 1% APW and incubated overnight at 37°C with shaking at 100 rpm. Kelp samples were also processed following this procedure. After incubation, 20 µl of the top-most layers of the overnight APW cultures were streaked onto TCBS agar (Sigma-Aldrich) and incubated overnight at 37ºC. Yellow sucrose-fermenting colonies were presumptively identified as *V. cholerae*, asceptically removed from the TCBS agar, and stored at -80°C in 2-ml cryotubes with 1.5-ml Luria-Bertani broth amended with 20% sterile glycerol.

PCR primers used are listed in table 7.3. PFGE was performed using methods developed for *V. cholerae* (Cooper et al., 2006). Sialic acid utilization was determined using the Biolog Phenotype MicroArrays™ (Biolog Inc., Hayward, California). Protreolysis, motility, and hemolysis assays were conducted following standard methods (Son and Taylor, 2011).

The focus the study was both detection of *V. cholerae* in a country where cholera has never occurred and presence of virulence factors in these strains, as well as the ability of the strains to express virulence factors at temperatures of the environment where they were isolated.

## ***Results and Discussion***

### Sampling

One liter water samples were filtered through 0.22 µm nitrocellulose membranes. Each membrane was placed in 200 ml of 1% alkaline peptone water (APW) and incubated overnight at 37ºC with shaking at 100 rpm. The meat of each mussel was removed aseptically and 10 g were weighed and diluted in 100 ml of PBS + 2% NaCl, and homogenized in a stomacher (Seward, West Sussex, UK). The homogenate (10 ml) was transferred to a flask with 200 ml of 1% APW and incubated overnight at 37°C with shaking at 100 rpm. Kelp samples were also processed following this procedure. After incubation, 20 µl of the top-most layers of the overnight APW cultures were streaked onto TCBS agar and incubated overnight at 37ºC. Yellow sucrose-fermenting colonies were presumptively identified as *V. cholerae*, asceptically removed from the TCBS agar,

and stored at -80°C in 2-ml cryotubes with 1.5-ml Luria-Bertani broth amended with 20% sterile glycerol.

Sampling locations chosen near geothermal activity were labeled geothermal-influenced, while locations far from geothermal activity were labeled non-geothermal-influenced (Figure 7.1). At sites positive for *V. cholerae*, all surface water, macroalage, and mussel samples collected were positive for *V. cholerae* (Tables 7.1 and 7.2). *V. cholerae* were recovered at all geothermal-influenced sites from the samples collected at low-tide, but not at any of the non-geothermal-influenced sites or geothermal-influenced sites at high tide, when surface water temperatures are low because of cold seawater inflow (Table 7.1). Friedman's test demonstrated that median surface water temperatures at times of sampling when *V. cholerae* was isolated (median = 27°C, mean = 25.5°C) came from a different distribution than surface water temperatures when *V. cholerae* was not isolated (median = 5.5°C, mean = 6.5°C) ($P < 0.05$). Three of the sites positive for *V. cholerae* are located significantly distant from cities or towns, centers of tourism, or major international shipping routes (Berserkseyri and Stykkishólmur in Breiðafjörður, and Vatnsnes in northern Iceland). Reykjavík harbor (the old port of Reykjavík) and Álftanes and Hliðnes near the shipping port of Hafnarfjörður near Reykjavík, are not geothermal-influenced and those samples were negative for *V. cholerae.*

**Isolate Characterization and Diversity**

In total, 380 presumptive *V. cholerae* isolates were recovered primarily from geothermal-influenced sites. One non-geothermal-influenced site, Hvalfjörður, yielded 22 yellow colonies on TCBS that were determined by PCR not to be *V. cholerae*. Other

non-geothermal-influenced sites yielded no growth on TCBS after incubation of water samples overnight at 37°C and incubation of inoculated TCBS plates at 37°C for 24 to 48 hours. Mesophilic bacteria that grow on TCBS at 37°C are either absent from or in low abundance in these areas or, most likely, may be present in the viable but nonculturable (VBNC) state and not detectable by the methods employed in this study. The remaining 358 isolates were isolated from samples collected at the geothermal-influenced sites and all, except 22 isolates from water samples collected at Stykkishólmur, were presumptively identified as *V. cholerae* by biochemical identification following a previously published protocol (Choopun et al., 2002). A subset (127) of the isolates were subjected to PCR and were confirmed *V. cholerae*. From these results it is concluded that *V. cholerae* was recovered from five of the coastal sampling sites in nine of 19 sample collections that were carried out (Table 7.1).

Confirmation of *V. cholerae* by PCR was achieved by targeting *V. cholerae* intergenic spacer and *toxR* gene (Chun et al., 1999; Vora et al., 2005), and further analysis was done to detect virulence factors and mobile genetic elements. All *V. cholerae* isolates were non-O1/non-O139 serogroups, and *ctxAB* negative (Table 7.2), and did not encode the *Vibrio* seventh pandemic island I (VSP-I) on either chromosome or the *Vibrio* pathogenicity island 1 (VPI-1) also known as the TCP island. All strains were evaluated for presence of the *Vibrio* seventh pandemic island II (VSP-II) by employing the PCR typing scheme for presence of the island and its variants (Taviani et al., 2010). Amplification was not observed for any of the VSP-II variants. However, 50 strains (39%) demonstrated amplification of the 451 bp region that has been reported in all described *V. cholerae* VSP-II variants, except the *V. cholerae* CIRS101 variant. It is

concluded that the Iceland strains encode a VSP-II variant similar to that described in *V. cholerae* RC385, an environmental isolate from the Chesapeake Bay and present in other *V. cholerae* non-O1/non-O139 strains isolated from the mid-Atlantic coast of the United States and from Bangladesh (Taviani et al., 2010). PCR targeting VSP-II flanking regions of the 77 VSP-II-negative isolates did not result in amplification, demonstrating that another genomic island not yet described is inserted at this locus in these isolates.

All but three strains encoded sialidase (NanH) of the *Vibrio* pathogenicity island 2 (VPI-2) that acts as a sialic acid scavenger by cleaving two sialic acid groups from the triasialoganglosides of the intestinal mucus, thereby releasing sialic acid and making the epithelial cell gangliosides in the human gut more accessible to cholera toxin (Moustafa et al., 2004; Almagro-Moreno and Boyd, 2009). We used *nanH* as a marker of VPI-2 which often encodes a suite of sialic acid transport and catabolism genes along with *nanH*, all shown to be expressed in models of *V. cholerae* infections (Almagro-Moreno and Boyd, 2009).

All strains encoded *hlyA*, *rtxA*, HA/P, which are involved in *V. cholerae* virulence in humans (Finkelstein et al., 1992; Olivier et al., 2007). Fifty eight strains (45%) encoded cholix toxin, a novel ADP-ribosylating toxin and 114 (89%) also encoded the integrase of an integrative and conjugative element (ICE). DNA of all strains encoded the hemagglutinin/protease (HA/P), known to be involved in mucin penetration, detachment, spread of the infection through the gastrointestinal tract, and full expression of enterotoxicity (Silva et al., 2006; Shinoda, 2011). This protease has also been shown to be involved in the degradation of chironomid egg masses (Halpern et al., 2003) which inhabit the aquatic environment of the Iceland coast (Ingólfsson, 1995; Sæther, 2009;

Kaiser et al., 2010). Chironomids can inhabit the sandy intertidal zone, where we were able to isolate *V. cholerae* in Berserkseyri.

Diversity of a subset of strains isolated from different locations in Iceland was analyzed by pulsed field gel electrophoresis (PFGE), revealing genomic variability within Icelandic *V. cholerae* populations (Figure 7.4). The genomic patterns did not match those of other environmental non-O1/non-O139 and clinical isolates reported in other regions of the world. Interestingly, two sets of strains from different sampling sites: (1) strain 226 isolated from a sand pore water sample collected at Berserkseyri on 12/08/2006 and strain 334 isolated from mussels in Ægisíða isolated on 2/09/2007; (2) strain 310 isolated from macroalgae in Seltjarnarnes on 1/09/2007 and strain 295 from macroalgae in Ægisíða, on 1/09/2007 yielded identical banding patterns. These two patterns suggest genomic clonality of a subset of strains circulating around Iceland between regions of geothermal activity. However, no other conserved set of PFGE patterns was observed among the other Iceland strains, indicating significant diversity of the *V. cholerae* strains.

Phenotypic diversity among seven strains randomly selected for analysis was observed at 25 and 34°C, showing different patterns of carbon substrate utilization on Biolog PM2A plates (Figure 7.3). Results of other phenotypic tests showed only one strain to be bioluminescent, a phenotype encoded in GI-64 of the global *V. cholerae* mobilome reported by (Chun et al., 2009). Antibiotic disk diffusion assays of a subset of 44 strains showed all were susceptible to chloramphenicol (30 μg), ciprofloxacin (5 μg), vibriostatic agent O/129 (150 μg), streptomycin (10 μg), sulfamethoxazole/trimethoprim (23.75/1.25 μg), and tetracycline (30 μg), while 25% were resistant to ampicillin (10 μg). These results demonstrate genetic and phenotypic diversity of the *V. cholerae* populations

of the Iceland coast, further evidence that *V. cholerae* in Iceland is neither imported nor derived from a single source. Moreover, similarity in some of the features suggests that strains of *V. cholerae* can circulate spatially along the coast of Iceland.

**Expression of Virulence Factors**

The role of virulence factors encoded in the *V. cholerae* strains was evaluated by expression *in vitro* at the temperatures of the geothermal-influenced environmental sampling sites. The phenotypes of 44 strains at were tested at 4, 14, 25, and 34, and 50°C. Motility, previously reported to be associated with virulence of *V. cholerae* virulence (Guentzel and Berry, 1975; Watnick et al., 2001; Krukonis and DiRita, 2003) was tested by inoculating motility media incubated for 24 hours. Those incubated at $\leq$ 14°C were evaluated after 96 hours. Motility was most active at 34°C (mean = 12.5 mm), followed by 25°C (mean = 6.25 mm), with no motility observed at temperatures $\geq$50°C or $\leq$ 14°C (Figure 7.4). Hemolysis was detected on blood agar plates incubated at 34°C (mean = 3 mm) and weak hemolysis at 25°C (mean =1 mm). Growth without hemolysis was observed for the blood agar plates incubated at 14°C. Neither hemolysis nor growth was observed for plates incubated at 50°C or 4°C.

The sialic acid utilization cluster of VPI-2 was studied using seven randomly selected strains. These were inoculated in a medium containing sialic acid as sole carbon source and all utilized sialic acid when incubated at 34, 25, and 14°C. Sialic acid utilization was estimated as follows: area under the curve $_{\text{Strain}}$ / area under the curve $_{\text{Background}}$ in a Biolog assay. Utilization was higher at 34°C (4.12) than at 25°C (3.57) (Table7.4). Quantitative evaluation at of sialic acid utilization at 14°C could not be

measured since this temperature falls below the Biolog temperature range. Plates with sialic acid as a carbon source were inoculated and incubated at 14°C, with end point color change recorded at 72 hours compared with a negative control. All strains tested utilized sialic acid at 14°C. Since marine bivalves contain free sialic acid in their hemolymph , this pathogenicity island may be instead ecologically significant in utilization of sialic acid of the mussels with which the bacterium is associated.

The *V. cholerae* strains were actively proteolytic at 14, 25 and 34°C, with the largest zones of proteolysis on the milk agar plates incubated at 34°C (6.1 mm). Milk agar plates inoculated with the *V. cholerae* isolates and incubated at 4°C for 2 weeks were also positive for proteolysis. To confirm proteolysis aliquots of overnight cultures were centrifuged at 20,000 x g for 20 minutes and the supernatant was removed and placed on ice for 1 hour, after which 10 µl was transfered to the top layer of chilled milk agar and incubated at 4°C for two weeks. Proteolysis was observed, with proteolysis as a virulence factor being functional at environmental temperatures relative to Iceland, a cholera-free region.

## ***Conclusions***

It is concluded that *V. cholerae* is naturally occurring and readily isolated from environmental samples collected along the coast of Iceland, a country where cholera has never been reported. *V. cholerae* was readily culturable in areas of geothermal activity where water temperature was elevated, but *in vitro* growth experiments did not demonstrate growth of these isolates at 50°C, i.e., the bacteria are not hyperthermophilic. We hypothesize that *V. cholerae* is present in the VBNC state in areas that are not

geothermal-influenced or when tides are high (and the water temperature is low) near sites of geothermal activity. *V. cholerae* in the mesophilic layer resulting from mixing of hot and cold water remained culturable. A mesophilic layer is present at all times at sites of geothermal-influence but during high tide it is overlaid with cold sea water. The similarity of the PFGE patterns of *V. cholerae* isolated from different locations in Iceland indicates circulation of *V. cholerae* strains around coastal Iceland that become culturable with temperature upshift of geothermal heated water. Such temperature upshifts have been shown to resuscitate VBNC cells of *V. vulnificus* to the culturable state (Oliver, 2005) and very likely *V. cholerae* (Chaiyanan et al., 2007).

DNA of *V. cholerae* isolated during this study were found to encode many of the virulence factors associated with intestinal and extraintestinal infections caused by *V. cholerae*. Icelanders have very little exposure to the usual pathways of cholera transmission, since municipal water is supplied by groundwater recharged by precipitation or glacial melt and uncooked molluscan shellfish, a common vehicle of vibrioses in developed countries, is not typically consumed in Iceland (Petursson, 1968; Guðfinnsson, 2007).

Historically, Iceland has experienced both sporadic cases and epidemics of other infections (plague, smallpox, and influenza) evidence that it is not isolated from pathogens that are global in their epidemiology (Hjaltelin, 1871; Karlsson, 1996; Dowell and Bresee, 2008, Cliff et al., 2009; Sigurdsson et al., 2009). It is one of the very few countries never to have recorded even a single case of cholera or related *Vibrio* infection (Islam et al., 1993). The absence of any record of cholera, a reportable disease in Iceland (Gunnarsdóttir, 2008; Icelandic Directorate of Health), combined with the extensive

health record keeping for all citizens, with the presence of potentially virulent *V. cholerae* strains in the most remote locations in Iceland, suggest *V. cholerae* is autochthonous to Iceland. The ecology of *V. cholerae*, as has been documented elsewhere by other investigators does not require human transmission for persistence (Kenyon et al., 1984; Louis et al., 2003; Schuster et al., 2011). However, sporadic cases of cholera or *V. cholerae* infections in or near areas of cholera outbreaks or ballast water exchange in shipping ports where cholera has occurred has been suggested to be a source of *V. cholerae*. To our knowledge this is the first report of *V. cholerae* with functional virulence factors isolated from a region where a case of cholera has never occurred. It is concluded that *V. cholerae* was not introduced to Iceland, i.e. from ballast water or infected persons, but rather is a component of its natural ecological and microbiological environment. The results further suggest global distribution of *V. cholerae* in the aquatic environment without the necessity of anthropogenic activity as a source. The presumed absence of *V. cholerae* in regions where cholera has never occurred or has not occurred for many years earlier, is most likely because the presence of *V. cholerae* has not been recognized as naturally occurring in the aquatic environment and, therefore, not monitored.

Both presence of the microorganism and the high degree of conservation of virulence factors where cholera has not been documented strongly indicate these factors have an ecological function other than pathogenicity for humans. The results of this study fully support the autochthonous nature of this bacterium in aquatic systems on a global scale and not its confinement to those regions where cholera or *V. cholerae* infections are endemic or sporadic or only to warmer tropical and subtropical regions of

the world. Clearly, human intestinal amplification and shedding is not required for the presence of *V. cholerae* in the natural aquatic environment and conservation of virulence factors does not serve solely to maintain cell viability between human infections, but rather they play a role in the natural ecology of *V. cholerae*.

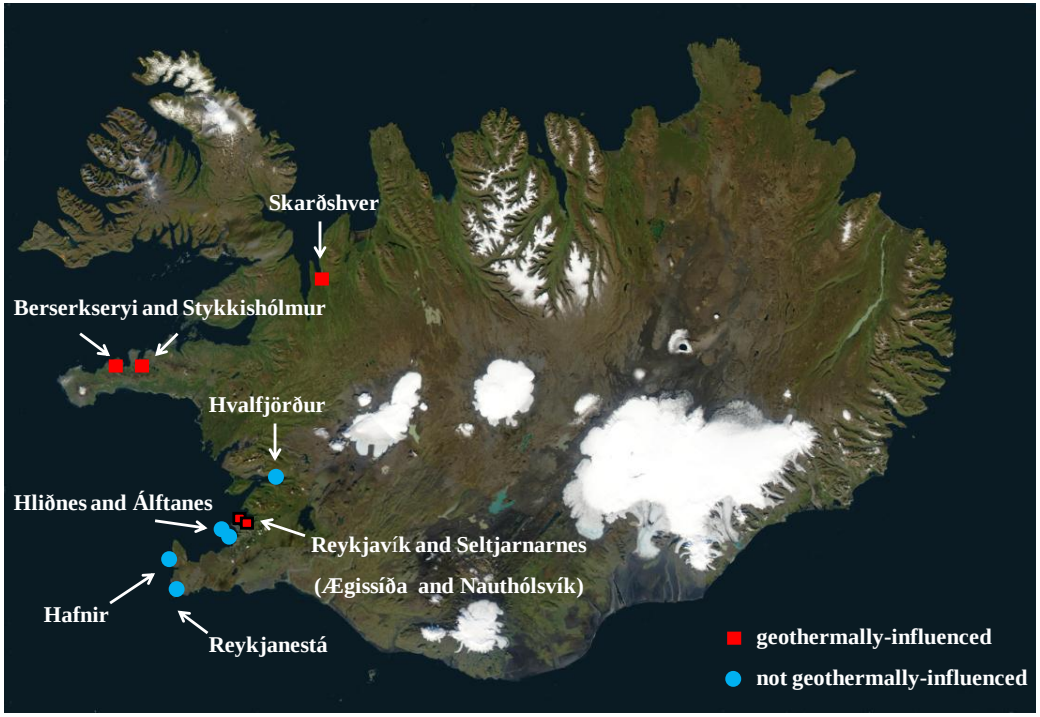



Figure 7.1. Sampling sites showing presence and absence of geothermal activity. Ægissíða and Nauthólsvík are both located within the city limits of Reykjavík. Symbols that correspond to Reykjavik are bordered in black. Seltjarnarnes is a suburb of Reykjavik. Water, mussel, and macroalgae samples were collected at stations along the coast of Iceland (Figure 7.1). Photo Credit: National Aeronautics and Space Administration (NASA). Use of this image is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license (http://commons.wikimedia.org/wiki/File:Iceland_sat_cleaned.png).

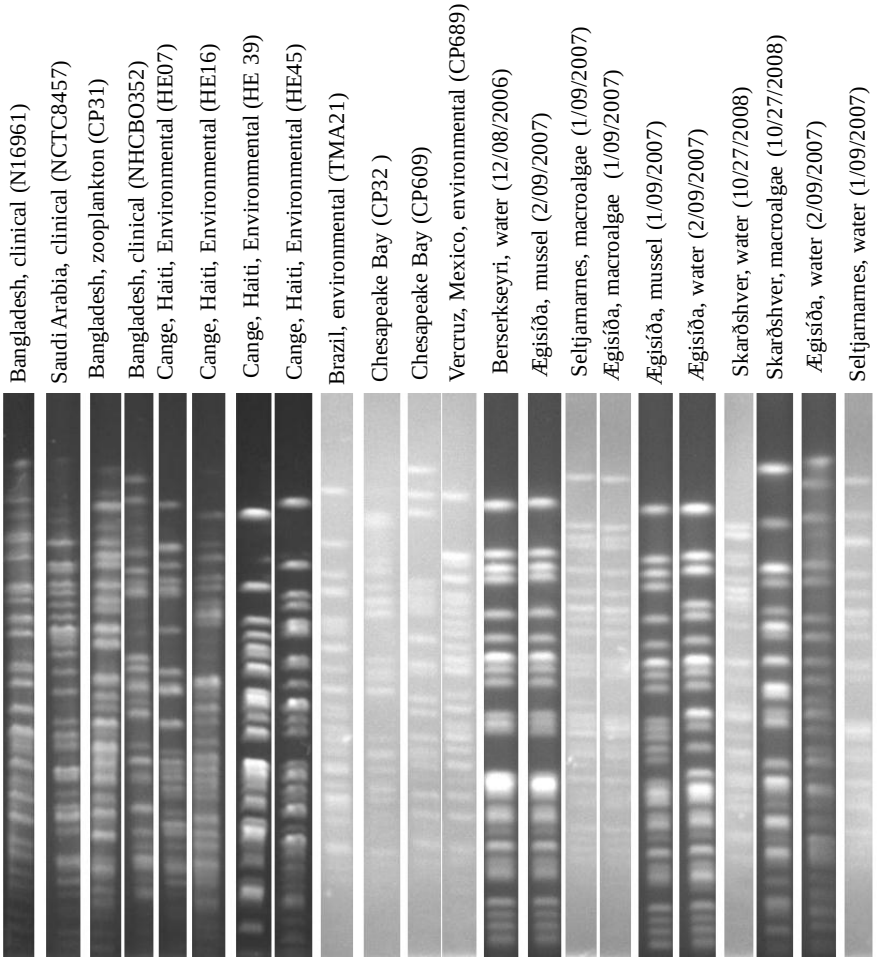



Figure 7.2. Pulse-Field Gel Electrophoresis images of strains collected along the coast of Iceland as well as globally distributed clinical and environmental strains. PFGE was performed using methods developed for *V. cholerae* (Cooper et al., 2006).

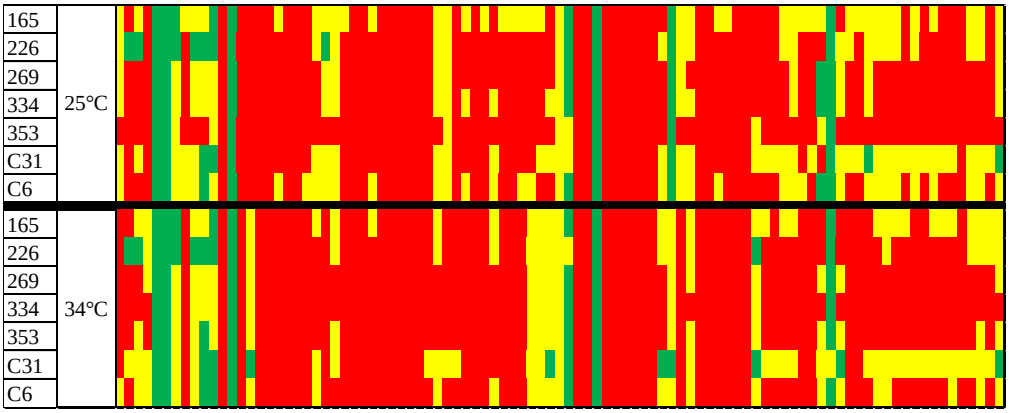



Figure 7.3. Carbon utilization patterns of 7 randomly selected *V. cholerae* strains using Biolog PM2A plate, Biolog Phenotype MicroArrays™ (Biolog Inc., Hayward, California). Strain ID is listed in far left column. Red = area under curve divided by background < 1. Yellow = area under curve divided by background 2 < 1. Green = area under curve divided by background > 2. Top row is carbon utilization profile at 25°C, bottom row is carbon utilization at 34°C.

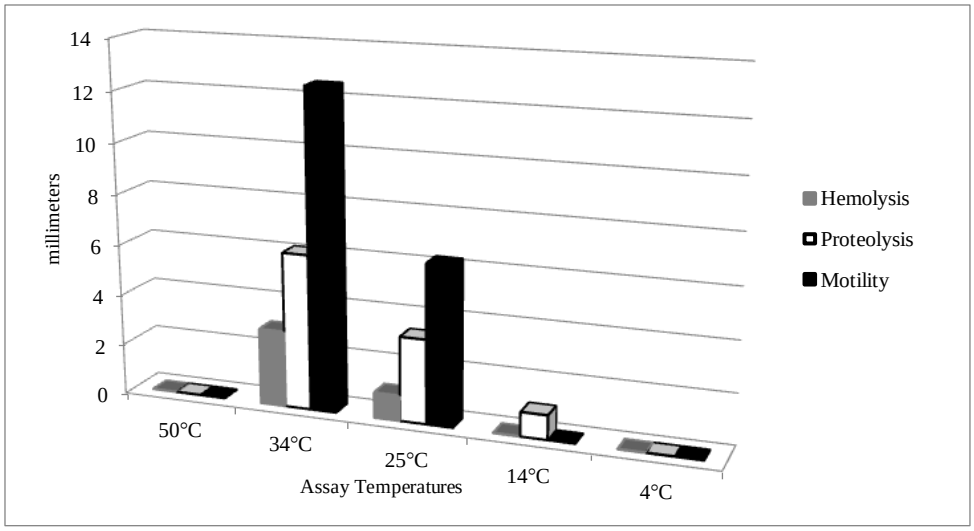



Figure 7.4. Expression of virulence factors on agar plates. Columns show the average expression; zones of hemolysis, proteolysis, and motility on blood, milk, and motility agar plates, at different temperatures, of strains isolated from different locations along the coast of Iceland (n = 44).

| Site | Temp (°C) | Salinity | Month | Year | Geothermally Influenced | Tide | Sample | *V. cholerae* Detected |
|---|---|---|---|---|---|---|---|---|
| Hvalfjörður | 6 | NR | Oct | 2006 | - | Low | Ma, W, M | - |
| Ægissíða | 5 | NR | Oct | 2006 | + | High | W | - |
| Hliðnes | 7 | NR | Oct | 2006 | - | Low | W | - |
| Álftanes | 7 | NR | Oct | 2006 | - | Low | W | - |
| Nauthólsvík | 7 | NR | Oct | 2006 | + | High | W | - |
| Seltjarnarnes | NR | NR | Nov | 2006 | + | High | Ma, W, M | - |
| Reykjanestá | 2 | 35 | Dec | 2006 | - | Low | Ma, W | - |
| Hafnir | 6 | 32 | Dec | 2006 | - | Low | Ma, W | - |
| Reykjavík Harbor | 8 | 30 | Apr | 2007 | - | High | W, M | - |
| Nauthólsvík | 11 | NR | Sept | 2008 | + | High | W | - |
| Stykkishólmur | 21 | 20 | Dec | 2006 | + | Low | Ma, W | + |
| Berserkseyri | 17 | 6 | Dec | 2006 | + | Low | W | + |
| Ægissíða | 26 | 7 | Jan | 2007 | + | Low | Ma, W, M | + |
| Seltjarnarnes | 18 | 20 | Jan | 2007 | + | Low | Ma, W | + |
| Ægissíða | 31 | 2 | Feb | 2007 | + | Low | Ma, W, M | + |
| Seltjarnarnes | 32 | 15 | Feb | 2007 | + | Low | W | + |
| Ægissíða | 34 | NR | Sept | 2008 | + | Low | Ma, W, M | + |
| Seltjarnarnes | 28 | NR | Sept | 2008 | + | Low | Ma, W | + |
| Skarðshver | 34 | NR | Nov | 2008 | + | Low | Ma, W | + |
| Skarðshver | 14 | NR | Nov | 2008 | + | Low | Ma, W | + |

NR = not recorded
Ma = macroalgae
W= water
M = mussel
+ = yes
- = no

Table 7.1. Sample sites, water temperature, salinity, presence of geothermal activity, tidal height, type of sample collected (water, kelp, mussels), and presence of *V. cholerae* are presented.

| Date | Site | Source | Strains | O1/O139 | *ctxA* | *ctxB* | *ace* | *zot* | *chxA* (cholix) | *toxR* | *ompU* | *hlyA* | *rtxA* | HA/P | *luxO* | *tcpA* | nag-ST | *nanH* | ICE | VSP-II | VSP-I |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 8 Dec 2006 | Stykkishólmur | Water | 10 | 0 | 0 | 0 | 0 | 0 | 10 (100) | 10 (100) | 10 (100) | 10 (100) | 10 (100) | 10 (100) | 10 (100) | 0 | 0 | 8 (80) | 10 (100) | 7(70) | 0 |
| 8 Dec 2006 | Stykkishólmur | Macroalgae | 10 | 0 | 0 | 0 | 0 | 0 | 10 (100) | 10 (100) | 10 (100) | 10 (100) | 10 (100) | 10 (100) | 10 (100) | 0 | 0 | 10 (100) | 10 (100) | 6 (60) | 0 |
| 8 Dec 2006 | Berserksey ri | Water | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 10 (100) | 10 (100) | 10 (100) | 10 (100) | 10 (100) | 10 (100) | 0 | 0 | 10 (100) | 10 (100) | 7(70) | 0 |
| 9 Jan 2007 | Ægisíða | Water | 11 | 0 | 0 | 0 | 0 | 0 | 0 | 11 (100) | 11 (100) | 11 (100) | 11 (100) | 11 (100) | 11 (100) | 0 | 0 | 11 (100) | 11 (100) | 0 | 0 |
| 9 Jan 2007 | Ægisíða | Mussel | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 8 (100) | 8 (100) | 8 (100) | 8 (100) | 8 (100) | 8 (100) | 0 | 0 | 8 (100) | 8 (100) | 2(25) | 0 |
| 9 Jan 2007 | Ægisíða | Macroalgae | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 8 (100) | 8 (100) | 8 (100) | 8 (100) | 8 (100) | 8 (100) | 0 | 0 | 8 (100) | 8 (100) | 1 (12) | 0 |
| 9 Jan 2007 | Seltjarnarnes | Macroalgae | 9 | 0 | 0 | 0 | 0 | 0 | 1 (11) | 9 (100) | 9 (100) | 9 (100) | 9 (100) | 9 (100) | 9 (100) | 0 | 0 | 9 (100) | 9 (100) | 4 (44) | 0 |
| 9 Jan 2007 | Seltjarnarnes | Water | 8 | 0 | 0 | 0 | 0 | 0 | 3 (37) | 8 (100) | 8 (100) | 8 (100) | 8 (100) | 8 (100) | 8 (100) | 0 | 0 | 8 (100) | 8 (100) | 3 (37) | 0 |
| 9 Feb 2007 | Ægisíða | Mussel | 11 | 0 | 0 | 0 | 0 | 0 | 5 (45) | 11 (100) | 11 (100) | 11 (100) | 11 (100) | 11 (100) | 11 (100) | 0 | 0 | 11 (100) | 11 (100) | 4 (36) | 0 |
| 9 Feb 2007 | Ægisíða | Water | 6 | 0 | 0 | 0 | 0 | 0 | 3 (50) | 6 (100) | 6 (100) | 6 (100) | 6 (100) | 6 (100) | 6 (100) | 0 | 0 | 6 (100) | 6 (100) | 2 (33) | 0 |
| 9 Feb 2007 | Ægisíða | Macroalgae | 4 | 0 | 0 | 0 | 0 | 0 | 2 (50) | 4 (100) | 4 (100) | 4 (100) | 4 (100) | 4 (100) | 4 (100) | 0 | 0 | 4 (100) | 4 (100) | 1 (25) | 0 |
| 9 Feb 2007 | Seltjarnarnes | Water | 9 | 0 | 0 | 0 | 0 | 0 | 7 (77) | 9 (100) | 9 (100) | 9 (100) | 9 (100) | 9 (100) | 9 (100) | 0 | 0 | 9 (100) | 7 (77) | 2 (22) | 0 |
| 4 Sep 2008 | Ægisiða | Water | 5 | 0 | 0 | 0 | 0 | 0 | 1 (20) | 5 (100) | 5 (100) | 5 (100) | 5 (100) | 5 (100) | 5 (100) | 0 | 0 | 5 (100) | 5 (100) | 0 | 0 |
| 4 Sep 2008 | Ægisiða | Mussel | 2 | 0 | 0 | 0 | 0 | 0 | 2 (100) | 2 (100) | 2 (100) | 2 (100) | 2 (100) | 2 (100) | 2 (100) | 0 | 0 | 2 (100) | 0 | 1 (50) | 0 |
| 4 Sep 2008 | Ægisiða | Macroalgae | 2 | 0 | 0 | 0 | 0 | 0 | 1 (50) | 2 (100) | 2 (100) | 2 (100) | 2 (100) | 2 (100) | 2 (100) | 0 | 0 | 2 (100) | 0 | 0 | 0 |
| 24 Sep 2008 | Seltjarnarnes | Water | 2 | 0 | 0 | 0 | 0 | 0 | 2 (100) | 2 (100) | 2 (100) | 2 (100) | 2 (100) | 2 (100) | 2 (100) | 0 | 0 | 2 (100) | 0 | 1 (50) | 0 |
| 24 Sep 2008 | Seltjarnarnes | Macroalgae | 5 | 0 | 0 | 0 | 0 | 0 | 5 (100) | 5 (100) | 5 (100) | 5 (100) | 5 (100) | 5 (100) | 5 (100) | 0 | 0 | 4 (80) | 3 (60) | 4 (80) | 0 |
| 27 Oct 2008 | Skarðshv er | Water | 4 | 0 | 0 | 0 | 0 | 0 | 4 (100) | 4 (100) | 4 (100) | 4 (100) | 4 (100) | 4 (100) | 4 (100) | 0 | 0 | 4 (100) | 3 (75) | 3 (75) | 0 |
| 27 Oct 2008 | Skarðshv er | Macroalgae | 2 | 0 | 0 | 0 | 0 | 0 | 2 (100) | 2 (100) | 2 (100) | 2 (100) | 2 (100) | 2 (100) | 2 (100) | 0 | 0 | 2 (100) | 1 (50) | 2 (100) | 0 |
| 27 Oct 2008 | Skarðshv er | Water | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 (100) | 1 (100) | 1 (100) | 1 (100) | 1 (100) | 1 (100) | 0 | 0 | 1 (100) | 0 | 0 | 0 |
| | | | 127 | 0 | 0 | 0 | 0 | 0 | 58 (45) | 127 (100) | 127 (100) | 127 (100) | 127 (100) | 127 (100) | 127 (100) | 0 | 0 | 124 (97) | 114 (89) | 50 (39) | 0 |

Table 7.2. Results of PCR analysis of *V. cholerae* isolates collected during this study.

| Target | | Forward Primer | Reverse Primer | Reference |
|---|---|---|---|---|
| VSP-I | *Vibrio* Seventh Pandemic Island I | GCCGAGAACTCTAAAGCGCTTCTC | CCAAGGTACAGATGAGTACCAGCA | (Grim et al., 2010) |
| VSP-I insertion site (Chr I) | | AAACTGGCGACCTTTGAGCAAGC | GATGGTAGCCTGACGCTGCATCTG | |
| VSP-I insertion site (Chr II) | | ATAGCGGGAGTTGGCTCTGCA | GGTGACTTGGTGCCCATCGTA | |
| pVSP2-I | *Vibrio* Seventh Pandemic Island II | CACCTGTCATGTTATGAGGTGCA | AACAGGTCTCTTATCGGCTTTGC | (Taviani et al., 2010) |
| pVSP2-II | | GCACAACTTGTAAGATAGCCTTGC | ACGCAAGACAAAACTACAGCTTGC | |
| pVSP2-III | | CCAGCAAACGGTCATTCGCT | TGGTTGGAAGGTGGGTTGTGT | |
| VSP-II insertion site | | AGATCAACTACGATCAAGCC | CGCAGTCACAGCTTAAAC | (O'Shea et al., 2004) |
| O1 | O1 antigen | GTTTCACTGAACAGATGGG | GGTCATCTGTAAGTACAAC | (Hoshino et al., 1998) |
| O139 | O139 antigen | AGCCTCTTTATTACGGGTGG | GTCAAACCCGATCGTAAAGG | |
| *ctxA* | A subunit of cholera toxin | ACAGAGTGAGTACTTTGACC | ATACCATCCATATATTTGGGAG | (Zhu et al., 2007) |
| *ctxAB* | CTXΦ | AGTCAGGTGGTCTTATGCC | TTGCCATACTAATTGCGG | |
| HA/protease | Hemagglutinin/Protease | ACGTTAGTGCCCATGAGGTC | ACGGCAAACACTTCAAAACC | Our Laboratory |
| nag-ST | non-agglutinating heat stable toxin | CAATCGCATTTAGCCAAACA | GCAAGCTGGATTGCAACATA | Our Laboratory |
| *chxA* | cholix toxin | TGGTGAAGATTCTCCTGCAA | CTTGGAGAAATGGATGCGCTG | (Purdy et al. 2010) |
| *tcpA* | A subunit of toxin co-regulated pilus | CACGATAAGAAAACCGGTCAAGAG | CGAAAGCACCTTCTTTCACGTTG<br>TTACCAAATGCAACGCCGAATG | (Rivera et al., 2001) |
| *ace* | accessory cholera enterotixin | TGATGGCTTTACGTGGCTTGTGATC | GCCTGTTGGATAAGCGGATAGATGG | |
| *zot* | zona occludens toxin | ATCTGCCTAACCACGCCTAACATTG | ACCGCCTTGCTCCCGACAG | |
| *toxR* | global regulator of virulence | ACCGCAGCCAGCCAATGTTG | TGGCAATGACTTCTATCGGCTTGAG | (Vora et al., 2006) |
| *ompU* | outer membrane protein U | TACGCTGGTATCGGTGGCACTTAC | TCCATGCGGTAAGAAGCGGCTAG | |
| *rtxA* | repeat in toxin | CTGAATATGAGTGGGTGACTTACG | GTGTATTGTTCGATATCCGCTACG | |
| *luxO* | global regulatory gene | CGCTGTATCGTTCTTACCTCACACC | GCTCGCCGCAGAGTCAATGG | |
| *nanH* | sialidase | CTTCCTCCAATACGGTTCTTGTCTCTTATGC | TTCGGCTACCATCGGCAACTTGTATC | |
| *hlyA* | hemolysin A | GGCAAACAGCGAAACAAATACC<br>GAGCCGGCATTCATCTGAAT | CTCAGCGGGCTAATACGGTTTA | (Rivera et al., 2001) |
| ICE | integrative and conjugative element | GCTGGATAGGTTAAGGGCGG | CTCTATGGGCACTGTCCACATTG | (Hochhut et al., 2001) |

Table 7.3. PCR primers used in the study.

| Assay Temperature (°C) | Hemolysis (mm) | Proteolysis (mm) | Motility (mm) | Sialic Acid Metabolism |
|---|---|---|---|---|
| 50 | 0 | 0 | 0 | ND |
| 34 | 3 | 6.11 | 12.52 | 4.12 A |
| 25 | 1 | 3.32 | 6.25 | 3.57 |
| 14 | 0 | 0.95 | 0 | + |
| 4 | 0 | weak | 0 | ND |

A = area under the curve $_{Strain}$ / area under the curve $_{Background}$

ND = not done

Table 7.4. Results of virulence assays. Numbers in columns show average results in millimeters (n=44 for hemolysis, proeteolysis, and motility assays).

*Reproduction License Agreement*

Mar 24, 2012

John Wiley and Sons License Terms and Conditions

---

---

This is a License Agreement between Bradd J Haley ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

**All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.**

License Number

2873811187878

License date

Mar 21, 2012

Licensed content publisher

John Wiley and Sons

Licensed content publication

Environmental Microbiology Reports

Licensed content title

Vibrio cholerae in a historically cholera-free country

Licensed content author

Bradd J. Haley,Arlene Chen,Christopher J. Grim,Philip Clark,Celia Municio Diaz,Elisa Taviani,Nur A. Hasan,Elizabeth Sancomb,Wessam Mahmoud Elnemr,Muhammad A. Islam,Anwar Huq,Rita R. Colwell,Eva Benediktsdóttir

Licensed content date

Apr 1, 2012

Start page

no

End page

no

Type of use

Dissertation/Thesis

Requestor type

Author of this Wiley article

Format

Electronic

Portion

Full article

Will you be translating?

No

Order reference number

VIBRIOICELAND

Total

0.00 USD

Terms and Conditions

**TERMS AND CONDITIONS**

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or a society for whom a Wiley

Company has exclusive publishing rights in relation to a particular journal (collectively WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your Rightslink account (these are available at any time at http://myaccount.copyright.com)

Terms and Conditions

1. The materials you have requested permission to reproduce (the "Materials") are protected by copyright.

2. You are hereby granted a personal, non-exclusive, non-sublicensable, non-transferable, worldwide, limited license to reproduce the Materials for the purpose specified in the licensing process. This license is for a one-time use only with a maximum distribution equal to the number that you identified in the licensing process. Any form of republication granted by this licence must be completed within two years of the date of the grant of this licence (although copies prepared before may be distributed thereafter). The Materials shall not be used in any other manner or for any other purpose. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Material. Any third party material is expressly excluded from this permission.

3. With respect to the Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Materials without the prior permission of the respective copyright owner. You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Materials, or any of the rights granted to you hereunder to any other person.

4. The Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc or one of its related companies (WILEY) or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any

such right, license or interest with respect thereto.

5. NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

6. WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

7. You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

8. IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

9. Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

10. The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to

any other or subsequent breach by such other party.

11. This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

12. Any fee required for this permission shall be non-refundable after thirty (30) days from receipt.

13. These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

14. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

15. WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

16. This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

17. This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

**Wiley Open Access Terms and Conditions**

All research articles published in Wiley Open Access journals are fully open access: immediately freely available to read, download and share. Articles are published under the terms of the Creative Commons Attribution Non Commercial License. which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes. The license is subject to the Wiley Open Access terms and conditions:
Wiley Open Access articles are protected by copyright and are posted to repositories and

 Linking to article content in e-mails redistributed for promotional, marketing or educational purposes;
 Use for the purposes of monetary reward by means of sale, resale, licence, loan, transfer or other form of commercial exploitation such as marketing products
 Print reprints of Wiley Open Access articles can be purchased from: corporatesales@wiley.com

Other Terms and Conditions:

BY CLICKING ON THE "I AGREE..." BOX, YOU ACKNOWLEDGE THAT YOU HAVE READ AND FULLY UNDERSTAND EACH OF THE SECTIONS OF AND PROVISIONS SET FORTH IN THIS AGREEMENT AND THAT YOU ARE IN AGREEMENT WITH AND ARE WILLING TO ACCEPT ALL OF YOUR OBLIGATIONS AS SET FORTH IN THIS AGREEMENT.

v1.7

**If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK500744278.**
**Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.**

**Make Payment To:**
**Copyright Clearance Center**
**Dept 001**
**P.O. Box 843006**
**Boston, MA 02284-3006**

**For suggestions or comments regarding this order, contact RightsLink Customer Support: customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.**

**Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.**

# Chapter 8: Molecular Diversity and Predictability of *Vibrio parahaemolyticus* in the Black Sea


***Abstract***

*Vibrio parahaemolyticus* not only is a leading cause of seafood-related gastroenteritis but is also an autochthonous member of the marine and estuarine environment worldwide. One-hundred seventy strains of *V. parahaemolyticus* were isolated from water and plankton samples collected along the Georgian coast of the Black Sea during three years of monthly sample collection. Isolates were identified as *V. parahaemolyticus* by PCR detection of the *toxR* and *tlh* genes. All confirmed isolates were tested by PCR for the thermostable direct hemolysin (*tdh*) and thermostable-related hemolysin (*trh*). A subset of strains were serotyped, and tested for other virulence factors and markers of pandemicity. Twenty six different serotypes were represented and five of the serotypes were clinically relevant. Although all 170 isolates were negative for *tdh*, *trh*, and the Kanagawa Phenomenon, 18 possessed the GS-PCR sequence, 2 the ORF8 sequence, and 34 the 850 bp sequence of the *V. parahaemolyticus* pandemic strains. The *V. parahaemolyticus* population in the Black Sea was determined to be genomically heterogenous by rep-PCR and serodiversity did not correlate with rep-PCR genomic diversity. Statistical modeling was used successfully to predict presence of *V. parahaemolyticus* as a function of water temperature, with strongest concordance observed for Green Cape site samples (concordance 70%, $P < 0.001$). Results demonstrate a diverse population of *V. parahaemolyticus* in the Black Sea carrying

pandemic markers. Based on predictive modeling, increased water temperature is related to the number of *V. parahaemolyticus* in the water.

## ***Introduction***

*Vibrio parahaemolyticus*, a halophilic bacterium, is a causative agent of seafood related gastroenteritis, wound infections, and septicemia and is known to occur in marine, estuarine, and brackish water environments globally with sporadic occurrence reported in fresh water (Sarkar et al., 1985; DePaola et al., 2000; Wong et al., 2000; Alam et al., 2009). Infections typically occur after consumption of raw or undercooked seafood and wounds are most frequently acquired by handling seafood or exposure of open wounds to natural waters. Sporadic cases and outbreaks caused by *V. parahaemolyticus* increasingly have been reported in the United States, Asia, and Europe (Chowdury et al., 2000; Daniels et al., 2000; Ansaruzzaman et al., 2005). In addition to its notoriety as a causative agent of human infection, the organism is also an autochthonous member of the marine and brackish water microbial communities and, like other *Vibrio* spp., plays an important role in the carbon cycle by degrading chitin (Kaneko and Colwell, 1974; Kadokura et al., 2007). One of its main virulence factors, the type three secretion system-2 (TTSS2), plays an important role in preventing predation of its host by higher organisms, suggesting virulence factors evolved in the environment (Matz et al., 2011). Little work has been done on non-anthropocentric roles of this organism, but its ubiquity and association with animals demonstrate its ecology extends beyond the human body.

Several virulence factors are associated with *V. parahaemolyticus* pathogenicity. The majority of clinical strains encode the thermostable direct hemolysin (TDH) within

the *Vibrio parahaemolyticus* pathogenicity island (Vp-PAI) (20), one of the virulence factors of *V. parahaemolyticus* responsible for enterotoxicity (Honda, 1993; Guang-Qing et al., 1995). However, some clinical isolates do not encode TDH, but other hemolysins instead, such as the TDH-related hemolysin (TRH) and all strains encode the thermolabile hemolysin (TLH). It has also been reported that two of type three secretion systems (TTSS1 and TTSS2) are involved in pathogenicity of *V. parahaemolyticus* (Bhattacharjee et al., 2006; Ono et al., 2006; Kodama et al., 2007; Matlawska-Wasowska et al., 2011). The TTSS1 found in all *V. parahaemolyticus* strains examined to date has been shown to translocate an effecter protein (VP1686) into the cytosol of macrophages and induce DNA fragmentation and VP1680 has been shown to play a role in cytotoxicity in eukaryotic cells (Bhattacharjee et al., 2006; Ono et al., 2006). This secretion system is similar in genetic sequence and structure to that of *Yersinia pestis* and has been found to be functional in *V. parahaemolyticus* (Makino et al., 2003). Interestingly, *V. parahaemolyticus* strains lacking TDH, TRH, and TTSS2 have frequently been isolated from patients who are not colonized by TDH-, TRH-, and TTSS2-positive strains, suggesting either the TTSS1 or some other not yet identified virulence factor(s), or both, are responsible for illness in humans (Suthienkul et al., 1995; Okuda et al., 1997; Vuddhakul et al., 2000; Laohaprertthisan et al., 2003; Cabanillas-Beltran et al., 2006; Bhoopong et al., 2007; Meador et al., 2007; Serichantalergs et al., 2007; García et al., 2009; Harth et al., 2009; Chao et al., 2009; Chao et al., 2010; Shalu et al., 2010). Studies have shown that up to 11.5% of clinical cases are caused by strains that do not encode any of these virulence factors, suggesting these TDH-, TRH-, and TTSS2-negative strains

are responsible for significant morbidity (Suthienkul et al., 1995). These strains are ubiquitous in saline environments and should be considered potential pathogens.

Multiple serogroups of *V. parahaemolyticus* are associated with human infections and 13 O serogroups and 71 K serotypes have been described with 75 O:K combinations identified to date (Nasu et al., 2000) and regional dominance of specific serogroups also reported (Chatterjee and Sen, 1974; Abbott et al., 1989; Nair et al., 2007). Since 1996, there has been an increase in isolation of serogroup O3:K6 from clinical cases worldwide, notably from the majority of clinical cases in Asia, Europe, Africa, and Latin America. Pandemic strains are characterized by molecular features such as ORF8 marker of the f237 lysogenic phage (Nasu et al., 2000), a seven base pair polymorphism in the *toxRS* sequence detected by group-specific PCR (GS-PCR) (Matsumoto et al., 2000), histone-like DNA binding protein (HU-α ORF), and a unique arbitrarily primed PCR pattern (AP-PCR) (Okuda et al., 1997; Matsumoto et al., 2000). Recently, serovariants of pandemic strains positive for the GS-PCR sequence but lacking ORF8 marker have been isolated from clinical cases, suggesting gene transfer between strains occurs on a global scale (Okura et al., 2003).

*V. parahaemolyticus* frequently has been isolated from water samples collected from the Black Sea and sporadic cases of gastroenteritis caused by this bacterium and related vibrios have been reported in the Sea of Azov region (Libinzon et al., 1974; Libinzon et al., 1980; Libinzon et al., 1981; Shikulov et al., 1980; Clark et al., 1998; WHO, 2011). Based on the increasing incidence of infection in Europe, Baker-Austin et al. (2010) have called for monitoring both environmental and seafood-borne *V. parahaemolyticus*. The objective of this study reported was to model occurrence of *V.*

*parahaemolyticus* on the coast of the Black, Georgia (Former Soviet Union), related to environmental parameters and to characterize phenotypic and molecular diversity of this population of bacteria.

## ***Materials and Methods***

### Sample Collection and Processing

Routine surveillance of the aquatic environment of Georgia, a former member of the Soviet Union (FSU) and the Commonwealth of Independent States (CIS), was conducted from June, 2006, to October, 2008. Water samples were collected monthly, except July to September when water was collected biweekly, from five stations on the coast of the Black Sea (Figure 8.1). One hundred liters of water were filtered through 200- and 64-µm plankton nets, to separate size fractions of plankton. Simultaneously, one-liter of the plankton filtrate, the water fraction, was collected in sterile bottles and all samples were transported to the laboratory and processed within 6 hrs of collection. Water temperature, salinity, pH, and dissolved oxygen were recorded at the time of sampling. The water fraction (100 ml) was filtered using a 0.45 µm nitrocellulose membrane, which was incubated in alkaline peptone water (APW) at 37ºC for 24 hrs. An aliquot (1 to 5- ml) of each plankton fraction (64- and 200- µm) was also inoculated and incubated in APW at 37°C for 24 hrs. A 10 microliter loop of the enrichment cultures were streaked onto thiosulfate citrate bile salts (TCBS) agar plates which were incubated overnight at 37°C. All colonies that appeared yellow to green at 24 hrs were considered presumptive *Vibrio* spp., picked with a sterile toothpick, and streaked to isolate colonies on Luria-Bertani (LB) agar. Presumptive *V. parahaemolyticus* colonies were confirmed

by streaking onto CHROMagar™ *Vibrio* (mauve colonies) the latter were confirmed by PCR (Table 8.1).

**Virulence Factors**

Virulence factors and pandemicity were detected employing PCR for *tdh*, *trh*, *tlh*, Mtase, ORF8 sequence of the f237 phage, pandemic group-specific *toxRS* sequence (GS-PCR), histone-like DNA binding protein (HU-α ORF), 850-bp fragment region of O3:K6 isolates, VP1346 (*yop*) and VP1339 (*escC*) of TTSS2 and VP1680 and VP1686 of TTSS1 (Table 8.1). DNA (25.0 ng) was mixed with 2.5 mM of dNTP, 15 mM of PCR buffer, and 5 U $\mu L^{-1}$ of Taq DNA polymerase, using 20 μm of appropriate primer for each analysis. Amplicons were visualized on 1.5% agarose gel stained with ethidium bromide and examined under a UV transilluminator.

**Estimation of Molecular Diversity by REP-PCR**

To determine the molecular diversity of the *V. parahaemolyticus* isolates rep-PCR was executed on a randomly selected subset of strains following the methods of Chokesajjawatee et al. (2007). PCR products were separated on a 1% agarose gel in TAE buffer. The resulting fingerprint patterns were documented using the GelDoc-It™ Imaging System (Ultra-Violet Products, Upland, CA).

**Phenotypic Analyses**

To determine the hemolytic activity on Wagatsuma agar, cells were first streaked for isolation on nutrient agar supplemented with 7% NaCl at 37ºC for 18 to 24 hours. A

single colony was removed and incubated at 37ºC and 100 rpm for 18 to 24 hours in nutrient broth supplemented with 7% NaCl. Ten microliters of this inoculum were then thinly streaked across Wagatsuma agar in a straight line and incubated at 37ºC for 18 to 24 hours. Cells causing β-hemolysis of red blood cells were determined to be Kanagawa Phenomenon (KP) positive. Urease activity was evaluated by inoculating isolated colonies into urea agar base supplemented with urea (Oxoid, Hampshire, England). Antibiotic susceptibility assays were conducted by growing isolates overnight in Heart-Infusion broth, swabbing each culture uniformly across plates with Mueller-Hinton agar (Becton Dickinson and Company, Franklin Lakes, NJ, USA), allowing the inoculated plates to dry and then placing antibiotic disks on top of the agar. Plates were incubated overnight and the diameters of inhibitions were measured and interpreted as per the manufacturer's specification. All media were supplemented with 2% NaCl.

**Serotype Determination**

Strains were streaked on LB agar with 3% NaCl and incubated overnight at 37ºC. One 10 µl loopful of growth was homogenized in 1 mL of saline solution (0.9% NaCl). This solution was divided into two 500 µl tubes, one of which was boiled for 2 hours. Ten microliters of the boiled cell solution was then mixed with 10 µl of each O-antisera and 10 µl of the cell suspension that had not been boiled was mixed with 10 µl of K-antisera on a glass slide and agglutination visually determined (Denka Seiken Co., Niigata-ken, Japan).

### Statistical Analyses

Differences between frequency of recovered serotype and the frequency of isolation between sampling sites were evaluated using a $\chi^2$ test. Predictive models of *V. parahaemolyticus* detection determined by examining the relationship between presence/absence (dependent variable) and recorded environmental parameters (independent variables) at the time of sample collection. These data were transformed to determine distance from optimality by determining median values of all parameters for those samples in which *V. parahaemolyticus* had been detected (optimal parameters) and subtracting this value from all observations. The absolute values of differences were used as independent variable in binary logistic regression analysis. For all measures of association, p values $\leq 0.05$ were considered significant. Statistical analyses were conducted on R and SAS software (Cary, NC, USA).

## ***Results***

### Detection of *V. parahaemolyticus*

In total, 170 isolates of *V. parahaemolyticus* were recovered from Black Sea water and plankton samples collected at stations located along the Georgian coast. All yellow, green or olive colonies on TCBS agar and magenta on CHROMagar™ *Vibrio* that were oxidase positive were purified for biochemical analyses and identification by PCR. A total of 101 *V. parahaemolyticus* isolates were recovered from water and from plankton. From plankton samples, 30 isolates were from the 64 µm fraction and 39 from the 200 µm fraction (Figure 8.2).

*V. parahaemolyticus* was isolated from 40 water samples and 19 and 26 of 64- and 200- μm plankton fractions. Based on Cochran's Q test, water samples yielded *V. parahaemolyticus* significantly more frequently than either of the plankton fractions. The difference in *V. parahaemolyticus* isolation frequency was not significantly different between plankton fractions. When these distributions were binned to water temperature quartiles (11, 19.8, and 25.8°C), water samples collected between 11 and 19.8°C were significantly more likely to yield *V. parahaemolyticus* than plankton. Median temperatures for *V. parahaemolyticus*-positive water samples and *V. parahaemolyticus*-negative plankton, both water and plankton *V. parahaemolyticus*-positive, and water *V. parahaemolyticus*-negative and plankton *V. parahaemolyticus*-positive were 20.3, 25.8, and 26.7°C, respectively (Table 8.2). Isolates were more frequently recovered in warmer summer months than cooler winter months, at which time the water temperature was significantly higher ($P < 0.001$).

**Serodiversity**

Twenty-seven serotypes of *V. parahaemolyticus* were detected: the majority of these being O2:K28 (8 isolates), O3:K31 (7), O3:KUT (7), O4:KUT, and untypeable (24) (Table 8.2), The untypeable strains were most frequently isolated (Friedman's $\chi^2$, $p<0.05$). *V. parahaemolyticus* O3 O-antigenic type was the most common, comprising 35% of the isolates. .

*V. parahaemolyticus* of the O3 O-antigenic type were recovered only from samples collected in the Supsa Estuary, Batumi Boulvard, Green Cape, and Choroki Estuary sites. The serotype O2:K28 isolates were more common in the Supsa Estuary

and O3:K31 isolates in the samples collected at Batumi Boulvard, O4:KUT from the Supsa Estuary, and untypeable strains from Green Cape. However, it is concluded that the *V. parahaemolyticus* serotypes were distributed randomly between sites based on results of Friedman's $\chi^2$ analysis.

**Virulence Factors and Markers of Pandemic Clones**

None of the strains of *V. parahaemolyticus* carried the gene for thermostable direct hemolysin (*tdh*), hence all were Kanagawa Phenomenon –negative and also were urease-negative (Table 8.2). None of the isolates possessed thermostable-related hemolysin (*trh*), TTSS-2, or MTase. Seven carried the ORF8 marker of the f237 filamentous phage but only four also carried the 850-bp pandemic sequence and no other pandemic markers. Two of the ORF8-positive strains were serotyped both were untypeable (UT). Nineteen strains carried the pandemic GS-PCR marker (*toxRS* sequence of pandemic strains), but only seven were 651 bp and 12 were ca. 750 bp. Three of the 651 bp, GS-PCR-positive strains were also positive for the 850 bp pandemic sequence, whereas six of the 750 bp, GS-PCR-positive isolates also encoded this region. Each of the 650 bp, GS-PCR-positive isolates were different serotypes and typed as O1:KUT, O3:KUT, O3:K31, O3:K33 O3:K65, and UT, the most notable was the O1:KUT, a type to exhibit pandemicity. This isolate was also positive for the 850 bp pandemic sequence but lacked all other markers of virulence except TTSS1.

### REP-PCR

Long range rep-PCR was performed on 45 of the strains (Figure 8.3). A dendrogram of banding patterns revealed 45 distinctive bands.

### Predictive Modeling

Using binary logistic regression a model was constructed which best fit the observed data for occurrence of *V. parahaemolyticus* water temperature since it had been found to be the single significant predictor (Table 8.3). For all sites combined, 37.3% of the variance in isolation of *V. parahaemolyticus* was explained by water temperature. In the Chorokhi and Supsa estuaries, only 22 and 32.1%, respectively, of the variance in *V. parahaemolyticus* isolation was explained by water temperature, a relationship higher for Batumi Bulvard and Green Cape sites (43.2% and 70.1%, respectively) (Table 8.3).

## ***Discussion***

*Vibrio parahaemolyticus* is a bacterium halophilic, autochthonous to marine and estuarine environments, and found in coastal environments worldwide. Although commonly isolated from brackish waters, its presence indicates public health concern since inhabitants of regions harboring the pathogen experience sporadic cases of illness, related to seafood with which the bacterium is associated. This risk is high regardless of whether pathogenicity islands are present since infections are caused by isolates lacking *tdh*, *trh*, and TTSS2. Sporadic outbreaks caused by *V. parahaemolyticus* historically have been reported around the Black Sea, indicating the value of monitoring for *V. parahaemolyticus* (Libinzon et al., 1974; Zakhariev, 1975).

The strains which were isolated represented 9 O-antigens and 27 k-antigens, as well as untypeable strains, a measure of the antigenic diversity of the natural isolates in this region. Gram-negative bacteria are capable of frequent mutations within antigen coding regions of the genome as well as lateral transfer possibly allowing strains to adapt to microenvironments of the environment (Lerouge et al., 2001; Woo et al., 2001). Molecular divergence was indicated by heterogeneity among O3:K31 and O2:K28 strains in rep-PCR analysis. This genomic indicates a need to classify strains by methods other than serology, which represents ca. 1% of the genome. The high degree of divergence among environmental *V. parahaemolyticus* strains in the Black Sea is corroborated by similar findings in geographically distant regions (Wong et al., 1999(a); Wong et al., 1999(b); Wong et al., 1999(c): Matsumoto et al., 2000; Alam et al., 2009).

*V. parahaemolyticus* seasonality was found to be a predictable pattern at all sites, with a clear trend of increasing numbers as water temperatures increased, May to September. *V. parahaemolyticus* was isolated from water samples, the temperatures of which was 8°C, a low temperature, but >93% of strains were isolated in waters > 17°C (Table 8.2). The statistical model for detection of *V. parahaemolyticus* pointed to water temperature at all sites with concordance highest at Green Cape (concordance = 70%, $P < 0.05$).

In conclusion, an antigenically diverse population of *V. parahaemolyticus* inhabits the Georgian coast of the Black Sea. Although none of the strains collected during this study Kanagaw phenomena-positive or *tdh* and *trh*-positive, the presence of TTSS1 effector proteins and thermolabile hemolysin were encoded, including a serovariant of the *V. parahaemolyticus* O3:K6 pandemic clone. These results, together with

epidemiological data showing strains lacking pathogenicity islands can cause disease suggest a risk can be associated with the occurrence of *V. parahaemolyticus* in the Black Sea coastal waters of Georgia. The warmer temperatures in the spring and summer lead to increased populations of *V. parahaemolyticus*.

| Target gene | Sequence (5'-3') | Amplicon size (bp) | Reference |
|---|---|---|---|
| pR72H element | F-TGCGAATTCGATAGGGTGTTAACC<br>R-CGAATCCTTGAACATACGCAGC | 387 | Lee et al (1995) |
| collagenase | F-GAAAGTTGAACATCATCAGCACGA<br>R-GGTCAGAATCAAACGCCG | 271 | Di Pinto et al (2006) |
| *tdh* | F-GTAAAGGTCTCTGACTTTTGGAC<br>R-TGGAATAGAACCTTCATCTTCACC | 269 | Bej et.al. (1999) |
| *trh* | F-TTGGCTTCGATATTTTCAGTATCT<br>R-CATAACAAACATATGCCCATTTCCG | 500 | Bej et.al. (1999) |
| *tlh* | F-AAAGCGGATTATGCAGAAGCACTG<br>R-GCTACTTTCTAGCATTTTCTCTGC | 450 | Bej et.al. (1999) |
| GS-PCR | F-TAATGAGGTAGAAACA<br>R-ACGTAACGGGCCTACA | 651 | Matsumoto et.al. (2006) |
| Orf8 | F-GTTCGCATACAGTTGAGG<br>R-AAGTACAGCAGGAGTGAG | ~700 | Nasu et.al., (2000) |

Table 8.1. PCR primers used in this study

| Serotype | # of Isolates | W/P[a] (64/200)[b] | *tdh* | KP | *tlh* | *trh* | VP1680 | VP1686 | ORF8 | GS-PCR |
|---|---|---|---|---|---|---|---|---|---|---|
| O1:K32 | 1 | 1/0 | 0 | 0 | 1[c] (100)[d] | 0 | 1 (100) | 1 (100) | 0 | 0 |
| O1:K58 | 1 | 0/1 (0/1) | 0 | 0 | 1 (100) | 0 | 1 (100) | 1 (100) | 0 | 0 |
| O1:KUT | 1 | 0/1 (0/1) | 0 | 0 | 1 (100) | 0 | 1 (100) | 1 (100) | 0 | 1 (100) |
| O2:K28 | 8 | 6/2 (0/2) | 0 | 0 | 7 (87.5) | 0 | 7 (87.5) | 7 (87.5) | 0 | 0 |
| O2:KUT | 1 | 0/1 (0/1) | 0 | 0 | 1 (100) | 0 | 1 (100) | 1 (100) | 0 | 0 |
| O3:K5 | 1 | 1/0 | 0 | 0 | 1 (100) | 0 | 1 (100) | 1 (100) | 0 | 0 |
| O3:K31 | 7 | 3/4 (3/1) | 0 | 0 | 7 (100) | 0 | 7 (100) | 7 (100) | 0 | 1 (1) |
| O3:K33 | 1 | 1/0 | 0 | 0 | 1 (100) | 0 | 1 (100) | 1 (100) | 0 | 1 (100) |
| O3:K51 | 2 | 0/2 (1/1) | 0 | 0 | 2 (100) | 0 | 2 (100) | 2 (100) | 0 | 0 |
| O3:K65 | 2 | 1/1 (1/0) | 0 | 0 | 2 (100) | 0 | 2 (100) | 2 (100) | 0 | 1 (50) |
| O3:KUT | 7 | 2/5 (3/2) | 0 | 0 | 7 (100) | 0 | 7 (100) | 7 (100) | 0 | 1 (14) |
| O4:K12 | 1 | 1/0 | 0 | 0 | 1 (100) | 0 | 1 (100) | 1 (100) | 0 | 0 |
| O4:K34 | 1 | 1/0 | 0 | 0 | 1 (100) | 0 | 1 (100) | 1 (100) | 0 | 0 |
| O4:K37 | 1 | 0/1 (0/1) | 0 | 0 | 1 (100) | 0 | 1 (100) | 1 (100) | 0 | 0 |
| O4:KUT | 7 | 2/5 (2/3) | 0 | 0 | 7 (100) | 0 | 7 (100) | 7 (100) | 0 | 0 |
| O5:K68 | 2 | 1/1 (0/1) | 0 | 0 | 2 (100) | 0 | 2 (100) | 2 (100) | 0 | 0 |
| O5:KUT | 2 | 2/0 | 0 | 0 | 2 (100) | 0 | 2 (100) | 2 (100) | 0 | 0 |
| O6:KUT | 1 | 1/0 | 0 | 0 | 1 (100) | 0 | 1 (100) | 1 (100) | 0 | 0 |
| O8:KUT | 2 | 2/0 | 0 | 0 | 2 (100) | 0 | 2 (100) | 2 (100) | 0 | 0 |
| O10:K61 | 1 | 1/0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| O10:K60 | 1 | 1/0 | 0 | 0 | 1 (100) | 0 | 1 (100) | 1 (100) | 0 | 0 |
| O10:KUT | 1 | 0/1 (1/0) | 0 | 0 | 1 (100) | 0 | 1 (100) | 1 (100) | 0 | 0 |
| O11:KUT | 1 | 1/0 | 0 | 0 | 1 (100) | 0 | 1 (100) | 1 (100) | 0 | 0 |
| OUT:K27 | 1 | 0/1 (0/1) | 0 | 0 | 1 (100) | 0 | 1 (100) | 1 (100) | 0 | 0 |
| OUT:K33 | 2 | 2/0 | 0 | 0 | 2 (100) | 0 | 2 (100) | 2 (100) | 0 | 1 (50) |
| OUT:K52 | 1 | 1/0 | 0 | 0 | 1 (100) | 0 | 1 (100) | 1 (100) | 0 | 0 |
| UT | 24 | 19/5 (3/2) | 0 | 0 | 23 (95.8) | 0 | 23 (95.8) | 24 (100) | 0 | 1 (4) |

a W = water, P = plankton
b plankton fraction size in μm
c number of positive isolates
d percent of total isolates of that serotype

Table 8.2. Molecular characterics of serotyped strains.

| Environmental Conditions when *V. parahaemolyticus* was Detected / Not Detected | | | | | |
|---|---|---|---|---|---|
| Media | Statistic | Salinity (‰) | Water Temp (ºC) | pH | DO (mg/L) |
| Water | Min | 3.4 A / 3.6 B | 8 / 7.7 | 6.2 / 6.3 | 2.1 / 2 |
| | Max | 20.8 / 20.8 | 28.5 / 29.7 | 8.6 / 8.5 | 7.2 / 7 |
| | Mean | 12.9 / 12 | 22.8 / 16.5 | 7.7 / 7.8 | 4.4 / 4.3 |
| | Median | 15.7 / 13 | 24.25 / 13 | 7.8 / 7.9 | 4.3 / 4.6 |
| | Std Dev | 5.0 / 7.3 | 4.9 / 7.3 | 0.7 / 0.6 | 1.2 / 1.3 |
| P64 | Min | 5 / 3.6 | 19.3 / 7.7 | 6.5 / 6.2 | 2 / 2 |
| | Max | 17.4 / 20.8 | 28.5 / 29.7 | 8.4 / 8.5 | 6.8 / 7.2 |
| | Mean | 13.6 / 12.3 | 25.4 / 17.3 | 7.9 / 7.8 | 4.1 / 4.4 |
| | Median | 16.5 / 14.2 | 26.6 / 17 | 8.2 / 7.9 | 4.2 / 4.5 |
| | Std Dev | 4.6 / 4.9 | 2.9 / 7 | 0.6 / 0.6 | 1 / 1.2 |
| P200 | Min | 3.4 / 3.6 | 18 / 7.7 | 6.2 / 6.5 | 2.1 / 2 |
| | Max | 20.8 / 20.8 | 29 / 29.7 | 8.4 / 8.5 | 7.2 / 7.2 |
| | Mean | 12.8 / 12.3 | 24.6 / 17 | 7.6 / 7.8 | 4.4 / 4.3 |
| | Median | 14.9 / 14 | 25.6 / 14.2 | 7.6 / 8 | 4.1 / 4.4 |
| | Std Dev | 5.4 / 4.7 | 3 / 7 | 0.7 / 0.5 | 1.3 / 1.2 |
| All Plankton | Min | 3.4 / 3.6 | 18 / 7.7 | 6.2 / 6.5 | 2 / 2 |
| | Max | 20.8 / 20.8 | 29 / 29.7 | 8.4 / 8.5 | 7.2 / 7.2 |
| | Mean | 13.2 / 12.3 | 25 / 16.2 | 7.7 / 7.8 | 4.2 / 4.3 |
| | Median | 16 / 14.1 | 25.8 / 13.6 | 7.8 / 8 | 4.1 / 4.5 |
| | Std Dev | 5 / 4.8 | 3 / 6.8 | 0.7 / 0.5 | 1.2 / 1.2 |
| All Sample Types | Min | 3.4 / 3.6 | 8 / 7.7 | 6.2 / 6.5 | 2 / 2 |
| | Max | 20.8 / 20.8 | 29 / 29.7 | 8.5 / 8.5 | 7.2 / 7 |
| | Mean | 12.7 / 12 | 22.7 / 15.3 | 7.7 / 7.8 | 4.4 / 4.3 |
| | Median | 15.2 / 13.2 | 24 / 12.4 | 7.7 / 8 | 4.4 / 4.6 |
| | Std Dev | 5 / 4.9 | 4.9 / 6.9 | 0.7 / 0.5 | 1.2 / 1.2 |

A = statistic when V. parahaemolyticus was detected
B = statistic when V. parahaemolyticus was not detected

Table 8.2. Recorded environmental parameters when *V. parahaemolyticus* was/was not detected for each sample type.

| Statistics | All Sites | Chorokhi | Batumi Bulvard | Green Cape | Supsa |
|---|---|---|---|---|---|
| Parameter | Temperature | | | | |
| Coefficient | 0.27 | 0.19 | 0.32 | 0.53 | 0.24 |
| Concordance | 37.33 | 22.01 | 43.18 | 70.08 | 31.23 |
| P-value | < 0.0001 | < 0.001 | < 0.01 | < 0.01 | < 0.01 |
| Deviance | 112.93 | 35.47 | 25.98 | 12.32 | 30.16 |

Table 8.3. Results of binary logistic regression analysis between *V. parahaemolyticus* and water temperature.

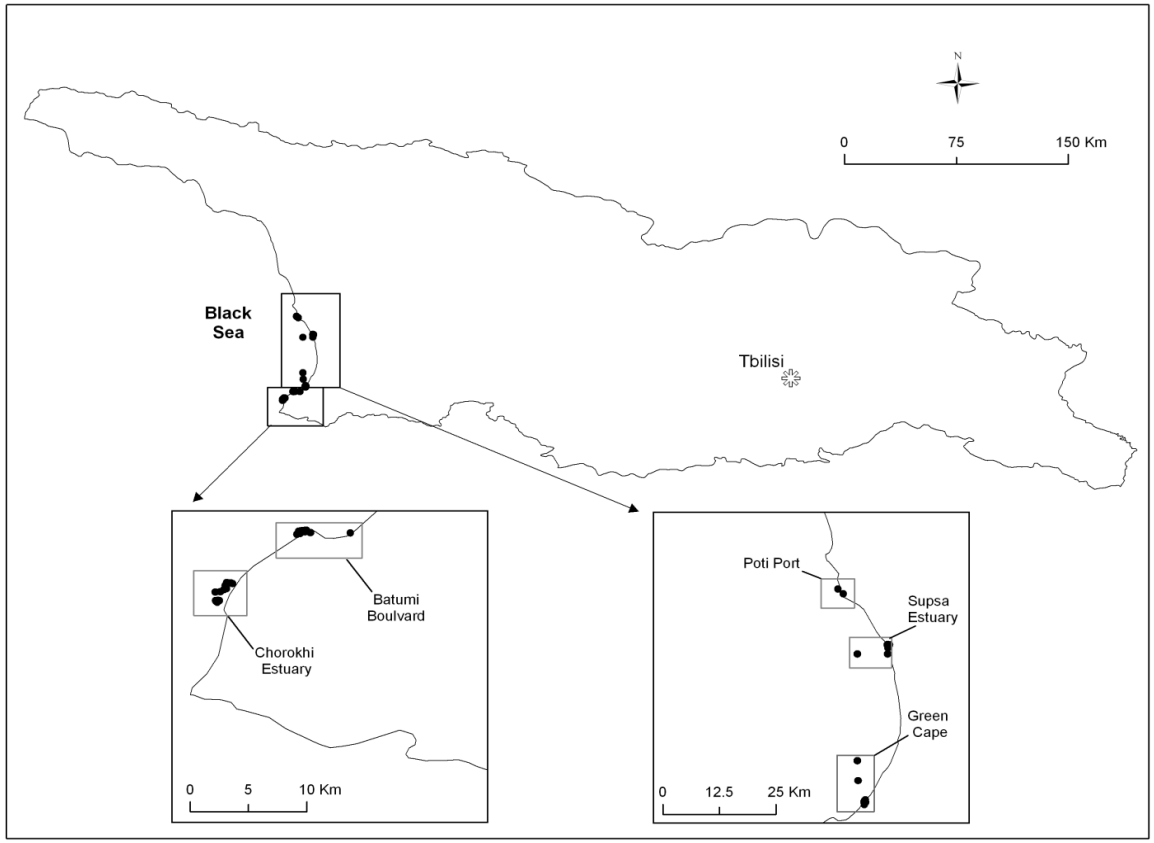



Figure 8.1. Map showing locations of sampling sites along Black Sea and freshwater lakes near Tbilisi.

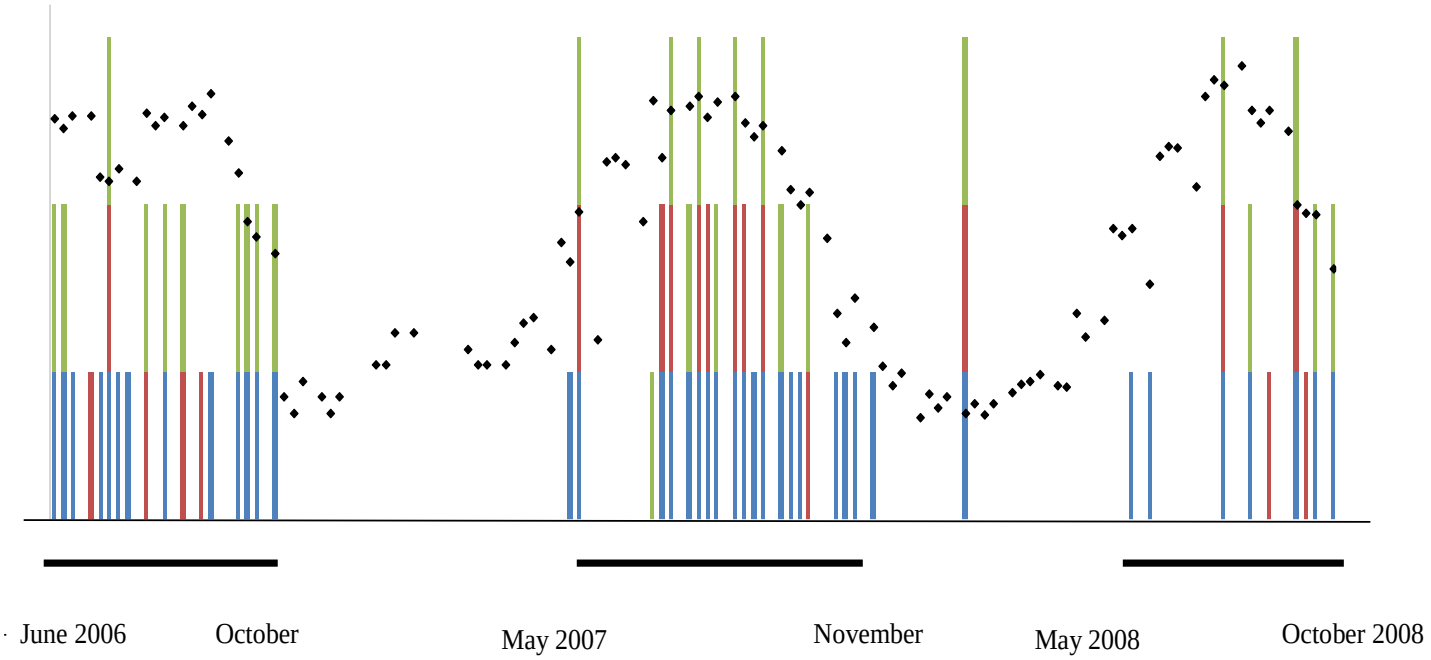
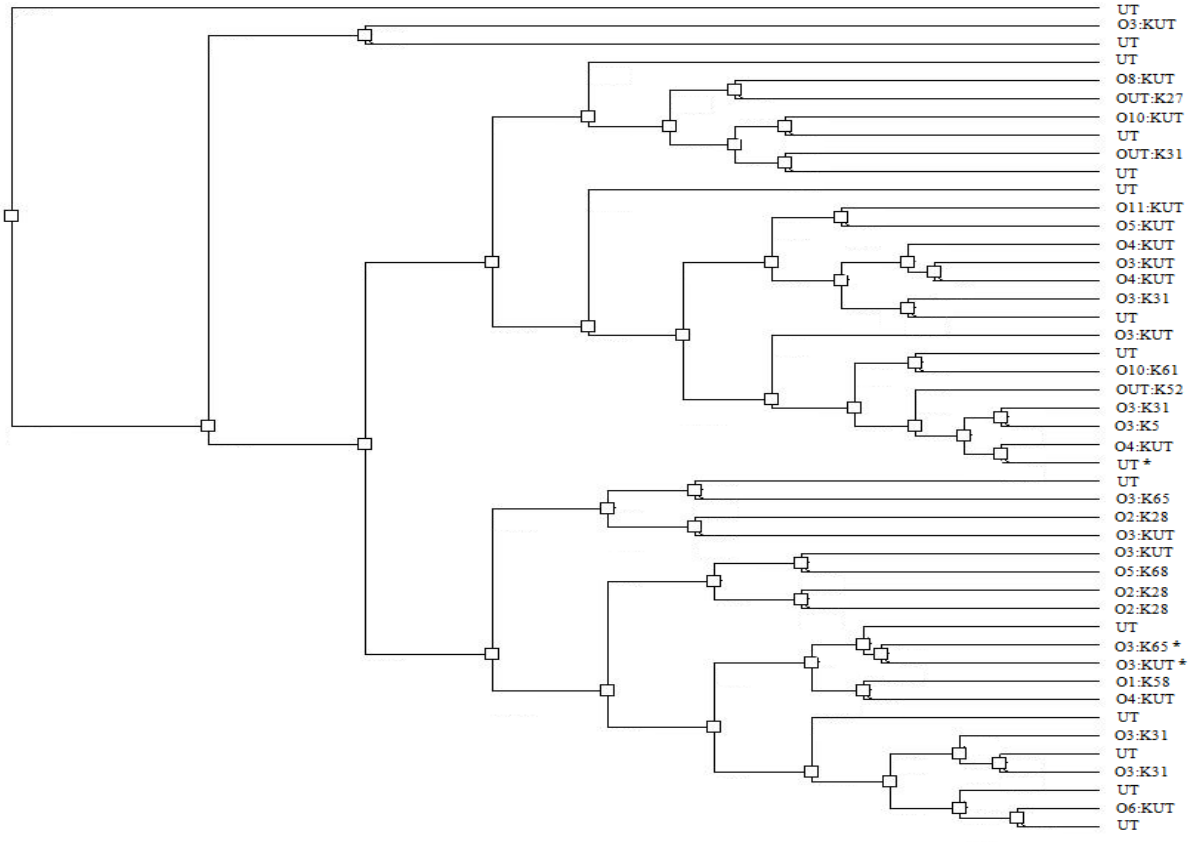



Figure 8.2. Water temperature (black diamonds) and *V. parahaemolyticus* detection in water (blue bars) and plankton (200 µm [green bars] and 64 µm [magenta bars]).

Figure 8.3. rep-PCR dendrogram of a subset of randomly selected strains in this study. Serotypes of the analyzed strains are shown at tree leaves. Asterisks indicate strains that were positive for GS-PCR



# Chapter 9: Detection of *Vibrio cholerae* in Environmental Waters Including Drinking Water Reservoirs of Azerbaijan


***Abstract***

Cholera, a globally prevalent gastrointestinal disease, remains a persistent problem in many countries including the former Soviet republics of the Caucasus region where sporadic outbreaks occurred recently. Historically, this region has experienced cholera during every pandemic since 1816, however, no known comprehensive evaluation of the presence of *Vibrio cholerae* in surface waters using molecular methods has been done. Here we present the first report of the presence of *V. cholerae* in surface waters of Azerbaijan and its seasonality, using a combination of bacteriological and molecular methods. Findings from the present study indicate a peak in the incidence of *V. cholerae* in warmer summer months relative to colder winter months. In the Caspian Sea, water temperature when optimal for growth of *V. cholerae* was significantly associated with detection of *V. cholerae*. *V. cholerae* was simultaneously detected at freshwater sites including two water reservoirs. Most importantly, detection of *V. cholerae* in these water reservoirs, source of municipal drinking water poses a potential health risk to the population due to the limited and insufficient treatment of water in Azerbaijan. Routine monitoring of environmental waters used for recreational purposes, and especially drinking water reservoirs is highly recommended as measure for public health safety.

## ***Introduction***

*Vibrio cholerae*, the causative agent of cholera, has been isolated from marine, estuarine, and fresh surface waters throughout the world. Historically, epidemic cholera was caused only by toxigenic strains of the O1 serogroup. In 1992, a new serogroup, O139, emerged and caused more than 10,000 deaths in India and Bangladesh (ICDDRB, 1993; Ramamurthy et al., 1993). *V. cholerae* non-toxigenic non-O1/non-O139 strains may also cause small outbreaks and sporadic cases of cholera, as well as extra-intestinal infections (Safrin et al., 1988; Ko et al., 1998; Lukinmaa et al., 2006; Shannon and Kimbrough, 2006; Chatterjee et al., 2009). These mild cases often go undiagnosed. Furthermore, inadvertently missed and inadequate disease reporting systems in even highly developed countries and failure of many sick individuals to seek medical attention for episodes of diarrhea result in considerable underreporting of intestinal infections on a global scale (Brabazon et al., 2008; MacDougall et al., 2008). Thus, it is reasonable to conclude that the global burden of illness caused by *V. cholerae*, along with other infectious agents, is significantly underestimated, therefore, potentially more cases occur than are revealed in the reports.

In Azerbaijan, a former Soviet Union (FSU) republic and currently a member of the Commonwealth of Independent States (CIS) in the Caucasus region, cholera has been reported sporadically with large outbreaks occurring at times (Gurbanov et al., 2012). This country reported the 4$^{th}$ most cases of cholera of all the Soviet republics during Soviet era of the ongoing 7$^{th}$ cholera pandemic (1961 to 1991) (Narkevich et al., 1993). Historically, cholera outbreaks in this country have been repeatedly linked to drinking water reservoirs that serve the majority of the urban populations (1985, 1993, and 1995)

(Gurbanov et al., 2012). The Caucasus region as a whole experiences cholera and other *Vibrio* infections, but available data and literature on these diseases and their causative agents in this region remain scarce (Narkevich et al., 1993). Overall, water resources in Azerbaijan are poor (Gurbanov et al., 2012) (Ewing, 2010) and information on water quality and quantity in this country are limited. Although a significant portion of the population in urban areas has access to piped-water, rural inhabitants do not, and therefore, water treatment often times occurs in one's household (Rosa and Clasen, 2010). This poor water quality motivated the United Nations to include Azerbaijan in the list of countries aiming to improve clean water access to the general public as one of its millennium development goals.

The occurrence of cholera and the epidemiologic features of this disease, its ability to spread globally and the periodicity of outbreaks are influenced by environmental factors that enhance the growth of *V. cholerae* in natural waters (Colwell, 1996; Binsztein et al., 2004; Huq et al., 2005; Pruzzo et al., 2008). Ecological studies of *V. cholerae* aid in the discernment of environmental factors that influence its growth and subsequent increased numbers in the environment. From this information, forecasts can be made that allow prediction of cholera on a global scale thus allowing public health officials get prepared in anticipation of an outbreak of cholera and work proactively for prevention. Several studies during the past two decades indicate that cholera epidemics can be predicted (Lobitz et al., 2000; Huq et al., 2005; Colwell, 2006).

To evaluate the presence and seasonality of *V. cholerae* in surface waters of Azerbaijan, water and plankton samples were collected from freshwater rivers and lakes and from the

Caspian Sea and subjected to modern bacteriological investigation, and the findings are reported here.

***Materials and Methods***

Samples were collected each month from nine sites, including a reservoir, lakes, a river, and the Caspian Sea, as shown in Figure 9.3 over a 15 month period, from May, 2010 to July, 2011. Water and air temperature, pH, salinity, and conductivity were measured on site, using a portable water meter (YSI Incorporated, Yellow Springs, Ohio). Surface water samples (100 ml) were collected by filling a sterile 4-liter bucket, the contents of which were subsequently passed through 20-µm and 64-µm plankton nets. The fractions were transferred to sterile bottles at the collection site. One liter of water was passed through a 20-µm net, providing "plankton-free" water sample (FW). After transport to the laboratory, the two plankton fractions were filtered using a 0.22-µm polycarbonate membrane and the filter was transferred into a 100-ml conical flask containing 25 ml alkaline peptone water (APW). One milliliter of FW was subsequently passed through a 0.22-µm polycarbonate membrane, which was also placed in 25 ml of APW. Each plankton fraction (5 ml) was homogenized with a sterile glass tissue grinder, passed through a 0.22-µm polycarbonate membrane, and placed in 25 ml APW. All APW inoculations were incubated for 37°C for 18 to 24 h without shaking. The microaerophilic pellicle that formed at the top of the tube of the APW (10 µl) was streaked onto TCBS, and incubated at 37°C for 18 to 24 h. Five to 15 yellow colonies were selected from each plate as presumptive *V. cholerae*, based on color and colony morphology, streaked onto gelatin agar, and incubated at 37°C for 18 to 24 h. From each

plate, a colony was selected and tested for oxidase and inoculated into Luria-Bertani broth overnight at 37°C. After incubation, 0.5 ml of the culture was transferred to a cryovial amended with sterile glycerol to a final concentration of 25%. Biochemical profiles of each isolate was determined using standard bacteriological methods (Huq et al., 2006). Strains that appeared to be *V. cholerae* were examined further using O1 and O139 anti-sera and confirmed to be *V. cholerae* by PCR by targeting the intergenic spacer unit (Chun et al., 1999). Strains were further tested for the presence of CTXΦ, *tcpA*, *nanH*, and nag-ST by PCR. Distance from optimal water temperatures were determined by binning water temperatures to the nearest 1°C and examination of the frequency distribution of *V. cholerae* detection. The subsequent median value (optimal water temperature) was then subtracted from the observed salinities and the absolute values of these differences were the distances from optimal water temperatures (Jacobs et al., 2010). Logistic regression was employed using the recorded environmental parameters and distances from optimal parameters (R).

## ***Results and Discussion***

Sampling sites were classified as freshwater or saline waters for purposes of this study (Figure 9.1). Water and plankton samples were collected from four fresh water sites; Kura River, Lake Karkhana, Jeyranbatan reservoir, and Mingachevir reservoir, and five sites along the Caspian Sea; Buzovnah, Bikgah, Novkhani, Shikhov, and Khanlar. The Kura River is the largest river in Azerbaijan and its head waters are a group of streams in northeastern Turkey. The river runs through Turkey and Georgia and bisects Azerbaijan where it flows southeast and joins the Araz River, the second largest river of

this country, before emptying into the Caspian Sea. Mingachevir and Jeyranbatan reservoirs are the 1st and 6th largest reservoirs, respectively, in this country in terms of volume and capacity. Mingachevir reservoir is primarily a source of irrigation water and Jeyranbatan provides municipal waters for Baku and other cities on the Absheron peninsula. The Caspian Sea is the largest inland water body by area in the world, is directly east of Azerbaijan where it is also bordered by Russia, Iran, Turkmenistan, and Kazakhstan. It is fed by over 130 rivers including the Volga River which is the largest in Europe. This water body serves as a source of seafood for the region as well as globally, and it overlays significant reservoirs of gas and oil making it a geopolitically sensitive region for its bordering countries.

Physical and chemical parameters were measured, such as, pH, conductivity, salinity, and water and air temperature at all sampling sites, both saline and fresh water, at the time of sample collection (Figure 9.1). Median water temperature, conductivity, and pH were not significantly different between Caspian Sea sites while salinities were significantly (Kruskall-Wallis, $P < 0.05$). Median salinities were significantly lower at Bilgah (10.2 ‰) than at Khanlar (10.7 ‰) and Shikov (10.75 ‰). When the freshwater sites were compared, median water temperatures at Jeyranbatan reservoir were significantly lower (14.4°C) than those at Lake Karkhana sites (28.3°C) (Kruskal-Wallis, $P < 0.05$). Median salinity at Mingachevir reservoir (0.2 ‰) was significantly lower than those at Jeyranbatan reservoir (0.4 ‰) and Lake Karkhana (1.1 ‰) (Kruskal-Wallis, $P < 0.05$). Median salinity at Lake Karkhana (1.1 ‰) was also higher than that at Kura River (0.3 ‰). Median conductivity readings were significantly lower at Lake Karkhana (711 µS) than all other sites and median pH values were significantly higher at Mingachevir

reservoir (8.3) than at Kura River (8.2) and Jeyranbatan reservoir (7.9) (Kruskal-Wallis, $P < 0.05$). At Caspian Sea sites, temperatures in February (10°C) and March (10°C) were significantly lower than June (27°C) and August (28°C) 2010 (Kruskal-Wallis, $P < 0.05$). Other parameters recorded in the Caspian Sea were not significantly different by month (Table 9.1).

A battery of biochemical tests led to identification of presumptive *V. cholerae* (oxidase positive, non-swarming, yellow colonies on thiosulfate citrate bile-salts sucrose [TCBS] agar and revealed a seasonal variation metabolic diversity of presumptive *V. cholerae* isolates (Figure 9.2). Strains utilizing rhamnose and fermenting lactose and inositol were more frequently isolated in winter and spring months of the year, respectively, while arabinose-utilizing strains peaked in occurrence in the spring (Figure 9.2). These data suggest a shift in the availability of nutrients and subsequent shifts in bacterial metabolic diversity in the Caspian Sea.

To confirm the presence of *V. cholerae*, molecular test, polymerase chain reactions (PCR) targeting the *V. cholerae*-specific intergenic spacer unit, was performed on a subset of presumptive *V. cholerae* isolates. Confirmed *V. cholerae* isolates were further tested for virulence factors by PCR and all were negative for the cholera toxin (*ctxA*), *tcpA* of *Vibrio* pathogenicity island-1 (VPI-1), *nanH* of *Vibrio* pathogenicity island-2 (VPI-2), and the non-agglutinating heat-stable enterotoxin (nag-ST). All isolates were also evaluated for the O1 and O139 serogroup *wbe** regions by PCR and all were determined to be *V. cholerae* non-O1/non-O139. *V. cholerae* was confirmed to be most frequently isolated during July 2010, with an occurrence of 80% in the plankton and water samples (Figure 9.3). Detection of *V. cholerae* was less frequent after July, but *V.*

*cholerae* could be detected well into the cooler winter months, with an anomalous peak in number in the Jeyranbatan reservoir in November 2010. It should be noted that *V. cholerae* was not readily detected at other sites during this same month or previous months. The frequency of *V. cholerae* detection peaked again during the warmer months of 2011, but not to the same level as observed during the previous year (Figure 9.3). Environmental parameters were compared when *V. cholerae* was detected and not detected and these analyses show that water temperatures were significantly higher at times when the organism was detected compared to water temperatures when it was not detected for the Caspian Sea sites (Kruskal-Wallis, $P < 0.05$). The median temperature at times of detection in this environment was 20.9ºC and 14.2ºC when not detected. These temperature relationships (*V. cholerae*-positive versus *V. cholerae*-negative water temperatures) were not significantly different for the fresh water sites. There were no significant differences in other environmental parameters when these data were compared for both the Caspian Sea and fresh water sites.

Of the three types of samples collected (20-µm plankton fraction, 64-µm plankton fraction, and water), *V. cholerae* was isolated from the water fraction less frequently than the plankton samples (14, 13, 11, 7, and 10 isolates of *V. cholerae* in the 20-µm plankton fraction [PL20], 64-µm plankton fraction [PL64], homogenized 20-µm plankton fraction [HOM20], homogenized 64-µm plankton fraction [HOM64], and the < 20-µm plankton fraction [FW], respectively). However, these data were not significant at the $P < 0.05$ level (Cochran's Q) but they do reaffirm the role of plankton as a reservoir of *V. cholerae* in the aquatic environment (Huq et al., 1983; Huq et al., 1990; Huq et al., 1995; Turner et al., 2009; Martinelli-Filho et al., 2011). Homogenized plankton fractions yielded slightly

lower numbers of *V. cholerae* by culture. Because homogenized plankton fractions allow enumeration of both the exterior and interior portions of plankton, enhanced detection of *V. cholerae* relative to non-homogenized plankton fractions would be expected. It has been observed previously that bacteria attached to the surface of zooplankton (copepods) will enter into a viable but nonculturable (VBNC) state whereby conventional bacteriological culture media do not support growth and can be detected only by direct detection methods (Huq et al., 2000). This phenomenon reduces the probability that *V. cholerae* can be isolated on TCBS agar, rather than reflecting an actual decrease in the number of *V. cholerae* present in a sample.

The ultimate goal of ecological studies of *V. cholerae* is to anticipate the occurrence of this organism in natural water in an attempt to prevent illness. Such predictions require long-term analyses, typically 3 years or longer, because of seasonal variation in year to year. However, short-term data can be useful in estimating the role of the environment on *V. cholerae* detection for the study period covering that specific data set, which was done in this study. Based on the principle of parsimony, we chose the simplest models to explain the observed data by using binary logistic regressions. For the Caspian Sea, models were built that used the data from all sites in different compositions (Table 9.2). This approach of combining sites was necessary due to the size of the data set and the variations of sites for model development were explored because *V. cholerae* isolation had occurred at anomalously low temperatures at Bilgah, Buzovnah, and Novkhani suggesting that *V. cholerae* dynamics are different between sites in this environment. The Khanlar and Shikhov-only (K,S) model was developed as these sites have similar salinities. We therefore, built alternative models using data, more typical for

*V. cholerae* isolation, based on global isolation patterns, and we present all 4 sets of models in this report. Similar to other studies, inclusion of fresh water and saline water data in the same model resulted in a poor model, so the two sets of data were separated for separate analyses (Louis et al., 2003). For the "all Caspian model" (all Caspian Sea sites included) several models that significantly explained the variations in the *V. cholerae* presence were fitted to the data, and water temperatures or distances from the optimal water temperature were at the core of all models excluding the two weakest, yet significant, models (not shown) (lowest two Akaike's Information Criteria [AIC] values) (Table 9.2). As water temperature increased, so did the detection of *V. cholerae* and when distance from optimal water temperature increased, detection of *V. cholerae* decreased. For the alternative Caspian Sea models the model fitted the data better, but the variation in *V. cholerae* detection explained by the model remained relatively low. In these two models temperature or distance from optimal temperature remained at the core of each model. At the Khanlar-Shikhov-Bilgah sites (K, S, Bi), when distance from optimal conductivity was added to the model, distance from optimal water temperature, the variation in *V. cholerae* described by the model improved from 31.7% to 50.4% (Table 9.2). For the freshwater sites alone no significant explanatory variables were found, which is similar for other studies (Louis et al., 2003).

For the Caspian Sea, we determined the predicted probabilities of *V. cholerae* detection based on the computed odds-ratios for the model of water temperature fitted to the *V. cholerae* detection data. Water temperatures were chosen as opposed to optimal water temperatures since historical water temperature data collected at the Caspian Sea could be fitted to these results. Predictive probabilities of *V. cholerae* detection

demonstrate a clear increase as water temperatures increases and ranges from $p = 0.045$ at the lowest recorded temperature (4°C) to $p = 0.56$ at the highest recorded temperature (28.4°C) (Figure 9.4). Quartiles of water temperatures recorded at time of collection were superimposed on this histogram (black bars) to demonstrate the increased probability of *V. cholerae* detection during seasonal shifts in water temperatures. Applying historical averages of water temperatures recorded in the Caspian Sea to this model suggests that historically *V. cholerae* can be detected most frequently in July, August, September, and October (Gurbanov et al., 2012). Our *V. cholerae* isolation data demonstrates that the organism was detected in 50% of sites sampled in July 2010, 66% of sites in July 2011, 40% of sites sampled in August, but at no sites in October. It should be noted that although the probability of *V. cholerae* detection increases with increasing temperatures and is highest in these months, it is still only 0.56 at its highest prediction based on water temperature. However, these data do demonstrate a clear trend that is confirmed by other studies in other parts of the world (Louis et al., 2003; Huq et al., 2005; Constantin de Magny et al., 2008; Constantin de Magny et al., 2009; Turner et al., 2009).

The detection of *V. cholerae* in Jeyranbatan reservoir water was a significant finding, since this is the source of drinking water for Baku and the entire Absheron peninsula. We detected *V. cholerae* during 3/7 (27%) sampling rounds. A second water reservoir, the Mingachevir reservoir, located in northern Azerbaijan, was also positive for *V. cholerae* in July, 2010. Although these isolates were not "epidemic strains", however, they are a significant because studies have shown, mobile genetic elements such as integrating filamentous phage CTXφ, encoding the cholera toxin (CT) gene, can make a

lateral transfer to nontoxigenic *V. cholerae* transforming to toxigenic (Waldor and Mekalanos, 1996). In addition, there have been reports on morbidity and mortality worldwide caused by non-toxigenic *V. cholerae* non-O1/non-O139 (Safrin et al., 1988; Ko et al., 1998; Lukinmaa et al., 2006; Shannon and Kimbrough, 2006; Chatterjee et al., 2009). When municipal waters are not treated prior to public distribution, and these reservoirs were the source of water, there were cholera outbreaks in Azerbaijan (Gurbanov et al., 2012). It is concluded that the bodies of water in Azerbaijan should be monitored as *V. cholerae* appears to be widely distributed in the aquatic environment of Georgia (Grim et al., 2010). It was interesting to note that water temperature and conductivity can significantly explain some deviance in the detection of *V. cholerae* in the Caspian Sea with predictive probabilities of *V. cholerae* detection increasing as temperatures increase. Applying historical averages of water temperatures recorded in the Caspian Sea to this model suggests that detection rates should be highest in August when condition are favorable it's growth. Results from a retrospective study on the data collected over 28 years, showed an increase of air temperatures directly corresponds to increased detection of *V. cholerae* (Gurbanov et al., 2012).

Our models fit the data with highly significant, albeit relatively low pseudo-$R^2$ values (% variation in *V. cholerae* explained by the environmental parameter) suggesting there are other factors not accounted for in this study may influence *V. cholerae* detection in this environment other than temperature. Pollution in this region is known to have a dramatic influence on macrobiota in the Caspian Sea and it may greatly influence microbial diversity in this environment as well (Dumont, 1995). There is evidence that pollution similar to that which occurs in the Caspian region influences aquatic microbial

communities and these sources of pollution, such as polyaromatic hydrocarbons (PAHs), can have as significant an influence on microbial communities as do seasonal shifts in environmental parameters (Djhanguiri et al., 1997; Lemke et al., 1997; Langworthy et al., 1998; Tolosa et al., 2004). However, evidence of this is lacking in the Caspian Sea due lack of research of this topic, primarily due to the loss of funding for scientific research after the fall of the Soviet Union. As previous models of some infectious diseases have utilized data on accelerated climate change due to industrial pollution future studies and models of aquatic bacterial pathogens should attempt to identify and quantify sources of pollution other than those which influence global air and water temperatures and estimate their influence on the ecological dynamics of these pathogens. In conclusion, we suggest employing appropriate method, it is important to monitor environmental waters, used for recreational purposes, and especially if it is a source of drinking water as a preventive measure for public health safety.


Figure 9.1. Samples were collected from nine sites, including a reservoir, lakes, a river, and the Caspian Sea, as shown in Figure 9.4 over a 15 month period, from May, 2010 to July, 2011. Bilgah, Buzovnha, Shikhov, Khanlar and Novkhani are located along the Caspian Sea, on or near the Absheron peninsula where Baku is located and the majority of the Azerbaijani population lives. Novkhani is located near the city of Sumqayit, a heavily polluted industrial town developed during the Soviet era (Islamzade, 1994). Capsian Sea sites were sampled monthly. Fresh water sites are the Kura River, Lake Karkhana, and Jeyranbatan and Mingachevir reservoirs. The Kura River, Lake Karkhana, and Mingachevir reservoir sites were sampled from May to October and the Jeyranbatan site was sampled monthly. Approximate locations of the three largest cities, Baku (the capitol), Sumqayit, and Ganja are bordered by a black box.

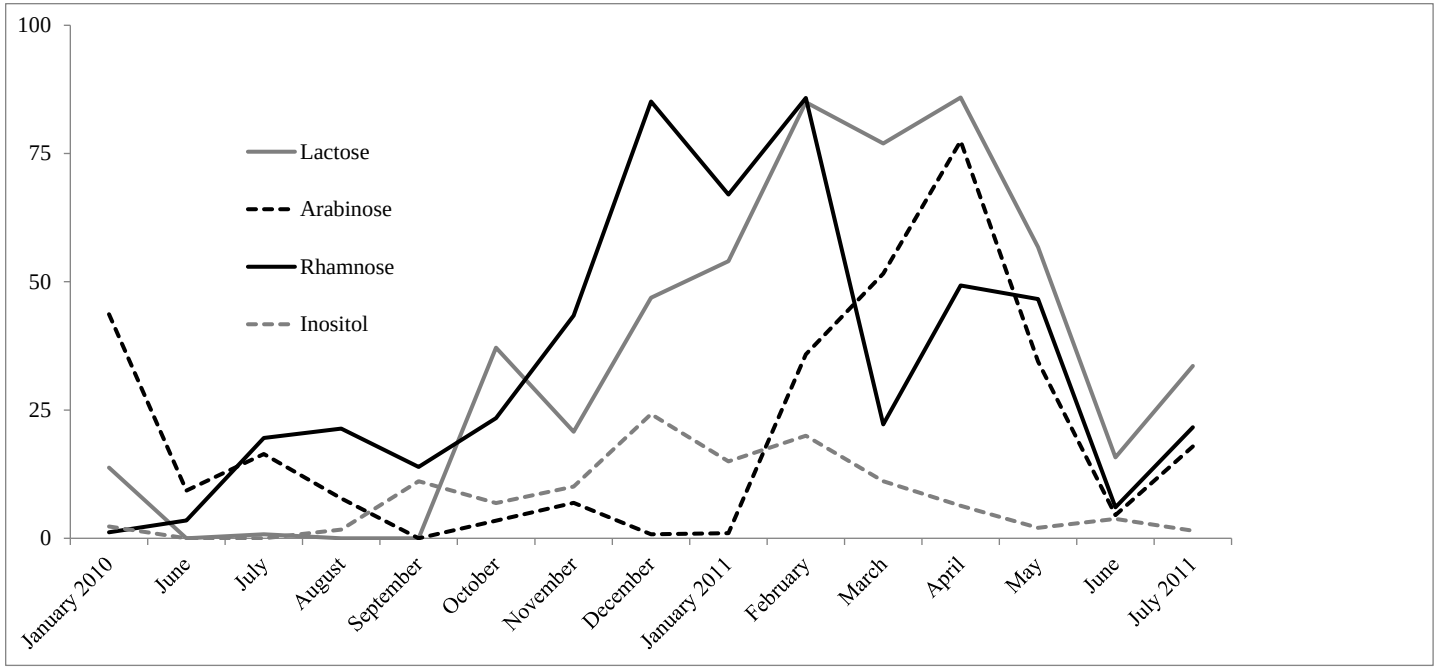



Figure 9.2. Seasonality in biochemical profiles of the presumptive *V. cholerae* strains.

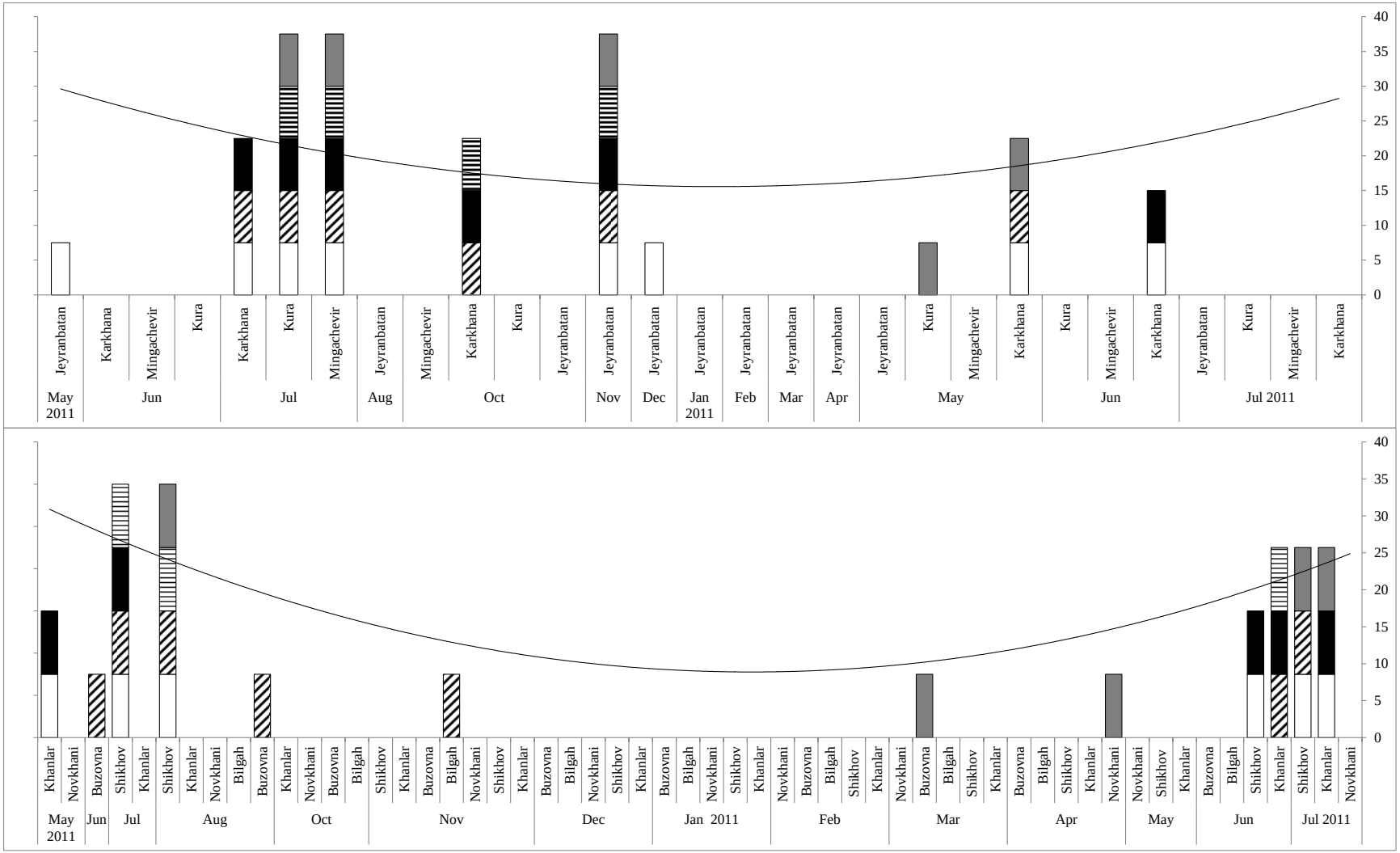



Figure 9.3. Monthly variation in *V. cholerae* detection at sampling locations in Azerbaijan. Freshwater location are indicated in th top figure and Caspian Sea locations are on the bottom figure. Stacked colored bars indicate the presence of *V. cholerae*

in < 20 μm free water fraction (FW, grey bars), 20-μm plankton fraction (PL20, white bars), 64-μm plankton fraction (PL64, diagonally striped bars), homogenized 20-μm plankton fraction (HOM20, black bars), and homogenized 64-μm plankton fraction (HOM64, horizontally striped bars). The black line (Y-axis on the right side) indicates percent of pooled locations positive for *V. cholerae* for each month.

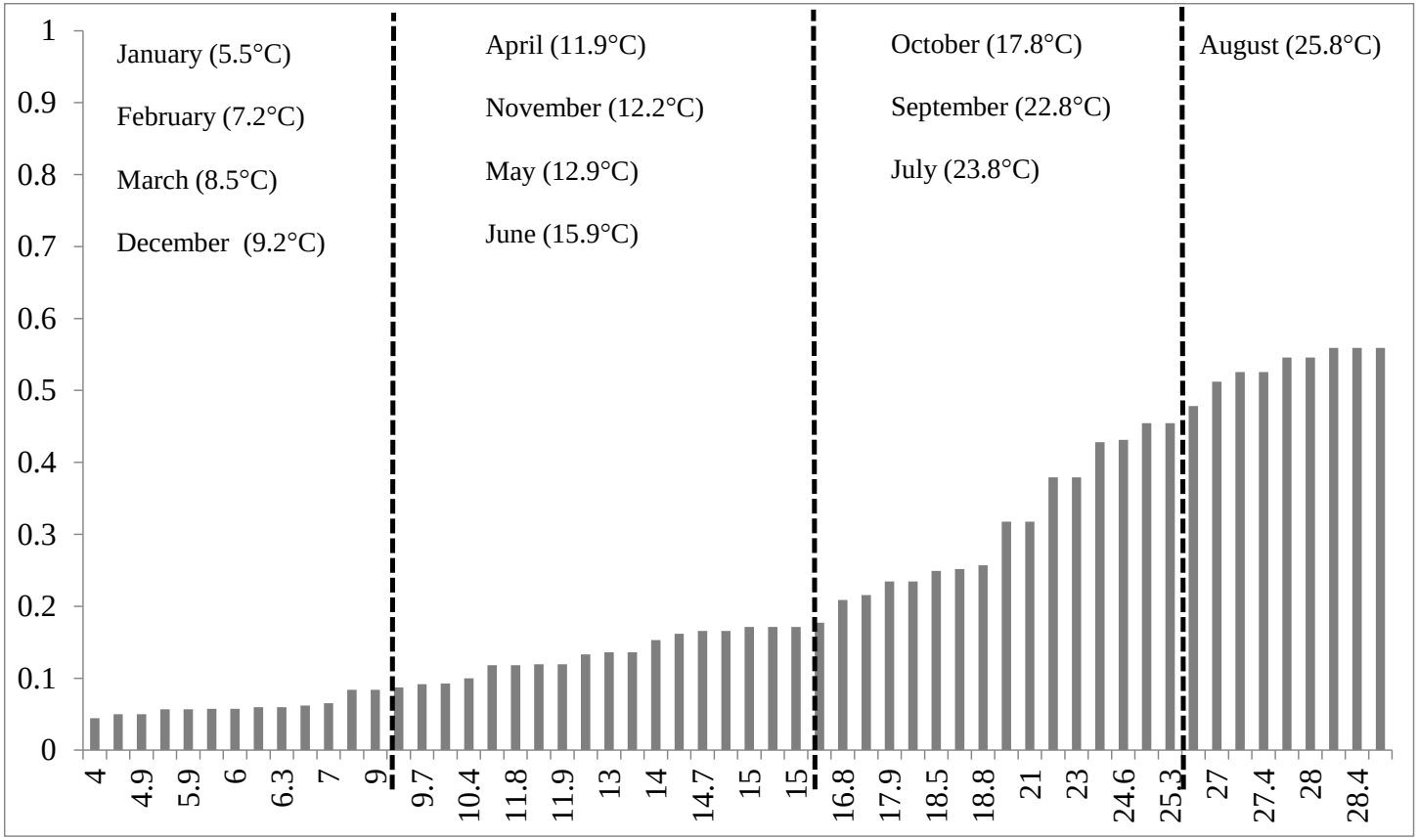



Figure 9.4. Predicted probabilities of *V. cholerae* detection (Y-axis) based on water temperatures (°C) in the Caspian Sea (X-axis). Hashed vertical lines demark quartiles of water temperatures and historical monthly temperature averages (Gurbanov et al., 2012), are assigned to the corresponding quartile to estimate the predicted probability of *V. cholerae* detection per month.

| Saline/Fresh | Season [A] | Month | Water Temperature (°C) | Air Temperature (°C) | Salinity (ppt) | Conductivity | pH | Percentage of Sites Positive for *V. cholerae* |
|---|---|---|---|---|---|---|---|---|
| Caspian Sea | Spring | May 2010 | 16 | 20.5 | 10.6 | 15.9 | 8.3 | 50 |
| | Summer | June | 27.5 | 32.7 | 10.6 | 18.5 | 8.3 | 50 |
| | | July | 27.7 | 32 | 14.6 | 14.6 | 8.3 | |
| | | August | 27.2 | 33 | 10.7 | 18.5 | 8.3 | |
| | Autumn | October | 17.7 | 22.8 | 10.4 | 16.6 | 8.3 | 18 |
| | | November | 14.9 | 14 | 10.1 | 15.5 | 8.3 | |
| | Winter | December | 12.0 | 13.4 | 10.9 | 16.4 | 8.3 | 5 |
| | | January 2011 | 9 | 10.2 | 10.6 | 15.2 | 8.3 | |
| | | February | 5.6 | 5.2 | 10.1 | 14.1 | 8.3 | |
| | | March | 5.8 | 6.2 | 10.4 | 13.6 | 8.2 | |
| | Spring | April | 10.9 | 12.3 | 10.6 | 15.6 | 8.1 | 17 |
| | | May | 16.6 | 20.3 | 10.7 | 16.8 | 8.2 | |
| | Summer | June | 22 | 27.8 | 10.2 | 18 | 8.2 | 57 |
| | | July2011 | 25 | 27 | 10.8 | 18.2 | 8.2 | |
| Fresh Water | Spring | May 2010 | 18 | 24.0 | 0.4 | 735.5 | 7.9 | 10 |
| | Summer | June | 25.4 | 28.3 | 0.6 | 481.9 | 8.0 | 43 |
| | | July | 26.9 | 28 | 0.5 | 447.0 | 8.0 | |
| | | August | 28.7 | 32 | 0.3 | 699.5 | 8.1 | |
| | Autumn | October | 19.8 | 23.3 | 0.5 | 461.8 | 8.0 | 40 |
| | | November | 14.4 | 16 | 0.4 | 664.5 | 7.9 | |
| | Winter | December | 11.3 | 13.0 | 0.4 | 686.0 | 8.3 | 25 |
| | | January 2011 | 13 | 13.0 | 0.4 | 615.0 | 7.9 | |
| | | February | 6.2 | 4.0 | 0.4 | 718.5 | 7.9 | |
| | | March | 8.8 | 8.0 | 0.4 | 711.0 | 8.8 | |
| | Spring | April | 12.5 | 15.1 | 0.3 | 519.8 | 8.3 | 20 |
| | | May | 15.6 | 20.0 | 0.3 | 548.0 | 8.3 | |
| | Summer | June | 25 | 30.0 | 0.2 | 443 | 8.3 | 0 |
| | | July2011 | 26 | 34 | 0.5 | 516.3 | 8.2 | |

[A] based on Huseynov and Malikov (2009)(Huseynov and Malikov, 2009)

Table 9.1. Average monthly environmental parameters and percentage of sampled sites that were positive for *V. cholerae* (far right column) at Caspian Sea and Freshwater sites. Seasons in Azerbaijan as determined by Huseynov and Malikov (2009) are listed in the second column from the left. The figure is shaded by quartiles of recorded water temperatures. The darkest shade represents the coldest recorded water temperatures with decreasing darkness of shade representing increasing water temperatures.

| Sites | Environmental Parameters | Coefficient | P-Value | Deviance Explained (%) |
|---|---|---|---|---|
| Bu, Bi ,N ,K, S | Distance from Optimal Water Temperature | -0.1907 | 0.0085 | 14.61 |
| | Water Temperature | 0.1255 | 0.0075 | 13.34 |
| K,N,S | Distance from Optimal Water Temperature | -0.2948 | 0.00976 | 25.43 |
| | Water Temperature | 0.14178 | 0.01336 | 16.96 |
| K, S, Bi | Distance from Optimal Water Temperature + Distance from Optimal Conductivity | -0.7903<br>1.7112 | 0.00612<br>0.02685 | 50.39 |
| | Distance from Optimal Water Temperature | -0.03518 | 0.00772 | 31.7 |
| | Water Temperature | 0.1779 | 0.00724 | 23.02 |
| K,S | Distance from Optimal Water Temperature | -0.4405 | 0.0111 | 43.76 |
| | Water Temperature | 0.2 | 0.009 | 27.8 |

Table 9.2. Statistical models of *V. cholerae* detection with logististic regression. Bu = Buzovnah, Bi = Bilgah, N = Novkhani, K = Khanlar, S = Shikhov.

# Chapter 10: Comparison of Methods for Quantifying Bacterial Indicators in an Urban Brackish Water Environment


***Abstract***

The United States EPA and European Community Bathing Water Directive recommend testing the levels of *Escherichia coli* and enterococci in surface waters as proxies for the presence of human enteric pathogens. Similarly, international and United States regulations for ships' ballast water discharge include acceptable limits for *E. coli* and enterococci. In this report we present the results of a comparative study of standard membrane filtration methods and recently developed enzyme substrate methods, Colilert and Enterolert (IDEXX Laboratories, Inc.), for detection of *E. coli* and enterococci in an urban brackish water environment at the Port of Baltimore. Enterolert and Colilert assays showed significant and positive correlation with analogous membrane filtration methods, $\rho = 0.60$ for modified $mTEC_{44.5^oC}$, and, $\rho = 0.55$ for mEA-$BEA_{44.5^oC}$. Microbial concentrations were significantly higher for membrane filtration assays incubated at Enterolert and Colilert recommended temperatures (41ºC and 35ºC, respectively), thereby producing stronger correlations, $\rho = 0.89$ for modified $mTEC_{35^oC}$ and $\rho = 0.65$ for mEA-$BEA_{41.5^oC}$. These results indicate that the membrane substrate methods, Enterolert and Colilert, can overestimate the target bacterial populations because of lower incubation temperatures compared to standard methods, most likely by allowing growth of non-fecal bacteria.

***Introduction***

Waterborne outbreaks of gastroenteritis are typically associated with exposure to anthroponotic and zoonotic pathogens in recreational, drinking, and irrigation waters. Many of these organisms are autochthonous to the aquatic environment while others are not and both types have been shown to persist and remain infectious for extended periods of time in these environments (Buswell et al., 1998; Fayer et al., 1998, Mezrioui et al., 1995). Some of these pathogens have been suggested to be introduced into the aquatic environment via urban and agricultural runoff, as well as from the presence of wildlife and their utilization of watersheds (Haley et al., 2009; Lipp et al., 2001: MacKenzie et al., 1994). However, viability and culturability of these pathogens may vary outside of the host. This presents a problem when attempting to predict the overall public health safety of the aquatic environment, with respect to pathogenic micoorganisms. To address this problem, suites of bacteria of fecal-origin, such as *E. coli* and enterococci, are used as proxies for the presence of enteric pathogens of human and animal origin.

In 1976 the USEPA and the European Community Bathing Water Directive called for the use of coliforms as indicators of water quality; however it has been demonstrated that environmental samples contain a large fraction of these bacteria that are not of fecal origin, such as *Klebsiella* and *Citrobacter* (Leclerc et al., 2001), thus making their use as indicators of water quality questionable. Epidemiological studies later demonstrated that the numbers of enterococci and *Escherichia coli* bacteria in samples collected at several freshwater and coastal beaches were directly

related to cases of gastroenteritis in the aquatic environment. Based on the results of these studies the USEPA adopted the use of enterococci and *Escherichia coli* as proxies for estimating public health safety of recreational water. *Escherichia coli* has been suggested to be a specific indicator of fecal pollution because of its abundance in fecal matter ($10^9$ $g^{-1}$) and inability to replicate outside the host under certain environmental conditions, while persisting at least the same length of time as other fecal pathogens discharged into aquatic environments. Enterococci are considered to be reliable indicators of fecal pollution because of their limited host range (humans, dogs, and chickens) (Wheeler et al., 2002), but this has been disputed and is currently under scrutiny (Jackson et al., 2007; Roslev et al., 2004).

To examine the abundance of *E. coli* and enterococci in the aquatic environment, several methods have been developed to improve accuracy and speed of detection. Erroneous high estimates of the concentrations of fecal indicator bacteria can lead to unnecessary closing of recreational waters causing substantial economic losses, while erroneous low concentrations will pose a public health threat for those exposed to the water. Current USEPA and European Union methods call for membrane filtration and incubation of the filters on selective media specific for growth of either enterococci or *E. coli*. Improvements in the test methods over have increased the specificity of detection, as well as reduced the time of analysis. For example, currently approved USEPA methods include the use of membrane-thermotolerant *E. coli* agar (mTEC) (USEPA Method 1103.1) and modified membrane-thermotolerant *E. coli* agar (modified mTEC) (USEPA Method 1603) for detection of thermotolerant *E. coli*. These tests require media preparation and quality

control testing prior to use. Furthermore, prepared media have a relatively short shelf life requiring media to be repeatedly prepared over time. To address this issue, rapid microbiological test kits for detection and quantification of *E. coli* and enterococci in water samples have been developed and are commercially available. The Enterolert and Colilert kits (IDEXX Laboratories, Inc., Westbrook, ME) involve two steps, inoculation of sample in dehydrated selective media and incubation of this inoculum in a sealed, multiwelled tray, yielding an estimation of cell concentration by the Most Probably Number (MPN) method.

This study compares two standard membrane filtration methods for quantification of thermotolerant *E. coli* and enterococci with Colilert and Enterolert systems. Results of each method are compared with heterotrophic bacteria (HPC) counts, thus the objective of this study was to determine whether the cell concentration estimations could be correlated and thereby provide the same results.

## ***Materials and Methods***

### Sample Collection

Water samples were collected twice weekly between August and October 2008 and between April and May 2009 from the Maritime Environmental Resource Center (MERC) ballast water treatment test facility, located onboard the M/V *Cape Washington* docked in the Port of Baltimore, Patapsco River, Maryland, USA. Five replicate one liter samples were collected at each sample collection time in sterile 1

liter polypropylene bottles and transported on ice. Samples were processed in the laboratory for microbial examination within four hours of sample collection.

**Enumeration of *Escherichia coli* on Modified mTEC Agar**

Following USEPA Method 1603; volumes of 1mL, 10 mL and 100 mL of water were passed through a 0.45 µm nitrocellulose membrane filter which was then placed on modified thermotolerant *E. coli* agar (modified mTEC) (Becton Dickson, Sparks, MD). Plates containing membranes were pre-incubated for two hours at 35ºC to enhance recovery of injured and stressed cells and incubated in a sealed Whirl-Pack® bags (Nasco, Fort Atkinson, WI) at 44.5ºC in a waterbath for 22-24 hours. The protocol was followed in parallel, except that the final incubation temperature was 35 ºC, to employ the same incubation conditions as for the IDEXX Colilert detection kit. After incubation, red and magenta colonies were enumerated as *E. coli* and recorded as *E. coli* colony forming units per 100 mL (CFU 100 $mL^{-1}$) of sample water.

**Enumeration of Enterococci on mE and BE Agar**

For enumeration of enterococci by membrane filtration, 10 mL and 100 mL of water were passed through a 0.45 µm nitrocellulose membrane, which was then transferred onto mEnterococcus agar (mEA, Becton Dickson, Sparks, MD) and incubated at 44.5ºC for 48 hours. Membranes containing light and dark red colonies were transferred to bile-esculin agar (BEA, Becton Dickson, Sparks, MD) and incubated for 4 hr (Figueras et al., 1996). The process was followed in parallel, with

a final incubation temperature of 41ºC to employ the same incubation conditions of the IDEXX Enterolert detection kit. After incubation, dark colonies with a black halo were scored as enterococci and reported as enterococci colony forming units per 100 mL (CFU 100 $mL^{-1}$).

### Enumeration of *Escherichia coli* and Enterococci Using IDEXX Kits

*E. coli* and enterococci concentrations were also estimated using the IDEXX Colilert and Enterolert kits, respectively (IDEXX Laboratories, Inc., Westbrook, ME). One-hundred milliliters of sample water were mixed in IDEXX resuspension bottles with one pack of the appropriate media for each test. This mixture was poured into a Quantitray and sealed. Enterolert trays were incubated at 41 ± 0.5ºC for 24 hours. Colilert trays were incubated at 35 ± 0.5ºC for 24 hours. When *E. coli* metabolizes 4-methyl-umbelliferyl-B-D-glucuronide (MUG), the nutrient-indicator, in the enrichment medium, the sample fluoresces. When enterococci metabolizes 4-methyl-umbelliferyl-B-D-glucoside (MUD), the nutrient-indicator, in the enrichment medium, the sample fluoresces. Fluorescence, determined with a 6-watt, 365 nm hand-held UV lamp within 5 inches of the sample in a dark environment, indicated a positive score for the presence of both enterococci and *E. coli*. Indicator densities were recorded as Most Probable Number (MPN) per 100 mL.

### Enumeration of Heterotrophic Bacteria

Heterotrophic bacteria were enumerated by plating a three-fold serial dilution of water sample, in triplicate, onto heterotrophic plate count (HPC) agar in 100 µl volumes. HPC agar plates were subsequently incubated at 30ºC for five days and heterotrophic bacteria densities were reported as CFU 100 $ml^{-1}$.

### Statistical Analyses

Data were tested for normality using the Kolmogorov-Smirnov and Anderson-Darling tests. Subsequently no data were determined to be normally distributed, even after logarithmic and square root transformations, ($P < 0.05$ for all analyses). Therefore non-parametric statistical analyses were performed on all data. To evaluate correlation between methods, Spearman's rank order correlations were determined between all enterococci detection methods and all *Escherichia coli* methods. Wilcoxon's signed ranks test was used to determine if different methods yielded different microbial concentrations. For all analyses, P-values less than 0.05 were considered statistically significant.

## ***Results and Discussion***

In this study *E. coli* concentrations were higher than enterococci densities in 48% of the samples collected, while enterococci levels were higher than *E. coli* in 13% of samples collected. Thirty-four percent of the samples tested for *E. coli*

exceeded 100 MPN 100 $mL^{-1}$ using the Colilert test and 5% and 0% of samples using membrane filtration at 35ºC and 44.5ºC, respectively. Moreover, 5% of the samples exceeded *E. coli* concentrations of 1,000 MPN 100 $mL^{-1}$ and 1% were recorded at the upper detection limit of 6,867 MPN 100 $mL^{-1}$, using the Colilert test. Enterococci concentrations were relatively low and did not exceed 100 MPN 100 $mL^{-1}$ with Enterolert or 22 CFU 100 $mL^{-1}$ with membrane filtration. Neither bacterial indicator at both incubation temperatures was detected in 42% of the samples in which all IDEXX and membrane filtration assays were run, which usually corresponded to treated water samples. Heterotrophic bacteria were detected in all samples including those in which enterococci and *E. coli* were not detected. The mean HPC count for all samples was $2.5 \times 10^4$ CFU $mL^{-1}$ and mean HPC concentration for samples in which no other bacterial indicator was detected was $5.5 \times 10^2$ CFU 1 $mL^{-1}$ (Table 10.1).

**Comparison of IDEXX Dilutions**

Following recommendations of the IDEXX system for analysis of water with greater than 5 ppt salinity, one 10-fold dilution (10-1) assay for each IDEXX assay was run in parallel with one undiluted (100) assay for eleven Enterolert samples and seventeen Colilert samples. Results of Wilcoxon's signed rank test showed no significant difference in microbial concentrations between dilutions for each assay.

**Enterococci Densities**

The number of enterococci ranged from undetectable to 22 CFU 100 $ml^{-1}$ at 44.5ºC (n = 134) and 21 CFU 100 $ml^{-1}$ at 41.5ºC (n = 116) on mEA-BEA (limit of detection = 1 CFU 100 $ml^{-1}$) and from undetectable to 93.5 MPN 100 $ml^{-1}$ (n = 134) using the Enterolert test (limit of detection = 1 MPN 100 $ml^{-1}$) (Table 10.1). These data suggest that water samples collected in this study were below the threshold for safe water since the enterococci counts did not exceed the one-time sampling maximum threshold of 104 CFU 100 $mL^{-1}$ for recreational beaches or the thresholds for moderately used marine waters which have less stringent standards . By the Wilcoxon Signed Rank test, enterococci counts were significantly higher on mEA-BEA plates incubated at 41.5ºC (mean = 1.6 CFU 100 $mL^{-1}$) than on mEA-BEA plates incubated at 44.5ºC (mean = 1.2 CFU 100 $mL^{-1}$) ($P < 0.0001$).

Thirty-seven percent of the Enterolert assays did not detect enterococci, while 68% of mEA-BEA assays incubated at 41.5ºC and 44.5ºC, respectively, did not detect enterococci. There were 53 samples in which the Enterolert assay detected enterococci (range = 1 to 53.3 MPN 100 $ml^{-1}$, mean = 9.4 MPN 100 $ml^{-1}$) and the mEA-BEA assay incubated at 41.5ºC did not detect enterococci and 32 samples in which the Enterolert assay detected enterococci (range = 1 to 63.5 MPN 100 $ml^{-1}$, mean = 11.4 MPN 100 $ml^{-1}$) and mE-$BEA_{44.5ºC}$ did not detect enterococci. There were only four assays in which mE-$BEA_{44.5ºC}$ detected enterococci (1 to 4 CFU 100 $ml^{-1}$) and the Enterolert assay did not detect enterococci. There were no assays in which mE-$BEA_{41.5ºC}$ detected enterococci and the Enterolert assay did not. Enterococci numbers determined by membrane filtration at both temperatures were

weakly correlated with each other as determined by Spearman's rank order correlation ($\rho = 0.60$, $P < 0.0001$).

### *Escherichia coli* Densities

*E. coli* numbers ranged from undetected to 157 CFU 100 100 $ml^{-1}$ at 35ºC (n = 116) and 58 CFU 100 $ml^{-1}$ (n = 134) at 44.5°C on modified mTEC, and from undetected to >6.8 x $10^3$ MPN 100 $ml^{-1}$ (n =129) using Colilert (lower limit of detection = 1 MPN 100 $mL^{-1}$, upper limit of detection = 6.8 x $10^3$ MPN 100 $mL^{-1}$) (Table 10.1). By the Wilcoxon Signed Rank test, the *E. coli* numbers were significantly higher on modified $mTEC_{35ºC}$ (mean = 25.4 CFU 100 $mL^{-1}$) than modified $mTEC_{44.5ºC}$ (mean = 6.1 CFU 100 $mL^{-1}$) ($P < 0.0001$). The *E. coli* numbers as determined by both methods correlated weakly with each other by Spearman's rank order correlation ($\rho = 0.68$, $P < 0.05$).

Thirty-six percent of Colilert assays did not detect *E. coli* while 43% and 53% of modified mTEC assays incubated at 35ºC and 44.5ºC, respectively, did not detect *E. coli*. There were twenty-six samples in which Colilert detected *E. coli* and modified $mTEC_{44.5ºC}$ did not detect *E. coli* (range = 1 to 6867 MPN 100 $mL^{-1}$, mean = 550.7 MPN 100 $mL^{-1}$). There were five samples in which Colilert detected *E. coli* and the modified $mTEC_{35ºC}$ assay did not detect *E. coli* (range = 1 to 2 MPN 100 $mL^{-1}$, mean = 1.2 MPN 100 $mL^{-1}$). There was one sample in which the $mTEC_{35ºC}$ assay detected E. coli (10 CFU 100 $ml^{-1}$) and the Enterolert assay did not detect *E. coli*. Excluding this one sample, all Colilert counts that detected *E. coli* were higher than mTEC counts for plates incubated at both temperatures. Similar observations have

been made in microcosm experiments comparing IDEXX, membrane filtration, and multiple tube fermentation assays (Noble et al. 2003).

**Correlation between Methods**

From the results of the statistical analyses employing Spearman's rank order correlation, the Enterolert and Colilert MPN methods of detection were correlated with the analogous membrane filtration method (mEA-BEA and modified mTEC) at all evaluated incubation temperatures ($P < 0.0001$) (Figures 10.1 and 10.2). The correlations were stronger between Colilert and modified mTEC ($\rho = 0.89$ for modified $mTEC_{35^oC}$, $\rho = 0.60$ for modified $mTEC_{44.5^oC}$) (Figure 10.1) (Table 10.2) than between Enterolert and mEA-BEA ($\rho = 0.65$ for $mEA\text{-}BEA_{41.5^oC}$, $\rho = 0.55$ for $mEA\text{-}BEA_{44.5^oC}$) (Figure 10.2) (Table 10.2). To investigate whether the low correlations were due to low microbial concentrations in the samples for which IDEXX detected the target and MF methods did not, all sample-points in which IDEXX detected a target and MF methods did not were removed and the analyses redone. Correlations in this second round of analyses showed only a moderately stronger correlation between enterococci methods ($\rho = 0.69$ and 0.56 for analyses at 41.5ºC and 44.5ºC respectively). However, for *E. coli* the analyses showed a weaker correlation ($\rho = 0.61$, $P < 0.0001$ for $mTEC_{35^oC}$) and non-significant correlation for $mTEC_{44.5^oC}$ ($P > 0.05$). The data indicate the IDEXX and MF assays were inconsistent in detection if either bacterial indicator was present near the limit of detection, an inconsistency that did not influence the correlation between methods, however.

**Correlation between Incubation Temperatures**

Although neither was strong, the correlation between Enterolert and mEA-BEA$_{44.5^{o}C}$ was weaker ($\rho$ = 0.55) than the correlation between Enterolert and mEA-BEA$_{41.5^{o}C}$ ($\rho$ = 0.65), suggesting incubation of mEA-BEA plates at the incubation temperature for Enterolert (41.5ºC) yielded results more consistent with Enterolert MPN concentrations. Correlations between Colilert and modified mTEC were moderately high at both temperatures but stronger for modified mTEC incubated at 35ºC ($\rho$ = 0.89 mTEC$_{35^{o}C}$ and $\rho$ = 0.60 for mTEC$_{44.5^{o}C}$). The USEPA standard method recommends incubating modified mTEC plates at 44.5ºC to select for thermotolerant *E. coli* strains theoretically originating from the gut of warm-blooded animals, as opposed to environmentally adapted strains that may grow at lower temperatures. The stronger correlation between the Colilert assay and modified mTEC$_{35^{o}C}$ suggests that the Colilert assay allows growth of a wider range of *E. coli* strains and does not select for thermotolerant strains.

**Correlation of HPC Counts with the Number Enterococci and *E. coli***

Heterotrophic bacteria counts showed no significant correlation with mEA-BEA concentrations at both 41ºC and 44.5ºC. All other bacterial indicator concentrations showed significant correlation with HPC counts, although the correlation was weak for enterococci by Enterolert ($\rho$ = 0.39) and *E. coli* densities determined by mTEC$_{44.5^{o}C}$ ($\rho$ = 0.54) (Table 10.2). The correlation between *E. coli* determined by mTEC$_{35^{o}C}$ ($\rho$ = 0.75) and Colilert ($\rho$ = 0.72) was higher (Table 10.2)

and was interpreted to reflect enhanced growth of environmental bacteria at the lower incubation temperature.

***Conclusions***

It is concluded ftom the results of this study that the estimates of the Colilert and modified mTEC assay incubated at 35ºC are significantly correlated. However, the USEPA recommends incubation for thermotolerant *E. coli* on modified mTEC agar should be 44.5ºC. Thus, although the correlation is strong, the Colilert results most likely include environmentally adapted enteric *E. coli*, as well as non-enteric (non-thermotolerant) *E. coli*. Thus, if one uses the recommended incubation temperature of the Colilert manufacturer and the incubation temperature for modified mTEC recommended by the USEPA, the thermotolerant *E.coli* estimates using modified mTEC will differ from total *E. coli* counts using Colilert. The other correlations between IDEXX and MF were significant, but weak, and the same observation was also noted for HPC and the bacterial indicator assays. These results demonstrate a serious inconsistency in results of methods recommended for detecting the level of indicators of fecal pollution in surface waters, with serious implications for the public health safety of bodies of water in the natural environment. Given the state, national, and international regulations for microbial indicators of public health significance it is perhaps time to consider substitution of more precise methods for direct detection of the human pathogens rather than indicators of their presence.

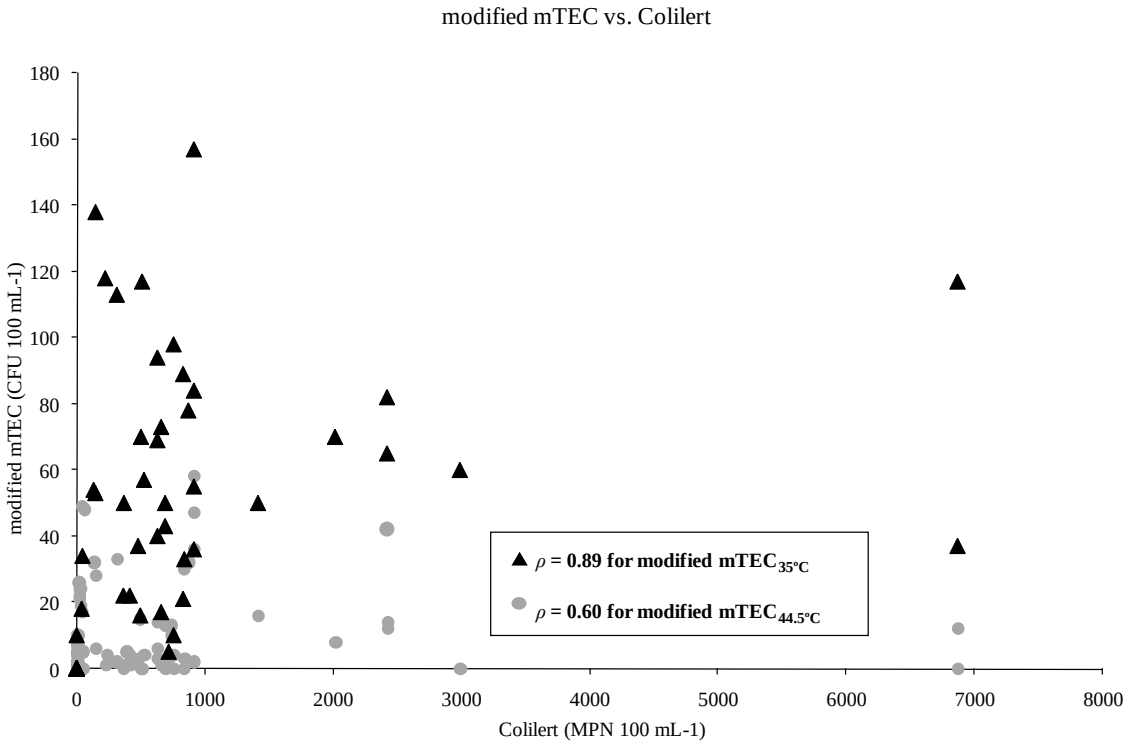



Figure 10.1. Scatter plot of *E. coli* concentrations using by modified mTEC agar and Colilert assays.

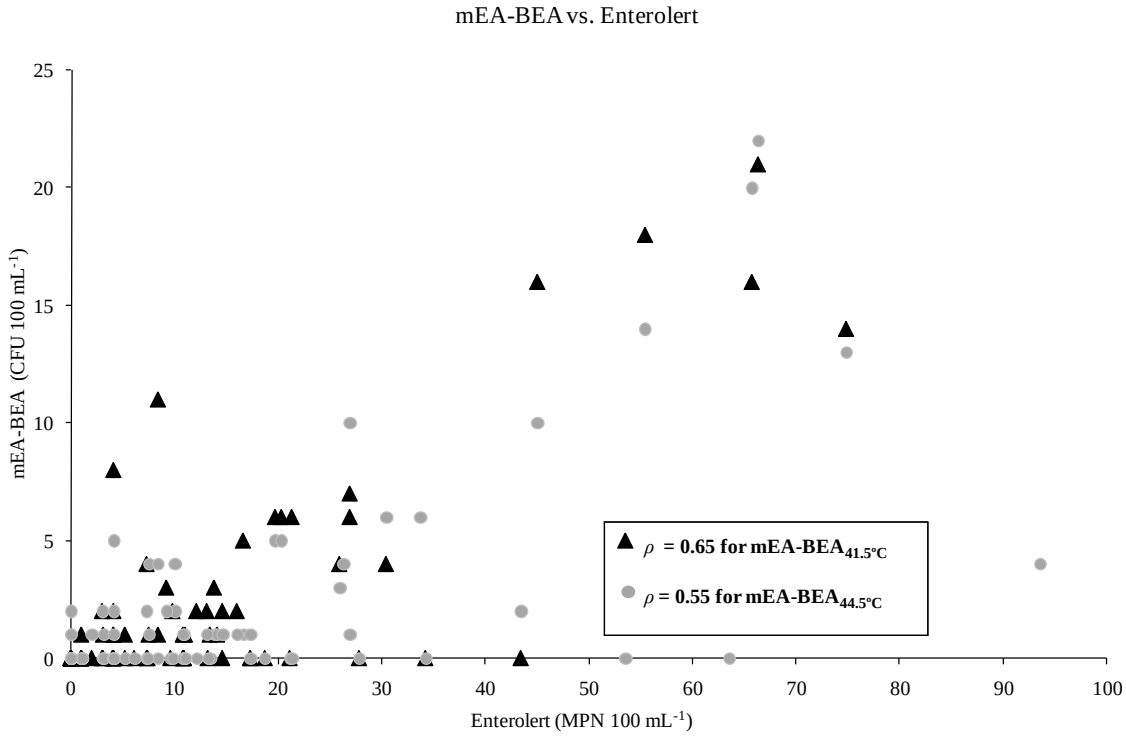



Figure 10.2. Scatter plot of enterococci concentrations using mEA-BEA agar and Enterolert assays.

| Statistic | Enterolert | mEA-BEA 41.5ºC | mEA-BEA 44.5ºC | Colilert | modified mTEC 35ºC | modified mTEC 44.5ºC | HPC |
|---|---|---|---|---|---|---|---|
| N | 134 [a] | 116 [b] | 134 [b] | 129 [a] | 116 [b] | 134 [b] | 120 [a] |
| Mean | 9.9 | 1.6 | 1.27 | 385.8 | 25.3 | 6.1 | 16648 |
| Median | 3.5 | 0 | 0 | 12 | 4 | 0 | 454500 |
| Maximum Value | 93.5 | 21 | 0 | 6867 | 157 | 58 | 16648.5 |
| Minimum Value | 0 | 0 | 11.0 | 0 | 0 | 0 | 21.5 |
| Range | 93.5 | 21 | 22 | 6867 | 157 | 58 | 454479 |
| Standard Deviation | 16.7 | 3.84 | 3.37 | 977.1 | 35.9 | 11.7 | 74244 |
| Percent Not Detected | 37 | 68 | 69 | 36.4 | 43 | 53 | 0 |

[a] = Colony Forming Units 100 ml $^{-1}$

[b] = Most Probably Number 100 ml $^{-1}$

Table 10.1. Descriptive statistics of the bacterial indicator assays used in this study.

| Assay | Enterolert | mEA-BEA 41.5ºC | mEA-BEA 44.5ºC | Colilert | modified mTEC 35ºC | modified mTEC 44.5ºC |
|---|---|---|---|---|---|---|
| mEA-BEA 41.5ºC | 0.65 [a] | 1 | 0.60 | ND | ND | ND |
| mEA-BEA 44.5ºC | 0.55 | 0.60 | 1 | ND | ND | ND |
| modified mTEC 35ºC | ND | ND | ND | 0.89 | 1 | 0.68 |
| modified mTEC 44.5ºC | ND | ND | ND | 0.60 | 0.68 | 1 |
| HPC | 0.39 | NS | NS | 0.72 | 0.75 | 0.54 |

[a] = Spearman's rank order correlation $\rho$-value

ND = Not Determined

NS = Not Significant ($P > 0.05$)

Table 10.2. Summary of correlation between methods determined by Spearman's rank order correlation analysis.

# Chapter 11: Summary and Conclusions

Results of this research contribute to knowledge of the environment as a natural reservoir of human pathogens from both an ecological and genomic perspective. Genomic analyses demonstrated a high level of similarity between genomes isolated from clinical cholera cases and isolates from the environment in regions where cholera occurs only sporadically. Moreover, environmental isolates whose genomes diverge significantly from genomes of selected clinical isolates encode many virulence factors and pathogenicity islands associated with cholera and *V. cholerae* infections. A newly identified species, *V. metecus*, was shown, in this study, to have a genome with < 95% nucleotide similarity with *V. cholerae*, yet encodes virulence factors and pathogenicity islands known to be integral to the disease cholera. These results suggest virulence genes are conserved among the *Vibrionaceae* and pathogenicity islands circulate across species boundaries in this family which implies they have a primary function in the environment.

From results of this study it is concluded that clinical *V. cholerae* isolates from the same outbreak are not truly clonal and have SNPs among them as well as differential expression of virulence. This study involved sequencing 73% of the clinical cases comprising a cholera outbreak in the United States and determining expression of the virulence genes. Implications drawn from the work involve understanding why cholera symptoms are not uniform among those who are infected. It is known that both the immune status of the host and the constellation of virulence factors of the pathogen contribute to the disease outcome. Since the outbreak that

was studied was not caused by *V. cholerae* from a human carrier, it is concluded that highly clonal strains can arise in the environment without human passage, enrichment, or dissemination.

The ecological analysis of *V. cholerae* in Iceland, a cholera-free country, yielded results showing that virulence genes and pathogenicity islands can be expressed at temperatures relevant to the environment from which the pathogen was isolated. For example, the virulence gene, hemaglutinnin/protease, involved in movement of *V. cholerae* in the intestine and in degradation of chironomid eggs was shown to be proteolytic at 14°C suggesting this "virulence factor" functions in the environment. Chironomids are frequently found in the intertidal zone along the coast of Iceland, where *V. cholerae* was isolated. Interestingly, the mobile pathogenicity island VPI-2, encoding a sialidase and involved in facilitating attachment of cholera toxin to gangliosides in the human intestinal epithelium, was found to be present in these strains and functional at 14°C, its environmental habitat temperature. The majority of strains isolated from water samples collected in Iceland were positive for this pathogenicity island and its conservation suggests an environmental role serving to select for its presence in that harsh environment.

The ecology of *V. metecus* in the Chesapeake Bay was studied and the results showed it has a different seasonality than *V. cholerae*. *V. metecus* densities were highest during colder months in contrast to *V. cholerae*, the densities of which were highest in warmer months of the year. *V. metecus* densities showed significant negative correlation with water temperature. Furthermore, although *V. metecus* could be isolated during warm months, in low numbers, it was readily isolated during cold

months of the year. *V. metecus* also showed a significant negative correlation with salinity suggesting it fares best in cold fresh water. This phenomenon is explained as follows. *V. metecus* grows more rapidly at temperatures at the lower end of the mesophilic temperature range and/or it is out-competed by co-occurring microorganisms during warmer months (summer), with higher numbers when these competing organisms are fewer (winter). Without data from both microcosm and mesocosm studies controlling for temperatures and inter-species competition it can be concluded that the ecology of *V. metecus* differs from that of *V. cholerae*.

The seasonality of *V. cholerae* in Azerbaijan was determined from studies carried out on water samples collected from the Caspian Sea, where it was found to be more abundant in the summer months, associated with an increase in water temperature. Results of this study showed that drinking water reservoirs in Azerbaijan are sources of cholera, corroborating earlier studies that suggested these reservoirs are the source of cholera outbreaks in this country.

*V. parahaemolyticus* occurrence in the Black Sea was significantly correlated with water temperatures. The *V. parahaemolyticus* strains that were isolated in this study were genetically diverse and many of the genomes encoded makers of the pandemic clone of V. parahaemolyticus but none were positive for major virulence factors associated with clinical cases, i.e., thermostable-direct hemolysin (TDH), thermostable-direct related hemolysin (TRH), and type three secretion system 2 (TTSS2). *V. parahaemolyticus* was found to be associated with plankton and its incidence associated with water temperature, as has been reported for *V. cholerae*.

In conclusion, the ubiquity of virulence factors in pathogenic bacteria in regions where associated illnesses are rare or do not occur at all are suggests their utility in fitness in the environment. Future work focused on determining those targets in the environment for these virulence factors would be highly informative and should yield valuable information about factors which contribute to the maintenance of highly pathogenic strains in the natural environment.

# Bibliography

Abbott, S. L., C. Powers, C. A. Kaysner, Y. Takeda, M. Ishibashi, S. W. Joseph, and J. M. Janda. (1989) Emergence of a restricted bioserovar of *Vibrio parahaemolyticus* as the predominant cause of *Vibrio*-associated gastroenteritis on the West Coast of the United States and Mexico. J Clin Microbiol 27:2891-2893.

Allen, MJ., and Geldreich, EE. (1975) Bacteriological Criteria for Ground-Water Quality. Ground Water. 13(1)45–52.

Nath, G., Singh, YK., Maurya, P., Gulati, AK., Srivastava, RC., Tripathi, SK. (2010) Does Salmonella Typhi primarily reside in the liver of chronic typhoid carriers? J Infect Dev Ctries. 3;4(4):259-61.

Saphra, I., and Winter, JW. (1957) Clinical Manifestations of Salmonellosis in Man — An Evaluation of 7779 Human Infections Identified at the New York Salmonella Center. N Engl J Med, 256:1128-1134.

Mezrioui, N., Baleux, B., Troussellier, M. (1995) A microcosm study of the survival of Escherichia coli and Salmonella typhimurium in brackish water. Water Research. 29(2)459–465.

Levin, BR., and Bull, JJ. (1994) Short-sighted evolution and the virulence of pathogenic microorganisms. Trends in Microbiology. 2(3)76–81.

Haley, B., Chen, A., Grim, CJ., Clark, P., Diaz, CM., Taviani, E., Hasan, NA., Sancomb, E., Elnemr, WM., Islam, M., Huq, A., Colwell, R., Benediktsdottir, E., (2012) Vibrio cholerae in a historically cholera-free country. DOI: 10.1111/j.1758-2229.2012.00332.x.

Woods, RJ., Barrick, JE., Cooper, TF., Shrestha, U., Kauth, M., and Lenski, R. (2011) Second-Order Selection for Evolvability in a Large Escherichia coli Population. Science. 331(6023)1433-1436.

Pukatzki, S., Ma, A., Sturtevant, D., Krastins, B., Sarracino, D., Nelson, W., Heidelberg, J., and Mekalanos, J. (2006) Identification of a conserved bacterial protein secretion system in Vibrio cholerae using the Dictyostelium host model system. PNAS.103(5) 1528-1533.

Reguera, G., and Kolter, R., (2005) Virulence and the Environment: a Novel Role for Vibrio cholerae Toxin-Coregulated Pili in Biofilm Formation on Chitin. J. Bacteriol. 187(10)3551-3555.

Vaitkevicius, K., Lindmark, B., Ou , G., Song, T., Toma, C., Iwanaga, M., Zhu, J., Andersson, A., Hammarström, ML., Tuck, S., and Wai, SN., (2006). A Vibrio

cholerae protease needed for killing of Caenorhabditis elegans has a role in protection from natural predator grazing. PNAS. 103(24) 9280-9285.

Adekambi T, Shinnick TM, Raoult D, Drancourt M: Complete *rpoB* gene sequencing as a suitable supplement to DNA-DNA hybridization for bacterial species and genus delineation. International Journal of Systematic and Evolutionary Microbiology 2008, 58(8):1807-1814.

Alam, M., W. B. Chowdhury, N. A. Bhuiyan, A. Islam, N. A. Hasan, G. B. Nair, H.,Watanabe, A. Huq, A. K. Siddique, R. B. Sack, M. Z. Akhter, C. J. Grim, K. M. Kam, C. K. Luey, H. Endtz, and R. R. Colwell. (2009). Serogroup, virulence, and genetic traits of *Vibrio parahaemolyticus* in the estuarine ecosystem of Bangladesh. Appl. Environ. Microbiol.

Almagro-Moreno, S., and Boyd., E.F. (2009) Sialic Acid Catabolism Confers a Competitive Advantage to Pathogenic *Vibrio cholerae* in the Mouse Intestine. Infect. Immun. 77(9): 3807-3816.

Ansaruzzaman, M., M. Lucas, J. L. Deen, N. A. Bhuiyan, X. Y. Wang, A. Safa, M. Sultana, A. Chowdhury, G. B. Nair, D. A. Sack, L. von Seidlein, M. K. Puri, M. Ali, C. L. Chaignat, J. D. Clemens, and A. Barreto. 2005. Pandemic serovars (O3:K6 and O4:K68) of *Vibrio parahaemolyticus* associated with diarrhea in Mozambique: spread of the pandemic into the African continent. J Clin Microbiol 43:2559-62.

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M: The RAST Server: rapid annotations using subsystems technology. BMC Genomics 2008, 9(1):75.

Baker-Austin, C., Stockley, L., Rangdale, R., Martinez-Urtaza, J. (2010) Environmental occurrence and clinical impact of *Vibrio vulnificus* and *Vibrio parahaemolyticus*: a European perspective. Environmental Microbiology Reports.2(1)7–18.

Barnhart BJ, Herriott RM: Penetration of deoxyribonucleic acid into *Haemophilus influenzae*. Biochimica et Biophysica Acta 1963, 76:25-39.

Bej, A. K., D. P. Patterson, C. W. Brasher, M. C. Vickery, D. D. Jones, and C. A. Kaysner (1999) Detection of total and hemolysin-producing *Vibrio parahaemolyticus* in shellfish using multiplex PCR amplification of *tl*, *tdh* and *trh*. J Microbiol Methods 36:215-225.

Beyhan, S., Tischler, A., Camilli, A., and Yildiz, F. (2006) Differences in Gene Expression between the Classical and El Tor Biotypes of *Vibrio cholerae* O1. Infect. Immun. 74(6)3633-3642.

Bhattacharjee, R. N., K. S. Park, Y. Kumagai, K. Okada, M. Yamamoto, S. Uematsu, K. Matsui, H. Kumar, T. Kawai, T. Iida, T. Honda, O. Takeuchi, and S. Akira (2006) VP1686, a *Vibrio* type III secretion protein, induces toll-like receptor-independent apoptosis in macrophage through NF-kappaB inhibition. J Biol Chem 281:36897-36904.

Bhoopong, P., Palittapongarnpim, P., Pomwised, R., Kiatkittipong, A., Kamruzzaman, M., Nakaguchi, Y., Nishibuchi, M., Ishibashi, M. and Vuddhakul, V. (2007) Variability of Properties of *Vibrio parahaemolyticus* Strains Isolated from Individual Patients. J. Clin. Microbiol. May 2007 vol. 45 no. 5 1544-1550.

Binsztein, N., Costagliola, M., Pichel, M., Jurquiza, V., Ramirez, F., Akselman, R. et al. (2004) Viable but nonculturable *Vibrio cholerae* O1 in the aquatic environment of Argentina. Appl Environ Microbiol 70: 7481-7486.

Boucher Y, et al. 2011. Local mobile gene pools rapidly cross species boundaries to create endemicity within global *Vibrio cholerae* populations. mBio 2(2):e00335-10. doi:10.1128/mBio.00335-10.

Boyd EF, Almagro-Moreno S, Parent MA: Genomic islands are dynamic, ancient integrative elements in bacterial evolution. Trends in Microbiology 2009, 17(2):47-53.

Boyd, E., Moyer, K., Shi, L., Waldor, M. (2000) Infectious CTXΦ and the *Vibrio* Pathogenicity Island Prophage in *Vibrio mimicus*: Evidence for Recent Horizontal Transfer between *V. mimicus* and *V. cholerae*. Infect. Immun. 68(3)1507-1513.

Brabazon, E.D., O'Farrell, A., Murray, C.A., Carton, M.W., and Finnegan, P. (2008) Under-reporting of notifiable infectious disease hospitalizations in a health board region in Ireland: room for improvement? Epidemiol Infect 136: 241-247.

Buswell, C. M., Herlihy, Y. M., Lawrence, L. M., McGuiggan, J. T., Marsh, P. D., Keevil, C. W. & Leach, S. A. (1998). Extended survival and persistence of *Campylobacter* spp. in water and aquatic biofilms and their detection by immunofluorescent-antibody and -rRNA staining. Applied and Environmental Microbiology, 64, 733-741.

Byun, R., Elbourne, L.D.H., Lan, R., and Reeves, P.R. (1999) Evolutionary Relationships of Pathogenic Clones of *Vibrio cholerae* by Sequence Analysis of Four Housekeeping Genes. Infect. Immun. 67:1116-1124.

Cabanillas-Beltrán, H.,LLausás-Magaña, E., Romero, R., Espinoza, A., García-Gasca, A., Nishibuchi, M., Ishibashi, M.,Gomez-Gil, B.(2006) Outbreak of gastroenteritis caused by the pandemic *Vibrio parahaemolyticus* O3 : K6 in Mexico. FEMS Microbiology Letters. 265(1) 76-80.

Chaiyanan, S., Chaiyanan, S., Grim, C., Maugel, T., Huq, A., and Colwell, RR. (2007) Ultrastructure of coccoid viable but non-culturable *Vibrio cholerae*. Environmental Microbiology. 9(2): 393–402.

Chao, G., Jiao, X., Zhou, X., Wang, F., Yang, Z., Huang, J., Pan, Z., Zhou, L., Qian, X. (2010) Distribution of genes encoding four pathogenicity islands (VPaIs), T6SS, biofilm, and type I pilus in food and clinical strains of *Vibrio parahaemolyticus* in China. Foodborne Pathog Dis. 7(6):649-58.

Chatterjee, B. D., and T. Sen (1974) *Vibrio parahaemolyticus* serotypes in Calcutta, India. Bull World Health Organ 50:559-561.

Chatterjee, S., Ghosh, K., Raychoudhuri, A., Chowdhury, G., Bhattacharya, M.K., Mukhopadhyay, A.K., Ramamurthy, T., Bhattacharya, S.K., Klose, K.E., Nandy R.K. (2009) Incidence, virulence factors, and clonality among clinical strains of non-O1, non-O139 *Vibrio cholerae* isolates from hospitalized diarrheal patients in Kolkata, India.. J Clin Microbiol. 47:1087-1095.

Chokesajjawatee, N., Y. G. Zo, and R. R. Colwell (2008) Determination of clonality and relatedness of *Vibrio cholerae* isolates by genomic fingerprinting, using long-range repetitive element sequence-based PCR. Appl Environ Microbiol 74:5392-5401.

Choopun N: The population structure of *Vibrio cholerae* in Chesapeake Bay. PhD Thesis. University of Maryland, College Park, Marine Estuarine and Environmental Science; 2004.

Choopun, N., Louis, V., Huq, A., and Colwell, RR. (2002) Simple Procedure for Rapid Identification of *Vibrio cholerae* from the Aquatic Environment. Appl. Environ. Microbiol. 68(2):995-998

Chowdhury, N. R., S. Chakraborty, T. Ramamurthy, M. Nishibuchi, S. Yamasaki, Y. Takeda, and G. B. Nair (2000) Molecular evidence of clonal *Vibrio parahaemolyticus* pandemic strains. Emerg Infect Dis 6:631-636.

Chun, J., Grim, C.J., Hasan, N.A., Lee, J.H., Choi, S.Y., Haley, B.J, et al. (2009) Comparative Genomics Reveals Mechanism for Short-term and Long-term Clonal Transitions in Pandemic *Vibrio cholerae*. Proc Natl Acad Sci USA. 106:15442-15447.

Chun, J., Huq, A., and Colwell, RR. (1999) Identification of *Vibrio cholerae* based on genes coding for 16S-23S rRNA internal transcriber spacers. Appl. Environ. Microbiol. 65:2202-2208.

Cinar HN, Kothary M, Datta AR, Tall BD, Sprando R, et al. (2010) *Vibrio cholerae* Hemolysin Is Required for Lethality, Developmental Delay, and Intestinal

Vacuolation in *Caenorhabditis elegans*. PLoS ONE 5(7): e11558. doi:10.1371/journal.pone.0011558.

Clark, C., Kravetz, A., Dendy, C., Wang, G., Tyler, K., and Johnson, W. (1998). Investigation of the 1994–5 Ukrainian *Vibrio cholerae* epidemic using molecular methods. Epidemiology and Infection. 121:15-29.

Cliff, A.D., Haggett, P., and Smallman-Raynor. M. (2009) The changing shape of island epidemics: historical trends in icelandic infectious disease waves, 1902-1988. Journal of Historical Geography. 35(3):545-567.

Colwell, R.R. (1996) Global climate and infectious disease: the cholera paradigm. Science 274: 2025-2031.

Colwell, R.R. (2006) Global Microbial Ecology of *Vibrio cholerae*. In Oceans and Health: Pathogens in the Marine Environment. Belkin, S., and Colwell, R.R. (eds). New York: Springer Science+Business Media, Inc., pp. 297-305.

Constantin de Magny, G., Long, W., Brown, C.W., Hood, R.R., Huq, A., Murtuggude, R., and Colwell, R.R. (2009) Predicting the Distribution of *Vibrio* spp. in the Chesapeake Bay: A *Vibrio cholerae* Case Study. EcoHealth 6: 378-389.

Constantin de Magny, G., Murtugudde, R., Sapiano, M., Nizam, A., Brown, C., Busalacchi, A., Yunus, M., Nair, G., Gil, A., Lanata, C., Calkins, J., Manna, B., Rajendran, K., Bhattacharya, M., Huq, A., Sack, R., and Colwell, R. (2008) Environmental signatures associated with cholera epidemics. Proc Natl Acad Sci. USA. 105(46)17676-17681.

Cooper, KLF., Luey, CKY., Bird, M., Terajima, J., Nair, GB., Kam, KM., et al. (2006) Development and validation of a PulseNet standardized pulsed-field gel electrophoresis protocol for subtyping of *Vibrio cholerae*. Foodborne Pathog. Dis. 3(1):51-58.

Dalsgaard, A., Serichantalergs, O., Forslund, A., Lin, W., Mekalanos, J., Mintz, E., Shimada, T., and Wells, J. (2001) Clinical and Environmental Isolates of *Vibrio cholerae* Serogroup O141 Carry the CTX Phage and the Genes Encoding the Toxin-Coregulated Pili. J. Clin. Microbiol. 39(11)4086-4092.

Daniels, N. A., L. MacKinnon, R. Bishop, S. Altekruse, B. Ray, R. M. Hammond, S. Thompson, S. Wilson, N. H. Bean, P. M. Griffin, and L. Slutsker (2000) *Vibrio parahaemolyticus* infections in the United States, 1973-1998. J Infect Dis 181:1661-1666.

Davis BM, Waldor MK: Filamentous phages linked to virulence of *Vibrio cholerae*. Curr Opin Microbiol 2003, 6(1):35-42.

Davis BR, Fanning GR, Madden JM, Steigerwalt AG, Bradford HB, Jr., Smith HL, Jr., Brenner DJ: Characterization of biochemically atypical *Vibrio cholerae* strains and designation of a new pathogenic species, *Vibrio mimicus*. J Clin Microbiol 1981, 14(6):631-639.

Davis, B. M., Kimsey, H. H., Chang, W., Waldor, M. K. (1999) The *Vibrio cholerae* O139 Calcutta bacteriophage CTXΦ is infectious and encodes a novel repressor. J Bacteriol. 181:6779-6787.

DePaola, A., C. A. Kaysner, J. Bowers, and D. W. Cook (2000) Environmental investigations of *Vibrio parahaemolyticus* in oysters after outbreaks in Washington, Texas, and New York (1997 and 1998). Appl Environ Microbiol 66:4649-4654.

Directorate of Health (Iceland). http://www.landlaeknir.is/Pages/876 (accessed on May 2, 2011).

Djhanguiri, F., Reimer, G.M., Holub, R., Mines, C.S.o., Aliev, C.S., Zolotovitska, T., and Sciences, A.A.o. (1997) Radioactive Pollution at the Oil Fields of the Apsheron Peninsula, Caspian Sea, Azerbajan. In SPE/EPA Exploration and Production Environmental Conference, 3-5 March 1997. Dallas, Texas: Society of Petroleum Engineers, Inc.

Dowell, S.F. and Bresee, JS. (2008) Pandemic lessons from Iceland. Proc. Natl. Acad. Sci. 105(4):1109-1110.

Dumont, H. (1995) Ecocide in the Caspian Sea. Nature 377: 673-674.

Dumontier, S., and Berche, P. (1998) *Vibrio cholerae* O22 might be a putative source of exogenous DNA resulting in the emergence of the new strain of Vibrio cholerae O139. FEMS Microbiology Letters. 164(1)91–98.

Dziejman, M., Balon, E., Boyd, D., Fraser, C. M., Heidelberg, J. F. & Mekalanos, J. J. (2002). Comparative genomic analysis of *Vibrio cholerae*: genes that correlate with cholera endemic and pandemic disease. Proc Natl Acad Sci USA. 99:1556-1561.

Ewing, A. (2010) Water Quality and Public Health Monitoring of Surface Waters in the Kura-Araks River Basin of Armenia, Azerbaijan, and Georgia. In: University of New Mexico.

Faruque, S. M. & Mekalanos, J. J. (2003). Pathogenicity islands and phages in *Vibrio cholerae* evolution. Trends Microbiol 11:505-510.

Fayer, R., Graczyk, T. K., Lewis, E. J., Trout, J. M. & Farley, C. A. (1998). Survival of infectious *Cryptosporidium parvum* oocysts in seawater and eastern oysters

(*Crassostrea virginica*) in the Chesapeake Bay. Applied and Environmental Microbiology, 64, 1070-1074.

Figueiredo, S. C., Neves-Borges, A. and Coelho, A. (2005). The neuraminidase gene is present in the non-toxigenic *Vibrio cholerae* Amazonia strain: a different allele in comparison to the pandemic strains. Mem Inst Oswaldo Cruz. 100:563-569.

Figueras, M. J., Inza, I., Polo, F. L., Feliu, M. T. & Guarro, J. (1996). A Fast method for the confirmation of fecal streptococci from M-enterococcus medium. Applied and Environmental Microbiology, 62, 2177-2178.

Finkelstein, R.A., Boesman-Finkelstein, M., Chang, Y., and Häse, C.C. (1992) *Vibrio cholerae* hemagglutinin/protease, colonial variation, virulence, and detachment. Infect. Immun. 60(2): 472-478.

Galen, J., Ketley, J., Fasano, A., Richardson, S., Wasserman, S., and Kaper, J. (1992) Role of *Vibrio cholerae* neuraminidase in the function of cholera toxin. Infect. Immun. 60(2)406-415.

García, K., Torres, R., Uribe, P., Hernández, C., Rioseco, M., Romero, J., and Espejo, R. (2009) Dynamics of Clinical and Environmental *Vibrio parahaemolyticus* Strains during Seafood-Related Summer Diarrhea Outbreaks in Southern Chile. Appl. Environ. Microbiol. 75(23)7482-7487.

Goryshin, I.Y., and Reznikoff, W.S. (1998) Tn5 *in vitro* transposition. J Biol Chem. 273:7367-7374.

Grim, C.J., Tediashvili, M., Whitehouse, C.A., Jaiani, E., Kokashvili, T., Janelidze, N. et al. (2010) Detection of toxigenic *Vibrio cholerae* O1 from freshwater environments in the former Soviet Republic of Georgia. Environ Microbiol Reports 2: 2-6.

Grim, CJ., Choi, J., Chun, J., Jeon, YS., Taviani, E., Hasan, NA., et al. (2010) Occurrence of the *Vibrio cholerae* Seventh Pandemic VSP-I Island and a New Variant. OMICS: A Journal of Integrative Biology. 14(1):1-7.

Guang-Qing, T., L. Tetsuya, Y. Koichiro, and H. Takeshi (1995) Ca2+ independent cytotoxicity of *Vibrio parahaemolyticus* thermostable direct hemolysin (TDH) on Intestine 407, a cell line derived from human embryonic intestine. FEMS Microbiology Letters 134:233-238.

Guðfinnsson, E.K. (2007) Address by the Icelandic Minister of Fisheries. Opening the Conference Iceland Mussel 2007. http://eng.sjavarutvegsraduneyti.is/minister/EKG_English/nr/1341 (accessed on February 12, 2011).

Guentzel, M.N., and Berry, L.J. (1975) Motility as a virulence factor for *Vibrio cholerae*. Infect. Immun. 11(5): 890-897.

Gunnarsdóttir, B.E. (2008) Assessment of iodine and mercury status of Icelandic adolescent girls, in Department of Food Science. University of Iceland: Reykjavik.

Gurbanov, S.H., Akhmadov, R., Shamkhalova, G., Akhmadova, S., Colwell, R.R., and Huq, A. (2012) Occurrence of *Vibrio cholerae* in Municipal and Natural Waters and Incidence of Cholera in Azerbaijan. EcoHealth. DOI: 10.1007/s10393-012-0756-8.

Haley BJ, Grim CJ, Hasan NA, Taviani E, Chun J, Brettin TS, Bruce DC, Challacombe JF, Detter JC, Han CS, et al.: The pre-seventh pandemic *Vibrio cholerae* BX 330286 El Tor genome: evidence for the environment as a genome reservoir. Environmental Microbiology Reports 2010, 2(1):208-216.

Haley, B. J., Cole, D. J. & Lipp, E. K. (2009). Distribution, diversity, and seasonality of waterborne salmonellae in a rural watershed. Applied and Environmental Microbiology, 75, 1248-1255.

Haley, B., Grim, C., Hasan, N., Chun, J., Huq, A., and Colwell, R. (2010). Comparative genomic analysis reveals evidence of two novel *Vibrio* species closely related to *V. cholerae*. BMC Microbiology. 10:154.

Halpern, M., Gancz, H., Broza, M., and Kashi, Y. (2003) *Vibrio cholerae* hemagglutinin/protease degrades chironomid egg masses. Appl. Environ. Microbiol. 69(7):4200-4204.

Han, C.S., and Chain, P. (2006) Finishing repeat regions automatically with Dupfinisher, p. 141-146. In H.R. Arabnia and H. Valafar (ed.), Proceedings of the 2006 International Conference on Bioinformatics and Computational Biology. CSREA Press, Las Vegas, NV.

Han, G.K., Khie, T.S. (1963) A new method for the differentiation of *Vibrio comma* and *Vibrio El Tor*. Am J Hyg. 77:184-186.

Harth, E., Matsuda, L., Hernández, C., Rioseco, M., Romero, J., González-Escalona, N., Martínez-Urtaza, J., and Espejo, R. (2009) Epidemiology of *Vibrio parahaemolyticus* Outbreaks, Southern Chile. Emerg Infect Dis. 15(2): 163–168.

Health Protection Agency (2006) Enumeration of enterococci by membrane filtration, National Standard Method. W 3 Issue 3.

Hjaltelin, J. (1871) Small-Pox Imported into Iceland by French Fishing Vessels, Stamped out by Quarantine and Sulphurous Fumigations. British Medical Journal. 2(566):519.

Hlady, W., and Klontz, K. (1996) The Epidemiology of *Vibrio* Infections in Florida, 1981–1993. J Infect Dis. 173(5):1176-1183.

Hochhut, B., Lotfi, Y., Mazel, D., Faruque, SM., Woodgate, R., and Waldor, MK. (2001) Molecular analysis of antibiotic resistance gene clusters in *Vibrio cholerae* O139 and O1 SXT constins. Antimicrob. Agents Chemother. 45(11):2991-3000.

Honda, T. (1993) The pathogenicity of *Vibrio parahaemolyticus* and the role of the thermostable direct haemolysin and related haemolysins. Reviews in Medical Microbiology 4:106-113.

Hoshino, K., Yamasaki, S., Mukhopadhyay, AK., Chakraborty, S., Basu, A., Bhattacharya, SK., et al. (1998) Development and evaluation of a multiplex PCR assay for rapid detection of toxigenic *Vibrio cholerae* O1 and O139. FEMS Immunol Med Microbiol. 20(3):201-207.

Huq A, Small E, West P, Huq M, Rahman R, Colwell R: Ecological relationship between *Vibrio cholerae* and planktonic copepods. Appl Environ Microbiol 1983, 45:275-283.

Huq, A., Colwell, R.R., Chowdhury, M.A., Xu, B., Moniruzzaman, S.M., Islam, M.S. et al. (1995) Co-existence of *Vibrio cholerae* O1 and O139 Bengal in plankton in Bangladesh. Lancet 345: 1249.

Huq, A., Colwell, R.R., Rahman, R., Ali, A., Chowdhury, M.A., Parveen, S. et al. (1990) Detection of *Vibrio cholerae* O1 in the aquatic environment by fluorescent-monoclonal antibody and culture methods. Appl Environ Microbiol 56: 2370-2373.

Huq, A., Grim, C., and Colwell, R. (2006) Detection, Isolation, and Identification of *Vibrio cholerae* from the Environment. In Current Protocols in Microbiology. Taylor, R. (ed). New York: John Wiley & Sons, pp. 6A.5.1-6A.5.38.

Huq, A., Rivera, I., and Colwell, R. (2000) Epidemiological Significance of Viable but Nonculturable Microorganisms. In Nonculturable Microorganisms in the Environment. Colwell, R., and Grimes, D. (eds). Washington, D.C.: ASM Press, pp. 301-323.

Huq, A., Sack, R., Nizam, S., Longini, I., Nair, G., Ali, S., Morris Jr., J., Huda Khan, M., Siddique, A., Yunus, M., Albert, M., Sack, D., and Colwell, R. (2005) Critical Factors Influencing the Occurrence of *Vibrio cholerae* in the Environment of Bangladesh. Appl. Environ. Microbiol. 71(8)4645-4654.

Huq, A., Small, E.B., West, P.A., Huq, M.I., Rahman, R., and Colwell, R.R. (1983) Ecological relationships between *Vibrio cholerae* and planktonic crustacean copepods. Appl Environ Microbiol 45: 275-283.

Huseynov, N., and Malikov, B. (2009) Regularity of distribution of precipitation at the airdomes of Azerbaijan Republic. Adv Geosci 20: 9-12.

ICDDRB, C.W.G. (1993) Large epidemic of cholera-like disease in Bangladesh caused by *Vibrio cholerae* O139 synonym Bengal. Lancet 342: 387-390.

Ichinose, Y., Ehara, M., Honda, T., and Miwatani, T. (1994) The effect on enterotoxicity of protease purified from *Vibrio cholerae* O1. FEMS Microbiol. Lett. 115(2-3):265-271.

Ingólfsson, A. (1995) Floating clumps of seaweed around Iceland: natural microcosms and a means of dispersal for shore fauna. Marine Biology. 122:13-21.

Islam, M.S., Drasar, BS., and Sack. RB. (1993) The aquatic environment as a reservoir of *Vibrio cholerae*: a review. J. Diarrhoeal Dis. Res. 11(4):197-206.

Islamzade, A. (1994) Sumgayit: Soviet's Pride, Azerbaijan's Hell. Azerbaijan International: 26-27, 30.

Jackson, C. R., Fedorka-Cray, P. J., Barrett, J. B., Hiott, L. M. & Woodley, T. A. (2007). Prevalence of streptogramin resistance in enterococci from animals: identification of *vatD* from animal sources in the USA. International Journal of Antimicrobial Agents, 30, 60-66.

Jacobs, J.M., Rhodes, M., Brown, C.W., Hood, R.R., Leigh, A., Long, W., and Wood, R. (2010) Predicting the Distribution of *Vibrio vulnificus* in Chesapeake Bay. NOAA Technical Memorandum NOS NCCOS 112. In. Oxford, Maryland: NOAA National Centers for Coastal Ocean Science, Center for Coastal Environmental Health and Biomolecular Research, Cooperative Oxford Laboratory.

Jermyn, W. S. & Boyd, E. F. (2002). Characterization of a novel *Vibrio* pathogenicity island (VPI-2) encoding neuraminidase (*nanH*) among toxigenic *Vibrio cholerae* isolates. Microbiology 148:3681-3693.

Jermyn, W. S. and Boyd, E. F. (2005). Molecular evolution of *Vibrio* pathogenicity island-2 (VPI-2): mosaic structure among *Vibrio cholerae* and *Vibrio mimicus* natural isolates. Microbiology 151:311-322.

Jonson, G., Holmgren, J., Svennerholm, A.M. (1991) Epitope differences in toxin-coregulated pili produced by classical and El Tor *Vibrio cholerae* O1. Microb Pathog. 3:179-188.

Kadokura, K., Rokutani, A., Yamamoto, M., Ikegami, T., Sugita, H., Itoi, S., Hakamata, W., Oku, T., and Nishio, T. (2007) Purification and characterization of *Vibrio parahaemolyticus* extracellular chitinase and chitin oligosaccharide deacetylase involved in the production of heterodisaccharide from chitin. Applied Microbiology and Biotechnology. 75(2)357-365.

Kaiser, TS., Neumann, D., Heckel, DG., Berendonk, TU. (2010) Strong genetic differentiation and postglacial origin of populations in the marine midge Clunio marinus (Chironomidae, Diptera). Molecular Ecology. 19: 2845–2857.

Kamruzzaman, M., P. Bhoopong, V. Vuddhakul, and M. Nishibuchi (2008) Detection of a functional insertion sequence responsible for deletion of the thermostable direct hemolysin gene (*tdh*) in *Vibrio parahaemolyticus*. Gene 421:67-73.

Kaneko, T., and Colwell, R. (1974) Distribution of *Vibrio parahaemolyticus* and Related Organisms in the Atlantic Ocean off South Carolina and Georgia. Appl. Environ. Microbiol. 28(6) 1009-1017.

Karlin S, Mrazek J, Campbell AM: Compositional biases of bacterial genomes and evolutionary implications. Journal of Bacteriology 1997, 179(12):3899-3913.

Karlsson, G. (1996) Plague without rats: the case of fifteenth-century Iceland. Journal of Medieval History. 22(3):263-284.

Kaspar C., and Tamplin, M. (1993) Effects of temperature and salinity on the survival of Vibrio vulnificus in seawater and shellfish. Appl Environ Microbiol. 59(8):2425–2429.

Kenyon, JE., Piexoto, DR., Austin, B., Gillies, DC. (1984) Seasonal variations of *Vibrio cholerae* (non-O1) isolated from California coastal waters. Appl. Environ. Microbiol. (47)6: 1243-1245.

Kim, Y. B., J. Okuda, C. Matsumoto, N. Takahashi, S. Hashimoto, and M. Nishibuchi (1999) Identification of *Vibrio parahaemolyticus* strains at the species level by PCR targeted to the *toxR* gene. J Clin Microbiol 37:1173-1177.

Kimura M: A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of molecular evolution 1980, 16(2):111-120.

Kimura, M. (1980) A simple method for estimating evolutionary rate of base substitutions through 30 comparative studies of nucleotide sequences. J Mol Evol 16:111-120.

Ko, W.C., Chuang, Y.C., Huang, G.C., and Hsu, S.Y. (1998) Infections due to non-O1 *Vibrio cholerae* in southern Taiwan: predominance in cirrhotic patients. Clin Infect Dis. 27:774-780.

Kodama, T., M. Rokuda, K. S. Park, V. V. Cantarelli, S. Matsuda, T. Iida, and T. Honda (2007) Identification and characterization of VopT, a novel ADP-ribosyltransferase effector protein secreted via the *Vibrio parahaemolyticus* type III secretion system 2. Cell Microbiol 9:2598-609.

Konstantinidis KT, Ramette A, Tiedje JM: The bacterial species definition in the genomic era. Philosophical Transactions B 2006, 361(1475):1929.

Konstantinidis KT, Tiedje JM: Genomic insights that advance the species definition for prokaryotes. Proceedings of the National Academy of Sciences 2005, 102(7):2567-2572.

Konstantinidis KT, Tiedje JM: Prokaryotic taxonomy and phylogeny in the genomic era: advancements and challenges ahead. Current opinion in microbiology 2007, 10(5):504-509.

Konstantinidis KT, Tiedje JM: Towards a Genome-Based Taxonomy for Prokaryotes. J Bacteriol 2005, 187(18):6258-6264.

Konstantinidis, K.T., and Tiedje, J.M. (2005) Towards a genome-based taxonomy for prokaryotes. J Bacteriol. 187:6258-6264.

Krukonis, ES., and DiRita, VJ. (2003) From motility to virulence: sensing and responding to environmental signals in *Vibrio cholerae*. Current Opinion in Microbiology. 6:186–190.

Kumar, S., Nei, M., Dudley, J., and Tamura, K. (2008) MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform. 9:299-306.

Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., and Salzberg, S.L. (2004) Versatile and open software for comparing large genomes. Genome Biol. 5: R12.

Kushmaro A, Banin E, Loya Y, Stackebrandt E, Rosenberg E: *Vibrio shiloi* sp. nov., the causative agent of bleaching of the coral *Oculina patagonica*. Int J Syst Evol Microbiol 2001, 51(Pt 4):1383-1388.

Lång, H., Jonson, G., Holmgren, J., Palva, E.T. (1994) The maltose regulon of *Vibrio cholerae* affects production and secretion of virulence factors. Infect Immun. 62:4781-4788.

Langille MGI, Brinkman FSL: IslandViewer: an integrated interface for computational identification and visualization of genomic islands. Bioinformatics 2009, 25(5):664-665.

Langworthy, D.E., Stapleton, R.D., Sayler, G.S., and Findlay, R.H. (1998) Genotypic and Phenotypic Responses of a Riverine Microbial Community to Polycyclic Aromatic Hydrocarbon Contamination. Appl Environ Microbiol 64: 3422-3428.

Laohaprertthisan, V., Chowdhury, A., Kongmuang, U., Kalnauwakul, S., Ishibashi, M., Matsumoto, C., and Nishibuchi, M. (2003) Prevalence and Serodiversity of the Pandemic Clone among the Clinical Strains of *Vibrio parahaemolyticus* Isolated in Southern Thailand. Epidemiology and Infection. 130(3) 395-406.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R et al. (2007) Clustal W and Clustal X version 2.0. Bioinformatics. 23(21):2947-2948.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H. et al. (2007) Clustal W and Clustal X version 2.0. Bioinformatics. 23: 2947-2948.

Leclerc, H., Mossel, D. A., Edberg, S. C. & Struijk, C. B. (2001). Advances in the bacteriology of the coliform group: their suitability as markers of microbial water safety. Annual Reviews in Microbiology, 55, 201-234.

Lemke, M.J., Brown, B.J., and Leff, L.G. (1997) The Response of Three Bacterial Populations to Pollution in a Stream. Microbial Ecology 34: 224-231.

Lerouge, I., T. Laeremans, C. Verreth, J. Vanderleyden, C. Van Soom, A. Tobin, and R. W. Carlson (2001) Identification of an ATP-binding cassette transporter for export of the O-antigen across the inner membrane in Rhizobium etli based on the genetic, functional, and structural analysis of an lps mutant deficient in O-antigen. J Biol Chem 276:17190-198.

Libinzon, A. E., G. D. Lebedev, I. V. Pavlova, A. F. Nagornaia, and N. V. Krasnova (1975) Isolation of parahemolytic vibrions from persons with acute gastrointestinal diseases. Zh Mikrobiol Epidemiol Immunobiol 0:25-29.

Libinzon, A. E., I. V. Domaradskii, Z. I. Us, A. I. Demina, and A. F. Nagornaia (1974) Parahemolytic vibrios and related halophilic microorganisms of the Black Sea. Zh Mikrobiol Epidemiol Immunobiol 00:80-84.

Libinzon, A. E., Luneva, E.I., Kashirova, A.K., Velichko, S.A., Sakhno, E.I (1980) Halophilic vibrios in human excreta and seawater. Gig Sanit Dec:73-74.

Libinzon, A., Brudnyĭ, R., Nagornaia, A., Demina, A., Krasnova, N. (1981) Halophilic Black Sea vibrios and their role in human pathology. Zh Mikrobiol Epidemiol Immunobiol. (2):97-101.

Lipp, E. K., Schmidt, N., Luther, M. E. & Rose, J. B. (2001). Determining the effects of El Nino-Southern Oscillation events on coastal water quality. Estuaries, 24, 491-497.

Lipp, E., Huq, A., and Colwell, R. (2002) Effects of Global Climate on Infectious Disease: the Cholera Model. Clin. Microbiol. Rev. 15(4)757-770.

Lobitz, B., Beck, L., Huq, A., Wood, B., Fuchs, G., Faruque, A., and Colwell, R. (2000) Climate and infectious disease: use of remote sensing for detection of *Vibrio cholerae* by indirect measurement. Proc. Nat. Acad. Sci. 97: 1438-1443.

Louis, V., Russek-Cohen, E., Choopun, N., Rivera, I., Gangle, B., Jiang, S. et al. (2003) Predictability of *Vibrio cholerae* in Chesapeake Bay. Appl Environ Microbiol 69: 2773-2785.

Lukinmaa, S., Mattila, K., Lehtinenc, V., Hakkinen, M., Koskela, M., Siitonen, A. (2006) Territorial waters of the Baltic Sea as a source of infections caused by *Vibrio*

*cholerae* non-O1, non-O139: report of 3 hospitalized cases. Diagn Micr Infec Dis. 54:1-6.

Mac Kenzie WR, H. N., Proctor ME, Gradus MS, Blair KA, Peterson DE, Kazmierczak JJ, et al. (1994). A massive outbreak in Milwaukee of *Cryptosporidium* infection transmitted through the public water supply. N Engl J Med 331, 1035.

MacDougall, L., Majowicz, S., Dore, K., Flint, J., Thomas, K., Kovacs, S., and Sockett, P. (2008) Under-reporting of infectious gastrointestinal illness in British Columbia, Canada: who is counted in provincial communicable disease statistics? Epidemiol Infect 136: 248-256.

Makino, K., K. Oshima, K. Kurokawa, K. Yokoyama, T. Uda, K. Tagomori, Y. Iijima, M. Najima, M. Nakano, A. Yamashita, Y. Kubota, S. Kimura, T. Yasunaga, T. Honda, H. Shinagawa, M. Hattori, and T. Iida (2003) Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism distinct from that of *V. cholerae*. Lancet 361:743-749.

Manning, S., Motiwala, A., Springman, A., Qi, W., Lacher, D., Ouellette, L., Mladonicky, J., Somsel, P., Rudrik, J., Dietrich, S., Zhang, W., Swaminathan, B., Alland, D., Whittam, T.(2008) Variation in virulence among clades of *Escherichia coli* O157:H7 associated with disease outbreaks. Proc Natl Acad Sci USA. (12)4868-4873.

Martinelli-Filho, J.E., Lopes, R.M., Rivera, I.N.G., and Colwell, R.R. (2011) *Vibrio cholerae* O1 detection in estuarine and coastal zooplankton. J Plankton Res 33: 51-62.

Matlawska-Wasowska K, Finn R, Mustel A, O'Byrne CP, Baird AW, Coffey ET, Boyd A (2010) The Vibrio parahaemolyticus Type III Secretion Systems manipulate host cell MAPK for critical steps in pathogenesis. BMC Microbiol. 30;10:329.

Matson, J.S., Withey, J.H., DiRita, V.J. (2007) Regulatory Networks Controlling *Vibrio cholerae* Virulence Gene Expression. Infect Immun. 75:5542-5549.

Matsumoto, C., J. Okuda, M. Ishibashi, M. Iwanaga, P. Garg, T. Rammamurthy, H. C. Wong, A. Depaola, Y. B. Kim, M. J. Albert, and M. Nishibuchi (2000) Pandemic spread of an O3:K6 clone of *Vibrio parahaemolyticus* and emergence of related strains evidenced by arbitrarily primed PCR and toxRS sequence analyses. J Clin Microbiol 38:578-585.

Matz C, Nouri B, McCarter L, Martinez-Urtaza J (2011) Acquired Type III Secretion System Determines Environmental Fitness of Epidemic *Vibrio parahaemolyticus* in the Interaction with Bacterivorous Protists. PLoS ONE 6(5): e20275. doi:10.1371/journal.pone.0020275.

Meador, C., Parsons, M., Bopp, C., Gerner-Smidt, P., Painter, J., Vora, G. (2007) Virulence Gene- and Pandemic Group-Specific Marker Profiling of Clinical Vibrio parahaemolyticus Isolates. J. Clin. Microbiol. April 2007 vol. 45 no. 4 1133-1139.
Meibom KL, Blokesch M, Dolganov NA, Wu CY, Schoolnik GK: Chitin induces natural competence in *Vibrio cholerae*. Science 2005, 310(5755):1824-1827.

Meibom, K., Li, X., Nielsen, A., Wu C., Roseman, S., Schoolnik G. (2004) The *Vibrio cholerae* chitin utilization program. Proc Natl Acad Sci USA. 101(8):2524-2529.

Mekalanos JJ, Sack DA: RS1 element of *Vibrio cholerae* can propagate horizontally as a filamentous phage exploiting the morphogenesis genes of CTXΦ. Infect Immun 2002, 70(1):163-170.

Merrell, D., Hava, D., Camilli, A. (2002) Identification of novel factors involved in colonization and acid tolerance of *Vibrio cholerae*. Molecular Microbiology. 43(6)1471-1491.

Mezrioui, N., Oufdou, K. & Baleux, B. (1995). Dynamics of non-O1 *Vibrio cholerae* and fecal coliforms in experimental stabilization ponds in the arid region of Marrakesh, Morocco, and the effect of pH, temperature, and sunlight on their experimental survival. Canadian Journal of Microbiology, 41, 489-498.

Moustafa, I., Connaris, H., Taylor, M., Zaitsev, V., Wilson, J. C., Kiefel, M. J., von Itzstein, M. & Taylor, G. (2004). Sialic acid recognition by *Vibrio cholerae* neuraminidase. J Biol Chem. 279:40819-40826.

Mukerjee, S. (1963) The Bacteriophage-Susceptibility Test in Differentiating *Vibrio cholerae* and *Vibrio* el tor. Bull Wld Hlth Org. 28:333-336.

Mukhopadhyay AK, Chakraborty S, Takeda Y, Nair GB, Berg DE: Characterization of VPI pathogenicity island and CTXΦ prophage in environmental strains *of Vibrio cholerae*. J Bacteriol 2001, 183(16):4737-4746.

Murphy, R. A. & Boyd, E. F. (2008). Three pathogenicity islands of *Vibrio cholerae* can excise from the chromosome and form circular intermediates. J Bacteriol. 190, 636-647.

Myers EW, Miller W: Optimal alignments in linear space. Comput Appl Biosci 1988, 4(1):11-17.

Myers, E.W., and Miller, W. (1988) Optimal alignments in linear space. Comput Appl Biosci. 4: 11-17.

Nair GB, Oku Y, Takeda Y, Ghosh A, Ghosh RK, Chattopadhyay S, Pal SC, Kaper JB, Takeda T: Toxin profiles of *Vibrio cholerae* non-O1 from environmental sources in Calcutta, India. Appl Environ Microbiol 1988, 54(12):3180-3182.

Nair, G. B., T. Ramamurthy, S. K. Bhattacharya, B. Dutta, Y. Takeda, and D. A. Sack (2007) Global dissemination of *Vibrio parahaemolyticus* serotype O3:K6 and its serovariants. Clin Microbiol Rev 20:39-48.

Nair, G.B., Safa, A., Bhuiyan, N.A., Nusrin, S., Murphy, D., Nicol, C. et al. (2006) Isolation of *Vibrio cholerae* O1 strains similar to pre-seventh pandemic El Tor strains during an outbreak of gastrointestinal disease in an island resort in Fiji. J Med Microbiol. 55: 1559-1562.

Narkevich, M.I., Onischenko, G.G., Lomov, J.M., Moskvitina, E.A., Podosinnikova, L.S., and Medinsky, G.M. (1993) The seventh pandemic of cholera in the USSR, 1961-1989. Bull World Health Organ 71: 189-196.

Nasu, H., T. Iida, T. Sugahara, Y. Yamaichi, K. S. Park, K. Yokoyama, K. Makino, H. Shinagawa, and T. Honda (2000) A filamentous phage associated with recent pandemic *Vibrio parahaemolyticus* O3:K6 strains. J Clin Microbiol 38:2156-2161.

Noble, R.T., Weisberg, S.B., Leicester M. K., McGee, C., Ritter, K., Vainik, P. & Walker, K. (2003). Comparison of beach bacterial water quality indicator measurement methods. Environmental Monitoring and Assessment, 81, 301-312.

Okuda, J., M. Ishibashi, E. Hayakawa, T. Nishino, Y. Takeda, A. K. Mukhopadhyay, S. Garg, S. K. Bhattacharya, G. B. Nair, and M. Nishibuchi (1997) Emergence of a unique O3:K6 clone of *Vibrio parahaemolyticus* in Calcutta, India, and isolation of strains from the same clonal group from Southeast Asian travelers arriving in Japan. J Clin Microbiol 35:3150-3155.

Okura, M., R. Osawa, A. Iguchi, E. Arakawa, J. Terajima, and H. Watanabe (2003) Genotypic analyses of *Vibrio parahaemolyticus* and development of a pandemic group-specific multiplex PCR assay. J Clin Microbiol 41:4676-82.

Oliver, J.D. (2005) *Vibrio vulnificus*. In Oceans and Health: Pathogens in the Marine Environment. Belkin, S., and Colwell, R.R. (eds.). New York, NY: Springer Science, pp. 253-276.

Olivier, V., Haines III, GK., Tan, Y., and Satchell, KJF. (2007) Hemolysin and the Multifunctional Autoprocessing RTX Toxin Are Virulence Factors during Intestinal Infection of Mice with *Vibrio cholerae* El Tor O1 Strains. Infect. Immun. 75(10):5035-5042.

Onifade, T., Hutchinson, R., Van Zile, K., Bodager, D., Baker, R., Blackmore, C. (2011) Toxin producing *Vibrio cholerae* O75 outbreak, United States, March to April 2011. Euro Surveill. 16(20):19870.

Ono, T., K. S. Park, M. Ueta, T. Iida, and T. Honda (2006) Identification of proteins secreted via *Vibrio parahaemolyticus* type III secretion system 1. Infect Immun 74:1032-42.

O'Shea, Y.A., Finnan, S., Reen, FJ., Morrissey, JP., O'Gara F., and Boyd. EF. (2004) The *Vibrio* seventh pandemic island-II is a 26.9 kb genomic island present in *Vibrio cholerae* El Tor and O139 serogroup isolates that shows homology to a 43.4 kb genomic island in V. vulnificus. Microbiology. 150(12):4053-4063.

Osorio, C., Crawford, J., Michalski, J., Martinez-Wilson, H., Kaper, J., and Camilli, A. (2005) Second-Generation Recombination-Based In Vivo Expression Technology for Large-Scale Screening for *Vibrio cholerae* Genes Induced during Infection of the Mouse Small Intestine. Infect. Immun. 73(2)972-980.

Pacha RE, Kiehn ED: Characterization and relatedness of marine vibrios pathogenic to fish: physiology, serology, and epidemiology. Journal of Bacteriology 1969, 100(3):1242-1247.

Park, K. S., T. Ono, M. Rokuda, M. H. Jang, K. Okada, T. Iida, and T. Honda (2004) Functional characterization of two type III secretion systems of *Vibrio parahaemolyticus*. Infect Immun 72:6659-65.

Parks AR, Peters JE: Tn7 elements: Engendering diversity from chromosomes to episomes. Plasmid 2009, 61(1):1-14.

Petursson, S. (1968) Kræklingurinn. Náttúrufræðingurinn. 37(1-2):12-23.

Pruzzo, C., Vezzulli, L., and Colwell, R.R. (2008) Global impact of *Vibrio cholerae* interactions with chitin. Environ Microbiol 10: 1400-1410.

Purdy, A.E., Balch, D., Lizárraga-Partida, ML., Islam, MS., Martinez-Urtaza, J., Huq, A., et al. (2010) Diversity and distribution of cholix toxin, a novel ADP ribosylating factor from *Vibrio cholerae*. Environmental Microbiology Reports. 2(1):198-207.

Ramamurthy, T., Garg, S., Sharma, R., Bhattacharya, S.K., Balakrish, G., Nair, T. et al. (1993) Emergence of novel strains of *Vibrio cholerae* with epidemic potential in southern and eastern India. Lancet 341: 703-704.

Rao, A., and Stockwell, B.A. (1980) The Queensland cholera incident of 1977. 1. The index case. Bull World Health Organ. 58: 663-664.

Rivera, I.N., Chun, J., Huq, A., Sack, RB., and Colwell, RR. (2001) Genotypes associated with virulence in environmental isolates of *Vibrio cholerae*. Appl. Environ. Microbiol. 67(6):2421-2429.

Rogers RC, C.R., Cossins YM, Murphy DM, Bourke AT. (1980) The Queensland cholera incident of 1977. Bull World Health Organ. 58: 665-669.

Rosa, G., and Clasen, T. (2010) Estimating the Scope of Household Water Treatment in Low- and Medium-Income Countries. Am J Trop Med Hyg 82: 289-300.

Roslev, P., Bjergbaek, L. A. & Hesselsoe, M. (2004). Effect of oxygen on survival of faecal pollution indicators in drinking water. Journal of Applied Microbiology 96, 938-945.

Sæther, OA. (2009) *Telmatogeton murrayi* sp. n. from Iceland and *T. japonicas Tokunaga* from Madeira (*Diptera: Chironomidae*). Aquatic Insects. 31(1) 31-44.

Safa, A., Bhuiyan, N.A., Murphy, D., Bates, J., Nusrin, S., Kong, R.Y. et al. (2009) Multilocus genetic analysis reveals that the Australian strains of *Vibrio cholerae* O1 are similar to the pre-seventh pandemic strains of the El Tor biotype. J Med Microbiol 58: 105-111.

Safrin, S., Morris, J.G., Jr., Adams, M., Pons, V., Jacobs, R., and Conte, J.E., Jr. (1988) Non-O:1 *Vibrio cholerae* bacteremia: case report and review. Rev Infect Dis 10: 1012-1017.

Saito, M., Kitamura, H., Sugiyama, K. (2001). Occurrence of gangliosides in the common squid and pacific octopus among protostomia Biochimica et Biophysica Acta (BBA) - Biomembranes. 1511. 2(2)271–280.

Saitou N, Nei M: The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 1987, 4(4):406-425.

Sarkar, BL., Nair, GB., Banerjee, AK., and Pal, SC. (1985) Seasonal distribution of Vibrio parahaemolyticus in freshwater environs and in association with freshwater fishes in Calcutta. Appl. Environ. Microbiol. 49(1) 132-136.

Schneider, D.  and Parker, C. (1982). Purificartion and characterization of the nucinase of *V. cholerae*. J Infect Dis. 145, 474-482.

Schubel, J., and Pritchard, D. (1986) Responses of upper Chesapeake Bay to variations in discharge of the Susquehanna River. Estuaries and Coasts. 9(4)236-249.

Schuster, BM., Tyzik, AI., Donner, RA., Striplin, MJ., Almagro-Moreno, S., Jones, SH., (2011) Ecology and Genetic Structure of a Northern Temperate *Vibrio cholerae*

Population Related to Toxigenic Isolates. Appl. Environ. Microbiol. 77(21): 7568-7575.

Serichantalergs, O., N. A. Bhuiyan, G. B. Nair, O. Chivaratanond, A. Srijan, L. Bodhidatta, S. Anuras, and C. J. Mason. (2007) The dominance of pandemic serovars of *Vibrio parahaemolyticus* in expatriates and sporadic cases of diarrhoea in Thailand, and a new emergent serovar (O3 : K46) with pandemic traits. J Med Microbiol 56:608-613.

Shalu, O., Golenishcheva, Y., Smolikova, L., Galtseva, G., Bozhko, B., Monakhova1, Y. (2010) Evaluation of the virulence of *Vibrio parahaemolyticus* strains isolated in Russia and contiguous states. Epidemiology and Infectious Diseases. No. 6.

Shannon, J.D., Kimbrough, R.C. (2006) Pulmonary Cholera Due to Infection with a Non-O1 *Vibrio cholerae* Strain. J Clin Microbiol 44: 3459–3460.

Shapiro, R., Altekruse, S., Hutwagner, L., Bishop, R., Hammond, R., Wilson, S., Ray, B., Thompson, S., Tauxe, R., and Griffin, P. (1998) The Role of Gulf Coast Oysters Harvested in Warmer Months in *Vibrio vulnificus* Infections in the United States, 1988–1996. J Infect Dis. 178(3):752-759.

Shikulov, V., Khaĭtovich, A., Bogatyreva, L. (1980) Isolation of halophilic vibrios from humans and the environment. Zh Mikrobiol Epidemiol Immunobiol. (6):38-40.

Shin OS, et al. 2011. Type III secretion is essential for the rapidly fatal diarrheal disease caused by non-O1, non-O139 *Vibrio cholerae*. mBio 2(3):e00106-11. doi:10.1128/mBio.00106-11.

Shinoda S, Nakagawa T, Shi L, Bi K, Kanoh Y, Tomochika K, Miyoshi S, Shimada T: Distribution of virulence-associated genes in *Vibrio mimicus* isolates from clinical and environmental origins. Microbiol Immunol 2004, 48(7):547-551.

Shinoda, S. (2011) Proteases Produced by *Vibrio cholerae* and Other Pathogenic Vibrios: Pathogenic Roles and Expression. In Epidemiological and Molecular Aspects on Cholera, Infectious Disease. Ramamurthy, T., Bhattacharya, S.K. (eds.) New York, NY: Springer Science, pp. 245-258.

Sigurdsson, G.H., Möller, AD., Kristinsson, B., Gudlaugsson, O., Kárason, S., Sigurdsson, SE., Kristjánsson, M., and Sigvaldason, K. (2009) Intensive care patients with influenza A (H1N1) infection in Iceland. Laeknabladid. 96(2):83-90.

Silva, A.J., Leitch, G.J., Camilli, A., and Benitez, J.A. (2006) Contribution of Hemagglutinin/Protease and Motility to the Pathogenesis of El Tor Biotype Cholera. Infect. Immun. 74(4): 2072-2079.

Son, M., and Taylor, R. (2011) Genetic Screens and Biochemical Assays to Characterize *Vibrio cholerae* O1 Biotypes: Classical and El Tor. Curr. Protoc. Microbiol. 22:6A.2.1-6A.2.17.

Suthienkul, O., Ishibashi, M., Iida, T., Nettip, N., Supavej, S., Eampokalap, B., Makino, M., and Honda, T. (1995) Urease Production Correlates with Possession of the trh Gene in *Vibrio parahaemolyticus* Strains Isolated in Thailand J Infect Dis. 172(5): 1405-1408.

Tacket, C., Taylor, R., Losonsky, G., Lim, Y., Nataro, J., Kaper, J., and Levine, M. (1998) Investigation of the Roles of Toxin-Coregulated Pili and Mannose-Sensitive Hemagglutinin Pili in the Pathogenesis of *Vibrio cholerae* O139 Infection. Infect. Immun. 66(2)692-695.

Tamura, K., Dudley, J., Nei, M., & Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution. 24:1596-1599.

Taviani, E., Grim, CJ., Choi, J., Chun, J., Haley, B., Hasan, NA., et al. (2010) Discovery of novel *Vibrio cholerae* VSP II genomic islands using comparative genomic analysis. FEMS Microbiology Letters. 308(2):130-137.

Thompson CC VA, Souza RC, Vasconcelos ATR, Vesth T, Alves N, Ussery DW, Iida T, Thompson FL: Genomic Taxonomy of the Vibrios. In: Vibrio2009. Rio de Janeiro, Brasil; 2009.

Thompson FL, and Swings, J.: Taxonomy of the Vibrios. In: Biology of the Vibrios. Edited by Thompson FL, Austin, B., and J. Swings. Washington, D.C: ASM Press; 2006: 29-43.

Thompson FL, Iida T, Swings J: Biodiversity of vibrios. Microbiol Mol Biol Rev 2004, 68(3):403-431.

Thompson, C., Souza, V., Vasconcelos, R., Vesth, A., Alves, N., Ussery, D., Iida, T., Thompson, F.: Genomic Taxonomy of the Vibrios. In: Vibrio2009. Rio de Janeiro, Brasil; 2009.

Thompson, C., Vicente, A., Souza R., Vasconcelos, A., Vesth, T., Alves, N., Ussery, D, Iida, T., Thompson, F.: Genomic taxonomy of vibrios. BMC Evolutionary Biology 2009, 9(1):258-274.

Tischler, A.D., Camilli, A. (2005) Cyclic Diguanylate Regulates *Vibrio cholerae* Virulence Gene Expression. Infect Immun 73:5873-5882.

Tobin-D'Angelo, M., Smith, A., Bulens, S., Thomas, S., Hodel, M., Izumiya, H., Arakawa, E., Morita, M., Watanabe, Marin, C., Parsons, M., Greene, K., Cooper, K., Haydel, D., Bopp, C., Yu, P., and Mintz, E. (2008) Severe Diarrhea Caused by Cholera Toxin–Producing *Vibrio cholerae* Serogroup O75 Infections Acquired in the Southeastern United States. Clin Infect Dis. 47(8):1035-1040.

Tolosa, I., de Mora, S., Sheikholeslami, M.R., Villeneuve, J.-P., Bartocci, J., and Cattini, C. (2004) Aliphatic and aromatic hydrocarbons in coastal Caspian Sea sediments. Marine Pollution Bull 48.

Tunkijjanukij, S., Giæver, H., Chin, CCQ., and Olafsen, JA. (1998) Sialic acid in hemolymph and affinity purified lectins from two marine bivalves. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology. 119(4):705-713.

Turner, J.W., Good, B., Cole, D., and Lipp, E.K. (2009) Plankton composition and environmental factors contribute to *Vibrio* seasonality. ISME J. 3:1082-1092.

Turnsek, M. (2010) A Newly Recognized Human Pathogen in the *Vibrio cholerae-mimicus* species complex. Vibrios in the Environment. Biloxi, MS.

Udden, S., Zahid, M., Biswas, K., Ahmad, Q., Cravioto, A., Nair, GB., Mekalanos, J., and Faruque, S. (2008) Acquisition of classical CTX prophage from *Vibrio cholerae*

O141 by El Tor strains aided by lytic phages and chitin-induced competence. Proc Natl Acad Sci USA. 105(33)11951-11956.

United States Environmental Protection Agency (2002). Method 1603: Escherichia coli (E. coli) in water by membrane filtration using modified membrane-thermotolerant *Escherichia coli* agar (modified mTEC). USEPA, Office of Water, Washington, D.C.

United States Environmental Protection Agency (2006). Method 1103.1: *Escherichia coli* (*E. coli*) in water by membrane filtration using membrane-yhermotolerant *Escherichia coli* agar (mTEC). USEPA, Office of Water, Washington, D.C.

United States Environmental Protection Agency. (1986). Ambient Water Quality Criteria for Bacteria - 1986. USEPA, Office of Water, Washington, D.C.

Vanden Broeck, D., Horvathb, C., De Wolfa, M. (2007) *Vibrio cholerae*: Cholera toxin. The International Journal of Biochemistry & Cell Biology. 39(10)1771–1775.

Vanlaere E, Baldwin A, Gevers D, Henry D, De Brandt E, LiPuma JJ, Mahenthiralingam E, Speert DP, Dowson C, Vandamme P: Taxon K, a complex within the *Burkholderia cepacia* complex, comprises at least two novel species, *Burkholderia contaminans* sp. nov. and *Burkholderia lata* sp. nov. International Journal of Systematic and Evolutionary Microbiology 2009, 59(1):102-111.

Vora, G., Meador, CE., Bird, MM., Bopp, CA., Andreadis, JD., and Stenger, DA. (2005) Microarray-based detection of genetic heterogeneity, antimicrobial resistance, and the viable but nonculturable state in human pathogenic *Vibrio* spp. Proc. Natl. Acad. Sci. 102(52):19109-19114.

Vuddhakul, V., Chowdhury, A., Laohaprertthisan, V., Pungrasamee1, P., Patararungrong, N., Thianmontri, P., Ishibashi, M., Matsumoto, C., and Nishibuchi, M. (2000) Isolation of a Pandemic O3:K6 Clone of a *Vibrio parahaemolyticus* Strain from Environmental and Clinical Sources in Thailand. Appl. Environ. Microbiol. 66(6) 2685-2689.

Waldor, M.K., and Mekalanos, J.J. (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. Science 272: 1910-1914.

Watnick, P.I., Lauriano, C.M., Klose, K.E., Croal, L., Kolter, R. (2001) The absence of a flagellum leads to altered colony morphology, biofilm development and virulence in *Vibrio cholerae* O139. Molecular Microbiology. 39(2): 223–235.

Wheeler, A. L., Hartel, P. G., Godfrey, D. G., Hill, J. L. & Segars, W. I. (2002). Potential of *Enterococcus faecalis* as a human fecal indicator for microbial source tracking. Journal of Environmental Quality. 31, 1286-1293.

WHO – World Health Organization (2011) http://www.euro.who.int/en/what-we-do/health-topics/emergencies/international-health-regulations/news/news/2011/06/ukraine-reports-14-cholera-cases. Accessed on 20 March 2012.

Wolfgang M, Lauer P, Park HS, Brossay L, Hebert J, Koomey M: PilT mutations lead to simultaneous defects in competence for natural transformation and twitching motility in piliated *Neisseria gonorrhoeae*. Molecular Microbiology 1998, 29(1):321-330.

Wong, H. C., C. C. Liu, T. M. Pan, T. K. Wang, C. L. Lee, and D. Y. Shih. (1999) Molecular typing of *Vibrio parahaemolyticus* isolates, obtained from patients involved in food poisoning outbreaks in Taiwan, by random amplified polymorphic DNA analysis. J Clin Microbiol 37:1809-1812.

Wong, H. C., C. Y. Ho, L. P. Kuo, T. K. Wang, C. L. Lee, and D. Y. Shih (1999) Ribotyping of *Vibrio parahaemolyticus* isolates obtained from food poisoning outbreaks in Taiwan. Microbiol Immunol 43:631-636.

Wong, H. C., M. C. Chen, S. H. Liu, and D. P. Liu. (1999) Incidence of highly genetically diversified Vibrio parahaemolyticus in seafood imported from Asian countries. Int J Food Microbiol 52:181-188.

Wong, H. C., S. H. Liu, L. W. Ku, I. Y. Lee, T. K. Wang, Y. S. Lee, C. L. Lee, L. P. Kuo, and D. Y. Shih (2000) Characterization of Vibrio parahaemolyticus isolates obtained from foodborne illness outbreaks during 1992 through 1995 in Taiwan. J Food Prot. 63:900-906.

Woo, P. C. Y., A. M. Y. Fung, S. S. Y. Wong, H.-W. Tsoi, and K.-Y. Yuen. (2001) Isolation and Characterization of a *Salmonella enterica* Serotype Typhi Variant and Its Clinical and Public Health Implications. J. Clin. Microbiol. 39:1190-1194.

Yamasaki, S., Shimizu, T., Hoshino, K., Ho, S., Shimada, T., Nair, G., Takeda, Y. (1999) The genes responsible for O-antigen synthesis of *Vibrio cholerae* O139 are closely related to those of Vibrio cholerae O22. Gene. 237(2)321-332.

Zakhariev, Z. (1976) The spread of parahemolytic vibrios in humans in the Burgass Region (Bulgaria). Zh Mikrobiol Epidemiol Immunobiol. 46-48.

Zhu, L., Cai, JP., Chen, Q., and Yu, SY. (2007) Specific detection of toxigenic *Vibrio cholerae* based on in situ PCR in combination with flow cytometry. Biomed. Environ. Sci. 20(1):64-69.

Zo YG: Phylogenomic and structural analyses of *Vibrio cholerae* populations and endemic cholera. PhD Thesis. University of Maryland, College Park, Marine Estuarine and Environmental Science; 2005.

# *Curriculum vitae*

**Bradd J. Haley**
Maryland Pathogen Research Institute
University of Maryland, College Park
Bioscience Research Building Bldg. 413
College Park, MD 20742

## EDUCATION:

**PhD in Environmental Molecular Biology** (May 2012)
- *The University of Maryland, College of Life and Chemical Sciences*
Dissertation: Polyphasic Analyses of Human Pathogens in the Natural Environment (Advisor: Dr. Rita R. Colwell).

**Master of Science in Environmental Health (MSEH)** (December 2006)
- *The University of Georgia, College of Public Health*
Master's Degree Thesis: Impact of Climate and Weather on *Salmonella* Densities in the Aquatic Environment (Advisor: Dr. Erin K. Lipp, Co-advisor: Dr. Dana Cole).

**BS in Environmental Geosciences** (August 2001)
- Department of Geology and Geophysics, *Boston College*

## RESEARCH INTERESTS:

- Microbial genomics
- Microbial ecology
- Environment and disease
- Microbial risk assessment

## PROFESSIONAL EXPERIENCE:

Graduate Research Assistant, *The University of Maryland*, Maryland Pathogen Research Institute, January 2008 to Present. College Park, MD.

Teaching Assistant for Introduction to Molecular Biology class (100 level), *The University of Maryland*, College of Life and Chemical Sciences, August 2007 to December 2007. College Park, MD.

J. William Fulbright Fellow, *University of Iceland (Haskoli Islands)*, Institute of Biology, October 2006 to June 2007. Reykjavik, Iceland.

Graduate Assistant, *The University of Georgia*, College of Public Health, August 2004 to September 2006. Athens, GA.

Teaching Assistant for Environmental Microbiology class (4000/6000 level), *The University of Georgia*, College of Public Health, January 2006 to May 2006. Athens, GA.

Assistant Laboratory Supervisor, *Forensic Analytical Specialties, Inc.*, Environmental Microscopy Laboratory, January 2002 to January 2004. San Francisco, CA.

Undergraduate Research Assistant, *Boston College*, Department of Geology and Geophysics, May 2000 to May 2001. Chestnut Hill, MA.

Classroom Assistant, *Boston College*, Boston College Campus School, academic years 1997, 1998, and 2000. Chestnut Hill, MA.

## PRESENTATIONS:

**B. Haley,** C.M. Diaz, and E. Benediktsdóttir. 2009. *Vibrio cholerae* at Geothermal Sites in Iceland. Federation of European Microbiological Societies (FEMS) Congress 2009, Göteborg, Sweden.

**B. Haley**, D. Cole, and E.K. Lipp. 2005. Environmental Prevalence of *Salmonella* in a Southeastern Georgia Watershed. American Society for Microbiology Southeastern Branch meeting, St. Pete's Beach, FL.

**B. Haley**, D. Cole, and E.K. Lipp. 2004. Factors Contributing to the Persistence of salmonellae in Georgia's Watersheds. College of Public Health Student Seminar, University of Georgia, Athens, GA.

## INVITED PRESENTATIONS:

**B. Haley.** 2011. Modern Environmental Microbiology and Simple Methods for Infectious Disease Prevention. Beta Iota Omicron, professional biology fraternity. University of Maryland, College Park.

**B. Haley**, D. Cole and E.K. Lipp. 2007. Weather and Waterborne Pathogens in Southeastern Georgia, USA. Invited Presentation at the Icelandic Society for Microbiology occasional meeting, Reykjavik, Iceland.

**POSTERS:**

A. Kahler, **B. Haley**, B. Mull, A. Chen, N. Purcell, R. Colwell, A. Huq, V. Hill. 2012. Evaluation of Human Pathogens in Surface Waters of Haiti. American Society for Microbiology General Meeting, San Francisco, CA.

**B. Haley**, E. Taviani, A. Chen, A. Phillipy, A. Huq, R.R. Colwell, I. Knight. 2012. Validation of Unique Signature Discovery in Select Agents and an Emerging Pathogen. ASM Biodefense and Emerging Pathogens Research Meeting. Washington, DC (12-BIO-A-342-ASM).

R. Akhmedov, S. Akhmadova, M. Rajabov, **B. Haley**, S. Gurbonov, R.R. Colwell, and A. Huq. 2011. Detection of *Vibrio cholerae* in Natural Waters Including Drinking Water Reservoirs in Azerbaijan. 3rd Annual Conference of the Biosafety Association for Central Asia & the Caucasus (BACAC). Tbilisi, Georgia

**B. Haley**, E. Taviani, J. Choi, C.J. Grim, A. Chen, P. Clark, L. Sancomb, M. Tamburri, R.R. Colwell, and A. Huq. 2011. Comparison of Methods for Quantifying Bacterial Indicators in Water from Urban Brackish Environment. American Society for Microbiology General Meeting, New Orleans, LA

P. Clark, A. Chen, N. Hasan, **B. Haley**, E. Taviani, L. Sancomb, A. Huq, and R.R. Colwell. 2011. Intestinal Microflora Isolated from Haitian Cholera Patients. American Society for Microbiology General Meeting, New Orleans, LA.

**B. Haley**, N.A. Hasan, E. Taviani, A. Chen, A. Huq, and R.R. Colwell. 2011. Endemic *Vibrio cholerae* in an Urban, Cholera-Free Region, not Influenced by Infected Persons, or Marine, Estuarine, and Ballast Discharged Waters. American Society for Microbiology General Meeting, New Orleans, LA.

P. Clark, **B. Haley**, A. Chen, N.A. Hasan, E. Taviani, A. Huq, R.R. Colwell. 2010. Occurrence of Human Pathogenic Vibrios Associated with Polychaetes in Oyster Beds in the Chesapeake Bay. Bioscience Day, University of Maryland, College Park.

A. Chen, N.A. Hasan, **B. Haley**, E. Taviani, P. Clark, D. White, S. Harvey, M. Tarnowski. 2011. Virulence Factors of *Vibrio parahaemolyticus* in the Chesapeake Bay, Maryland. American Society for Microbiology General Meeting, New Orleans, LA.

R. Akhmedov, S. Akhmadova, M. Rajabov, **B. Haley**, S. Gurbonov, R.R. Colwell, and A. Huq. 2011. Ecological investigation of *Vibrio cholerae* in Surface Waters of Azerbaijan. American Society for Microbiology General Meeting, New Orleans, LA.

J. Choi, **B. Haley**, N.A. Hasan, C.J. Grim, E. Taviani, A.C. Munk, T. Brettin, D. Bruce, et al. 2011. Genomic Islands of *Vibrio orientalis* and their homologs in the *Vibrionaceae*. American Society for Microbiology General Meeting, New Orleans, LA.

E. Taviani, M. Spagnoletti, D. Ceccarelli, **B. Haley**, N.A. Hasan, A. Chen, M.M. Columbo, A. Huq, R.R Colwell. 2011. Genomic Analysis of an Integrative Conjugative Element (ICE) Identified in Clinical *V. cholerae* O37: A New ICE or a Genomic Island? American Society for Microbiology General Meeting, New Orleans, LA.

**B. Haley**, A. Chen, P. Clark, E. Taviani, N.A. Hasan, C. Grim, E. Benediktsdottir, A. Huq, R.R. Colwell. 2010. Endemic *Vibrio cholerae* Populations in an Historically Cholera-Free Country. Vibrios in the Environment 2010. Biloxi, Mississippi.

N.A. Hasan, C. Grim, **B. Haley**, J. Chun, M. Alam, E. Taviani, M. Hoq, A. Munk, et al. 2010. Comparative Genomics of Clinical and Environmental *Vibrio mimicus*. Vibrios in the Environment 2010. Biloxi, Mississippi.

E. Taviani, A. Chen, **B. Haley**, N. A. Hasan, A. Huq, R. R. Colwell. 2010. Geographical distribution and molecular characterization of Integrative Conjugative Elements in environmental *Vibrio cholerae*. Vibrios in the Environment 2010. Biloxi, Mississippi.

**B. Haley**, R.R. Colwell, E. Benediktsdóttir. 2009. The endemicity of *Vibrio cholerae* in an historically cholera-free country. Icelandic Biological Society Annual Meeting (Líffræðiráðstefnan 2009), Reykjavik, Iceland.

G. Caburlotto, E. Taviani, **B. Haley**, T. Kokashvili, N. Janelidze, M. M. Lleo, A. Huq, et al. 2009. Design of unique signature sequences and whole genomic analysis of environmental *Vibrio parahaemolyticus*. Vibrio2009, Rio de Janeiro, Brazil

E. Taviani, C. J. Grim, **B. Haley** N.A. Hasan, J. Chun, A. Huq, and R.R. Colwell. 2009. Discovery of new variants of *Vibrio* seventh pandemic islands I and II through comparative genomics and their distribution among the *Vibrio cholerae* population. Vibrio2009, Rio de Janeiro, Brazil.

**B. Haley**, C. J. Grim, N.A. Hasan, E. Taviani, J. Chun, A. Huq, and R.R. Colwell. 2009. *Vibrio metecus* sp. nov. and *Vibrio parilis* sp. nov., novel *Vibrio* species previously characterized as variant *V. cholerae* and *V. mimicus*. Vibrio2009, Rio de Janeiro, Brazil.

G. Caburlotto, E. Taviani, **B. Haley**, T. Kokashvili, N. Janelidze, M. M. Lleo, A. Huq, et al. 2009. Discovery of *Vibrio parahaemolyticus* signature sequences and whole genomic analysis of environmental strains. American Society for Microbiology General Meeting, Philadelphia, PA.

**B. Haley**, T. Kokoshvili, N. Janelidze, E. Taviani, C. Grim, M. Tediashvili, A. Huq and R. Colwell. 2009. The Occurrence of *Vibrio parahaemolyticus* in the Black Sea, Georgia. American Society for Microbiology General Meeting, Philadelphia, PA.

**B. Haley** and E. Benediktsdóttir. 2007. *Vibrio cholerae* at Geothermal Sites in Iceland. Vibrio2007, Paris, France.

**B. Haley**, D. Cole, and E.K. Lipp. 2006. Survival of *Salmonella* spp. in Aquatic Environments. American Society for Microbiology General Meeting, Orlando, FL

**B. Haley**, D. Cole, and E.K. Lipp. 2006. Impact of Climate and Weather on *Salmonella* Densities in the Aquatic Environment. Ocean Sciences Meeting of the American
Society for Limnology and Oceanography, Honolulu, HI.

**B. Haley**, J.C. Venton, E. Vereen, T.J. Onifade, K. Caspary, E.K. Lipp. 2005. Evaluation of Enterococci and Fecal Coliform Bacteria as Indicators of Viral Pathogen Loads in Marine Recreational Waters. American Society for Microbiology General Meeting Atlanta, GA.

**B. Haley**, K. Harrison, and P.K. Strother. 2001. Using $\delta^{13}C$ to reconstruct paleontological environments and paleoclimatic conditions of the middle Cambrian Bright Angel Shale in the eastern Grand Canyon of northern Arizona. American Geophysical Union meeting, Boston, MA.

## BOOK CHAPTERS

A. Huq, **B. Haley**, N.A. Hasan, E. Taviani, A. Chen, and R.R. Colwell. Nonenteric Gamma Proteobacteria, Detection, Isolation, and Identification of *Vibrio cholerae* from the Environment *in* Current Protocols in Microbiology.

## PUBLICATIONS

E. Taviani, M. Spagnoletti, D. Ceccarelli, **B. Haley**, N. A. Hasan, A. Chen, M. M. Colombo, A. Huq, and R. R. Colwell. ICEVchBan8 and the extreme flexibility of the mosaic structure of ICEs. *In press* (*accepted*).

S. Gurbanov, R. Akhmadov, and G. Shamkhalova, S. Akhmadova, **B. Haley**, R.R. Colwell, and A. Huq. 2011. Occurrence of *Vibrio cholerae* in Municipal and Natural Waters and Incidence of Cholera in Azerbaijan. *In press (accepted).*

**B. Haley**, A. Chen, C. Grim, P. Clark, C. Diaz, E. Taviani, N. Hasan, E. Sancomb, W. Elnemr, M. Islam, A. Huq, R. Colwell and E. Benediktsdóttir. *Vibrio cholerae* in an historically cholera-free country. Environmental Microbiology Reports. *In Press.* doi:10.1111/j.1758-2229.2012.00332.x

W.R. Easterday, K.L. Kausrud, B. Star, L. Heier, **B. Haley**, V. Ageyev, R.R. Colwell & N.C. Stenseth. 2011. An additional step in the transmission of Yersinia pestis? The ISME Journal. 6: 231–236.

N.A. Hasan, C.J. Grim, **B. Haley**, J. Chun, M. Alam, E. Taviani, M. Hoq, et al. 2010. The genome sequence of clinical and environmental *Vibrio mimicus* and a comparative genomic analysis with related species. Proceedings of the National Academy of Sciences. 107(49):21134-21139

E. Taviani, C.J. Grim, J. Choi, J. Chun, **B. Haley**, N.A. Hasan, A. Huq, and R.R. Colwell. 2010. Discovery of a novel VSP-II Using Comparative Genomic Analyses. FEMS Microbiology Letters. 308(2): 130–137.

C.J. Grim, N.A. Hasan, E. Taviani, **B. Haley**, J. Chun, T.S. Brettin, D.C. Bruce, et al. 2010. Genome sequence of hybrid *V. cholerae* O1 MJ-1236, B-33 and CIRS101 and comparative genomics with *V. cholerae*. Journal of Bacteriology. 192: 3524-3533.

**B. Haley**, C.J Grim, N.A. Hasan, J. Chun, A. Huq, and R.R. Colwell. 2010. Comparative genomic analysis reveals evidence of two novel *Vibrio* species closely related to *V. cholerae*. BMC Microbiology. 10:154

C.J. Grim, J. Choi, J. Chun, Y.S. Jeon, E. Taviani, **B. Haley**, A. Huq, R.R. Colwell. 2010. Occurrence of *Vibrio cholerae* 7th Pandemic VSP-I island and a new variant. OMICs. 14(1):1-7.

**B. Haley**, C.J. Grim, N.A. Hasan, E. Taviani, J. Chun, T. Brettin, D. Bruce, et al. 2010. The Pre-7$^{th}$ Pandemic *Vibrio cholerae* BX 330286 El Tor Genome; Evidence for the Environment as a Genome Reservoir. Environmental Microbiology Reports. 2(1):208-216.

G. Caburlotto, **B. Haley**, M. Lleò Del Mar, A. Huq, R.R. Colwell. 2010. Serodiversity and ecological distribution of *Vibrio parahaemolyticus* in the Venetian Lagoon, Northeast Italy. Environmental Microbiology Reports. 2(1):151-157.

J. Chun, C.J. Grim, N.A. Hasan, J.H. Lee, S.Y. Choi, **B. Haley**, E. Taviani, et al. 2009. Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. Proceedings of the National Academy of Sciences. 106(36): 15442-15447.

**B. Haley**, D. Cole, and E.K. Lipp. 2009. Distribution, Diversity and Seasonality of Waterborne *Salmonella* in a Rural Watershed. Applied and Environmental Microbiology. 75(5):1248-1255.

## HONORS AND AWARDS:

- Student Travel Grant, American Society for Microbiology General Meeting, 2009.

- Outstanding Young Alumnus Award for Excellence in Academics, The Episcopal Academy, 2006.

- J. William Fulbright Fellowship, The United States Department of State, 2006 to 2007 (for study and research abroad in Iceland).

- University of Georgia Amazing Student, 2006 (nominee)

- President's Award for Outstanding Master's Degree Student Oral Presentation, American Society for Microbiology Southeastern Branch Meeting, 2005.

- Student Travel Grant, American Society for Microbiology Southeastern Branch Meeting, 2005.

## PROFESSIONAL AFFILIATIONS:

- American Society for Microbiology (ASM)
    - Washington, DC Branch
    - Region 12: Former Soviet Countries Branch

- Icelandic Microbiology Society

- International Society for Infectious Diseases (ISID)