GIFT

Environmental Factors, Microbial Community Dynamics and Emergence of *Vibrio cholerae*

Nur A. Hasan
Registration Number: 225 (Old Reg. No. 214)
Session: 2011-2012 (Old Session: 2006-2007)

466259

Dhaka University Library
466259

Department of Microbiology
University of Dhaka, Dhaka-1000
Maryland Pathogen Research Institute
University of Maryland College Park

Dissertation submitted to the Faculty of Biological Sciences of the University of Dhaka in partial fulfillment of the requirements for the degree of Doctor of Philosophy
2013

R
614.514
HAE
C. 1

N. B

Dhaka University Institutional Repository

CERTIFICATE

This is to certify that the thesis entitled "Environmental factors, microbial community dynamics and emergence of *Vibrio cholerae*" was carried out by Nur A. Hasan, registration number: 225, session: 2011-2012 for the fulfillment of the degree of the Doctor of Philosophy (Science), from the Department of Microbiology, University of Dhaka, Bangladesh.

This work was carried out under our supervision and the style and contents of the thesis have been approved and recommended for the award of Ph.D. degree.

466259

Professor M. Mozammel Hoq
Department of Microbiology
University of Dhaka
Dhaka-1000, Bangladesh

Professor Anwarul Huq
Maryland Pathogen Research Institute
University of Maryland
College Park, Maryland 20742, USA

ঢাকা বিশ্ববিদ্যালয় গ্রন্থাগার

Dhaka University Institutional Repository

ABSTRACT

466259

For centuries cholera has been one of the most feared diseases and *Vibrio cholerae*, a water-borne pathogen, is the causative agent of this severe watery diarrheal disease. It is the only enteric pathogen that has the potential to produce pandemic disease. The massive cholera outbreak that followed the earthquake of 2010 in Haiti reaffirmed the disease as a major public health threat for developing countries. Many impressive advances have been made in the past few decades on understanding the biology, pathophysiology, immunology, mechanism of pathogenesis, ecology and epidemiology of *V. cholerae* and of the disease cholera. Despite all these advances, the global trend is that the incidence of cholera is increasing steadily in many parts of the developing world. Since *Vibrio cholerae* is autochthonous to the aquatic environment, it is not possible to eradicate the disease but prevention is an achievable goal. This study, therefore, employed a holistic approach for acquisition of knowledge useful for lowering the burden of the disease. This was achieved by employing a combination of microbiological, hydroclimatological and molecular methods along with high-throughput genomic and metagenomic analyses, combined with broader ecological studies evaluating the role of the environment with respect to presence of the pathogens, dissecting the micro niche of the bacterium to elucidate interaction with microbial populations in the environment and human, interaction with plankton populations, as well as investigating the genomic evolution of the bacterium to catalog virulence markers and track their dissemination within and between toxigenic, non-toxigenic, and archived reference isolates. The results of environmental clinical sampling and analysis demonstrated improved detection and isolation of *V. cholerae* from its natural environment and revealed that the occurrence of toxigenic *V. cholerae* in the environment and cases of cholera were correlated with water temperature, water depth, salinity, and conductivity. Significant association was also observed between occurrence of cholera cases caused by *V. cholerae* O1 and its presence in the environment without any time lag, or with a lag of two to four weeks before appearance of cholera in Mathbaria and Bakerganj, Bangladesh, respectively. The study demonstrated a peak in *V. cholerae* abundance 1-2 months prior to an observed increase in pathogenic *V. cholerae* isolation and arrival of clinical cases to a nearby

ঢাকা
বিশ্ববিদ্যালয়
গ্রন্থাগার

Dhaka University Institutional Repository

health clinic. Investigating the distribution of eighteen virulence genetic markers, among 794 clinical and environmental *V. cholerae* samples, the study demonstrated that the distribution of certain virulence genes, was found to be similar among most *V. cholerae* O1, O139 and non-O1/O139, whereas distribution of other virulence marker genes markedly differed among toxigenic and non-toxigenic strains. Metagenomic analysis of 144 water samples collected over a period of one year from Mathbaria demonstrated that aquatic environments harbor genetically diverse and complex microbial communities including primary members of both freshwater and coastal ecosystems. The abundance and diversity of the microbial community including *Vibrio* population, exhibited a complex seasonal distribution, and specific consortia among the natural community, perhaps stimulated as a result of positive selection, might be crucial to trigger *V. cholerae* to overcome the natural competition and emerge in epidemic proportion. Comparative genome analysis of over 100 *V. cholerae* and its phylogenetic near neighbors demonstrated a variety of genomic events including inter and intra chromosomal rearrangements, gene absence and acquisitions, chromosomal integration of phages, and other mobile genetic elements (MGEs), all of which are important driving forces in the evolution of species *Vibrio*. The study also demonstrated that *V cholerae* harbors an open pan genome and the current pandemic is caused by strains belonging to a single phyletic line, diversified mainly via horizontal gene transfer from a *Vibrio* gene reservoir in the natural environment.

In conclusion, this work supports the findings of numerous prior studies that environmental factors determine cholera epidemicity and advocates that both quantitative and qualitative (genomic) data on *Vibrio* abundance in the environment are critical to estimate the risk of disease. The ubiquity of virulence factors in the environment and association of such mobile genetic elements with several other human pathogens comprises a major public health risk. The discovery of novel markers using those MGEs will advance forensic, diagnostic, therapeutic and epidemiological studies and identified genomic signatures will help not only in the development of a more accurate risk assessment should a human outbreak occur, but can also be readily utilized to monitor strain prevalence prior to human infection.

Dhaka University Institutional Repository

ACKNOWLEDGEMENTS

I would like to gratefully acknowledge my supervisor at the University of Dhaka, Prof. Md. Mozammel Hoq and at the University of Maryland Prof. Anwar Huq for their meticulous guidance, generous assistance and encouragement for successful completion of the study. I am ever grateful and highly indebted to Dr. Munirul Alam, senior scientist, icddrb, Prof. Thomas A. Cebula, Johns Hopkins University, and all of mu beloved teachers, particulrly Prof. Mahmuda Yasmin, Chairman, Department of Microbiology, University of Dhaka, for their constant support, suggestions and cooperation. My sincere appreciation and thanks also goes to my laboratory partners Guillaume Constantin de Magny, Christopher J. Grim, Elisa Taviani, Bradd J. Haley, Seon Yonug Choi, Jinna Choi, Arlene Chen, Wessam M. Elnemr, Phillip Clark, Huai Li, Shakur Abdullah, Garrett Longmire, Richard Isom, and Kelly Saeed for their invaluable help throughout the study. I would also like to express my highest gratitude and thanks to all my colleagues and collaborators at icddrb especially Abdus Sadique, Marzia Sultana, Muhammad A. Islam, Wasimul B. Chowdhury, M. Niaz Rahim, Kabir U. Ahmed, Mohammed A. Islam, Shahnewaz B. Mannan, and Shah Monjoor Rashed for all of their help.

Much of this work particularly the whole genome sequencing of the *Vibrio* isolates, would not have been possible without the help of my collaborators from the genome science group of Los Alamos National Laboratory, USA, Chris Munk, Thomas S. Brettin, David C. Bruce, Jean F. Challacombe, J. Chris Detter, Cliff S. Han, Olga Chertkov, and Elizabeth Saunders. My sincere gratitude and thanks also goes to Professor Jongsik Chun, and all his team members at the Seon National University for their valuable assistance and technical supports with bioinformatics pipeline. I am also grateful to my collaborators at the Institute for Genome Science, University of Maryland School of Medicine, Prof. Jacques Ravel, Mark Eppinger, Luke J. Tallon, Lisa Sadzewicz, and Naomi Sengamalay for their valuable help and assistance in sequencing *Vibrio cholerae* isolates.

The generous assistance of the Enteric and Food Microbiology laboratory, icddrb, Maryland Department of Environment (Kathy Brohawn, Rusty McKay and Sarah Harvey), Maryland Department of Natural Resources (Mitchell Tarnowski and

Dhaka University Institutional Repository

David White), Department of Environmental Science and Technology (Stephanie Lansing) and Center for Health Security and Infrastructure, University of Maryland (Michael McDonald) are 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.

I am delighted to pay tribute to my entire family for their tireless support, encouragement, and love all throughout.

Dhaka University Institutional Repository

© Copyright by

Nur A. Hasan

2013

Dhaka University Institutional Repository

CONTENTS

TABLE OF CONTENTS	:	i
LIST OF TABLES	:	ii
LIST OF FIGURES	:	iv
CHAPTER 1: GENERAL INTRODUCTION	:	1
CHAPTER 2: SEASONALITY OF CHOLERA IN THE COASTAL AQUATIC ENVIRONMENT OF BANGLADESH USING IMPROVED METHODS OF DETECTION, ISOLATION, AND IDENTIFICATION OF *Vibrio cholerae*	:	21
CHAPTER 3: VIRULENCE GENE PROFILE AND ANTIBIOTIC RESPONSES OF CLINICAL AND ENVIRONMENTAL *Vibrio cholerae* ISOLATED FROM COASTAL ENDEMIC FOCI OF BAY OF BENGAL	:	51
CHAPTER 4: INFLUENCE OF ENVIRONMENT FACTORS IN OCCURRENCE AND DISTRIBUTION OF *Vibrio cholerae*	:	73
CHAPTER 5: *Vibrio cholerae* SEASONALITY AS A FUNCTION OF COMPLEX CHANGES WITHING MICROBIAL COMMUNITIES OF THE AQUATIC ENVIRONMENT.	:	94
CHAPTER 6: COMPARATIVE GENOMICS OF CLINICAL AND ENVIRONMENTAL *Vibrio mimicus*	:	122
CHAPTER 7: DEEP SEA BACTERIA RELATED TO HUMAN PATHOGENIC *VIBRIO* SPECIES	:	153
CHAPTER 8: GENOMIC DIVERSITY OF 2010 HAITINA CHOLERA OUTBREAK STRAINS	:	185
CHAPTER 9: CONCLUSION	:	220
LIST OF PUBLICATIONS	:	226
GRANT SUPPORT	:	229

Dhaka University Institutional Repository

LIST OF TABLES

TABLES	LEGENDS		PAGE NO.
Table 2.1.1:	Primer sequences for the genes, *V. cholerae* O1 *rfb*, *V. cholerae* O139 *rfb*, and *ctxA*.	:	27
Table 2.1.2:	Occurrence of *V. cholerae* O1, *V. cholerae* O139, and cholera toxin gene, *ctx*A each of the fractionated samples.	:	29
Table 2.1.3:	Seasonal detection of *V. cholerae* O1, *V. cholerae* O139, and cholera toxin gene, *ctx*A in Bakergonj and Mathbaria, Bangladesh.	:	32
Table 3.1:	*V. cholerae* O1, O139 and non-O1/O139 isolated from clinical and environmental samples collected in Bakerganj and Mathbaria between March 2004 and January 2006.	:	54
Table 3.2:	Oligonucleotide primers, sequences, and amplicons used in PCR assays	:	56
Table 3.3:	Virulence gene profiles of *V. cholerae* isolates examined in this study.	:	58
Table: 3.4:	Antibiotic resistance of *V. cholerae* O1 biotype El Tor serotype Inaba and Ogawa isolated from environmental and clinical samples between March 2004 to May 2006 from Mathbaria and Bakergonj, in Bangladesh.	:	61
Table 3.5:	Cumulative presentation of resistant pattern, plasmid analysis, and genetic screening of *V. cholerae* O1 and O139 isolated from environmental and clinical samples from two coastal foci of endemic cholera, in Bangladesh.	:	65
Table 4.1:	Descriptive statistics of the environmental variables (monthly average) and *V. cholerae* isolation (average number of samples positive/month) used for statistical correlation.	:	82

Dhaka University Institutional Repository

Table 5.1:	Relative abundance of some well-characterized bacterial genera and species frequently identified in this study by 16S rRNA metagenomics. Altogether, a total of 1600 bacterial genera and 3598 species were identified through out the study.	:	102
Table 6.1:	General features of the *Vibrio mimicus* genome compared with other Vibrio spp.	:	128
Table 6.2:	Tandem repeats in *V. mimicus* MB451 identified using the tandem repeats finder program.	:	137
Table 7.1:	General features of the *Vibrio antiquarius* EX25 genome	:	163
Table 7.2:	Average nucleotide identity (ANI) of conserved genes between V. antiquarius and *V. alginolyticus, V. parahaemolyticus*, *V. harveyii*, and *V. vulnificus*.	:	165
Table 7.3:	Predicted biology of *V. antiquarius* EX25	:	168
Table 8.1:	List of strains used in this study. C; Clinical, E; Environmental S; Sewage	:	191
Table 8.2:	VNTR genotypes of a subset of *V. cholerae* strains isolated from Haiti and other geographic locations. Allele number indicates the number of repeat units of each locus.	:	195
Table 8.3:	Occurrence and distribution of mobile genetic elements (MGE's) in Haitian Strains.	:	197-98
Table 8.4:	Sites of nucleotide variation in CTX prophages and RS1 elements.	:	201
Table 8.5:	Distributions and genomic variations of other known pathogenicity islands in outbreak and reference isolates.	:	203-04

Dhaka University Institutional Repository

LIST OF FIGURES

FIGURES	PAGE NO.
Figure 2.1.1: Temporal dynamics of *V. cholerae* isolation (O1, O139, and non-O1/O139) in two coastal regions of endemic cholera	28
Figure 2.1.2: Temporal dynamics of *V. cholerae* O1, O139, and cholera toxin gene (*ctx*A) by direct multiplex PCR	31
Figure 2.1.3: Seasonal occurrence of cholerae cases in Bakergonj and Mathbaria collected through hospital surveillance	33
Figure 2.1.4: Seasonality of detection of *V. cholerae* O1, O139, and cholera toxin gene (*ctxA*) in Bakergonj (a) and Mathbaria (b)	34
Figure 2.2.1: Representative RNA colony blots hybridized with probe Vchomim1276	42
Figure 2.2.3: Detection and enumeration of total *V. cholerae* associated with plankton free water, Phytoplankton and Zooplankton	45
Figure 2.2.4: Enumeration of total bacterioplankton associated with plankton free Water, Phytoplankton and Zooplankton samples as measured by Acridine Orange Direct Count (AODC) and DAPI (4',6' diamino-2-phenylindole-2HCl) staining	45
Figure 2.2.5: Comparison of total bacterioplankton count as opposed to total culturable heterotrophic bacterial count associated with plankton free Water, Phytoplankton and Zooplankton.	46
Figure 3.1: Resistance to Tetracycline and Erythromycin linked to the resurgence of *V. cholerae* O1 biotype Ogawa El Tor replacing Inaba El Tor	64
Figure 3.2: Pulsed Field Gel Electrophoresis (PFGE) of NotI-digested genomic DNA revealed the strains to be of similar clonal lineage	67
Figure 4.1: Statistical correlations of the environmental variables with isolation of toxigenic and non-toxigenic *V. cholerae*	83
Figure 4.2: Temporal dynamics of Conductivity and Total Dissolved Solids (TDS) and their correlation with observed Cholera and Diarrheal cases in Mathbaria and Bakerganj, Bangladesh	84
Figure 4.3: Temporal dynamics of air and water temperature and their correlation with observed Cholera and Diarrheal cases in Mathbaria and Bakerganj, Bangladesh	85

Dhaka University Institutional Repository

Figure 4.4: Temporal dynamics of water depth correlated with the observed cholera cases in Mathbaria and Bakerganj, Bangladesh 85

Figure 4.5: Correlation of water salinity with observed cholera cases in Mathbaria and Bakerganj, Bangladesh 86

Figure 4.6: Correlation of water pH with observed cholera cases in Mathbaria and Bakerganj, Bangladesh 86

Figure 4.7: Correlation of cholera cases and environmental isolation of toxigenic *V. cholerae* with the metagenomic identification of microbial species over time in Mathbaria, Bangladesh 89

Figure 4.8: Temporal distribution of numbers of species identified by 16S metagenomics correlated with certain environmental variables 90

Figure 5.1: Global population diversity and overall community composition of coastal ecosystem of Mathbaria 100

Figure 5.2: Relative abundance of major microbial phyla (A) and classes (B) in coastal ecosystem of Mathbaria, Bangladesh 101

Figure 5.3: Variation in bacterial community composition across different habitats (site 1 to 6) of Mathbaria, Bangladesh 103

Figure 5.4: Characteristics microbiota of two distinct habitats (site 1 and 6) over time 104

Figure 5.5: Rank abundance curves computed at different taxonomic levels (Class to species) using the accumulated frequencies 108

Figure 5.6: Species richness accumulation curves computed at different taxonomic levels (Class to species) to demonstrate dynamics of species richness over accumulated sampling rounds 109

Figure 5.7: Shannon index plotted with recorded cholera cases in this study 110

Figure 5.8: Temporal diversity of Shannon index plotted along with environmental *V. cholerae* detection and recorded cholera cases 111

Figure 5.9: Plot of Vibrio sp. values against the number of cholera cases observed in Mathbaria, Bangladesh 112

Figure 5.10: Temporal variation of beta diversity with the occurrence of cholera cases and environmental isolation of toxigenic and non-toxigenic *V. cholerae* in Mathbaria, Bangladesh 113

Figure 5.11: Pairwise comparison between alpha, beta and gamma indexes at different taxonomic levels 114

Dhaka University Institutional Repository

Figure 5.12: Clustering of cholera cases with alpha, beta and gamma diversity indexes 115

Figure 5.13: Variation in microbial community composition during and in between seasonal cholerae epidemics in Mathbaria, Bangladesh 116

Figure 5.14: Relative abundance of major species detective during and in between seasonal cholera epidemics in Mathbaria, Bangladesh. 117

Figure 6.1: Circular maps of the chromosomes and plasmids of *V. mimicus* MB451 genome 129

Figure 6.2: Distribution of different subsystem features in Vibrios by RAST subsystem category distribution 130

Figure 6.3: Genome comparison of *V. mimicus* MB451 with reference genomes 132

Figure 6.4: Whole genome comparison of *Vibrio mimicus* by using MUMmer 133

Figure 6.5: Linear pairwise comparison between chromosomes of *Vibrio mimicus*: evidence of intra chromosomal rearrangement. 134

Figure 6.6: Distribution of genomic islands (GI's) in the large chromosome (C-I) of *V mimicus* MB451

Figure 6.7: Linear pairwise comparison of the *Vibrio mimicus* genome 139

Figure 6.8: Pathogenicity Islands found in the genome of *Vibrio mimicus* 140

Figure 6.9: Neighbor-joining tree based on alignment of homologous sequences of 75 conserved ORFs 144

Figure 6.10: Neighbor-joining tree based on alignment of homologous sequences of 925 conserved ORFs 145

Figure 6.11: Evolutionary Distance Tree based on the divergence of the conserved ORFs among *Vibrio* spp. 146

Figure 7.1: Evolutionary history was inferred using the Neighbor-Joining method employing 16S rDNA sequences obtained from the isolates from the East Pacific Rise (EPR) and other *Vibrio* species 160

Figure 7.2: Evolutionary history was inferred using the Neighbor-Joining method employing (A) 23S rDNA and (B) ToxR sequences obtained from the isolates from the East Pacific Rise (EPR) and other *Vibrio* species 161

Figure 7.3: Circular maps of the chromosomes of Vibrio antiquarius EX25 genome generated by CGView 164

Figure 7.4: Genome comparison of *Vibrio antiquarius* EX25 with other Vibrio genomes 166

Dhaka University Institutional Repository

Figure 7.5: Venn diagrams showing number of unique and shared orthologs between *Vibrio antiquarius* EX25 and other *Vibrio* genomes. 167

Figure 7.6. Schematic representation of a) Pre-CTXphi and b) VPI-2 like element identified in the genome of *Vibrio antiquarius* EX25. 170

Figure 7.8: Graphical view of the distribution of genomic islands (GI's) in a) the large chromosome (C-I); and b) the small chromosome (C-II) of *Vibrio antiquarius* EX25 171

Figure 7. 9: Schematic representation of the putative genomic islands identified in *Vibrio antiquarius* EX25 172

Figure 7.10: Linear pairwise comparison of *Vibrio antiquarius* EX25 genome 173

Figure 7.11: Chromosome wise alignment of *Vibrio antiquarius* EX25 genome with *V. parahaemolyticus* and *V harveyi* genomes by MUMmer demonstrates evidence of genomic inversion and rearrangement. 175

Figure 7.12: Phylogeny of EX25 using protein sequences of 31 housekeeping genes 176

Figure 8.1: Source and distribution of isolates collected during the 2010 Haitian cholera outbreak 193

Figure 8.2: Minimum spanning tree based on MLVA clustered *V. cholerae* O1 strains into three distinct groups 196

Figure 8.3: Neighbor-joining trees showing phylogenetic relationships based on 992 orthologous protein-coding genes 208

Figure 8.4: The two-dimensional PCA projection plots based on divergence of average nucleotide identity and tetra nucleotide frequency 209

Figure 8.5: Polymorphisms in Haitian outbreak and reference isolates as determined by whole genome mapping 210

Figure 8.6: Whole genome map similarity cluster of Haitian outbreak isolates with reference strains using UPGMA 211

Figure 8.7: SNP-discovery on chromosome I in *V. cholerae* O1 213

Figure 8.8: SNP-discovery on chromosome II in *V. cholerae* O1 214

Figure 8.9: SNP-based phylogenetic tree for *V. cholerae* O1, constructed using 277 intra- and intergenic SNPs identified in the *V. cholerae* O1 population 215

Dhaka University Institutional Repository

CHAPTER 1:

GENERAL INTRODUCTION

Dhaka University Institutional Repository

Infectious diseases have been the major health problems confronting humankind for as long as diseases have been documented or depicted in historical records, literature, art and scientific writing. The problem of infectious diseases is never static, as microorganisms co-exist with people often sharing a common environment. A number of important factors including global travel and trade, the growth of technology, adaptation of micro-organisms producing strains that are more virulent and resistant to treatment, increases in the numbers of people with weakened immune system, changes in environmental parameters and land use create change in this delicate balance and some increase the risk to human health from infectious diseases. As parts of the world become both politically and ecologically unstable, waterborne infectious diseases become a greater concern because the best prevention mechanism supplying the most vulnerable populations with clean drinking water becomes increasingly difficult [1, 2]. Cholera is an ancient waterborne disease that has been categorized as one of the "emerging and reemerging infections" [3] threatening many countries worldwide in the form of epidemics and pandemics. Estimates have suggested that cholera can affect 3–5 million people and cause 100,000–130,000 deaths a year as of 2010, although the reported number of cases is much lower. It is often easily treatable if adequate fluids are administered within a short time of the onset of symptoms, and vaccines against the disease are available, but vaccines are not fully protective, and their effectiveness is limited to time scales of months to a few years [4]. The biological and environmental factors that contribute to the dynamics of cholera epidemics continue to be the subject of intense study. A distinctive epidemiological feature of cholera is its appearance in a regular seasonal pattern in areas of endemic infection and in explosive outbreaks often starting in several distinct foci simultaneously [5, 6], indicating a possible role of environmental factors in triggering the epidemic process [7].

Cholera is transmitted by water carrying high concentrations of *Vibrio cholerae*, a gram-negative bacterium which has both human and environmental stages in its life cycle. The bacterium is an autochthonous member of the aquatic microbial community and its extensive genetic diversity and propensity for homologous recombination ensure that new variants will emerge to evade immune defenses. Therefore understanding the nature of genetic diversity and the means by which *V. cholerae* survives in the environment is critical to understand the parameters that

Dhaka University Institutional Repository

allow new pathogenic strains to emerge and to predict the onset of conditions favoring *V. cholerae* proliferation, thereby preventing or limiting the extent of future pandemics.

V. cholerae is conventionally classified into serogroups and biotypes. Up to the present time, more than 206 distinct serogroups have been identified according to the antigenic property of cell surface polysaccharides. Two serogroups, designated O1 and O139, are considered to be responsible for most of the epidemic and endemic cholera. *V. cholerae* of other serogroups are collectively called non-O1/O139 that may also cause localized outbreaks of cholera. The O1 serogroup is further divided into three sub-serovars: Ogawa; Inaba; and Hikojima based on antigenic properties.*V. cholerae* O1 strains are classified into two biotypes: 'classical' or 'El Tor' based on phenotypic differences like haemolysis of sheep erythrocytes, agglutination of chicken erythrocytes, Voges–Proskauer reaction, and sensitivity to polymyxin B and specific phages [8]. Additionally, differences in single genes including *tcpA*(toxin co-regulated pilinA), *ctxB* (cholera toxin subunit B), *rstR* (repeat sequence transcriptional regulator) and variations in certain gene clusters such as VSP-I and -II (*Vibrio* seventh pandemic island I and II), and RTX (repeat in toxin) [9] have also been applied to differentiate the two biotypes. Two biotypes differ also in their pathogenic potentials, survivability and infection patterns in their human hosts [6] as El Tor strains are associated more frequently with asymptomatic infections and fewer fatalities, better survival in the environment and in the human host, and more efficient host-to-host transmission compared with classical strains[10]. By contrast, the clinical manifestation of the disease caused by the classical biotype is more severe. Seven cholera pandemics are believed to have occurred since 1817. Many variants of pathogenic *V. cholerae* have emerged over the course of recent pandemics but only a few have significantly affected the epidemiology of the disease. The classical biotype strains that caused the sixth pandemic (1899–1923) are the first known pandemic biotype. Although El Tor strains existed and were associated with sporadic cholera during the sixth pandemic [11], it was not until 1961 that this biotype initiated the seventh pandemic which is still ongoing and causing explosive endemic and epidemic cholera outbreaks around the world. In 1992, a new pathogenic serogroup other than O1, designated as O139, caused cholera epidemics in India and Bangladesh. Strains belonging to this serogroup subsequently spread to other Asian countries, prompting

Dhaka University Institutional Repository

concerns that the emergence of O139 represented the beginning of the eighth pandemic. This serogroup is believed to have originated from the El Tor biotype, with major changes in the O-antigen biosynthetic genes [12].

Different molecular tools have been used to fingerprint and track toxigenic strains of *V. cholerae* demonstrating that O1 and O139 isolates group closely with one another, forming a tight cluster, while environmental non-O1/O139 isolates are generally diverse [13-15]. Moreover, detailed investigation of the *V. cholerae* genome and of specific housekeeping genes revealed that recombination may be the major source of this variability, a common feature in the evolution of *V. cholerae*. Furthermore, the three main pathogenic clones of *V. cholerae* O1 sixth pandemic [classical], seventh pandemic [El Tor], and U.S. Gulf Coast isolates appear to have evolved independently [14, 15]. Evidence indicates that toxigenic strains may arise from environmental, nontoxigenic progenitors in coastal areas by acquiring prophage carrying genes for cholera toxin [15, 16]. *V. cholerae* has been isolated routinely from many aquatic environments throughout the world, often in association with plankton, plants and aquatic animals [17, 18]. The prevalence of *V. cholerae* in the environment is influenced by temperature and salinity [18] as well as concentrations of dissolved organic carbon [19]. A number of genes have been previously implicated in environmental survival, and this has already led to a better understanding of the genetic basis for *Vibrio cholerae* adaptations including genes important for biofilm formation [20], zooplankton association [21], survival with filamentous blue cyanobacteria [22]. These genes clearly provide *Vibrio cholerae* with mechanisms to avoid environmental stresses and obtain nutrients in aquatic environments. Horizontal gene transfer has played an important role in the emergence of *V cholerae* strains with altered human virulence and evolution of pandemics mediating acquisition of large islands of genes encoding pathogenicity and survival functions, including the *Vibrio* pathogenicity island [23] and the cholera toxin phage genome [24], the superintegron, the *SXT* conjugative integrating element and O-antigen genes [7].

Ecology of *Vibrio cholerae*:

Despite the prevalence of the causative organism worldwide, only certain regions of the world, mainly in the tropics and subtropics, maintain endemicity of the

Dhaka University Institutional Repository

disease cholera. The nature of this phenomenon is most likely related to both environmental and socioeconomic factors. Following the sixth pandemic (1950), areas with endemic cholera had been reduced to southeastern India and Bangladesh, a significantly narrower geographic region [25]. During the current seventh pandemic, the range of endemicity has expanded and includes areas of Africa and South and Central America. Where endemic, cholera tends to demonstrate distinct seasonal trends strongly related to the ecology of *V. cholerae* in the environment, where high numbers are observed during times of warm water temperatures and zooplankton blooms [26]. In Bangladesh, there is a bimodal distribution in the frequency of cases over an annual cycle where onset of epidemics coincides with dry weather and the warmest water temperatures of the year (August or September) and the frequency of cases is reduced to background levels of a few cases as the temperature decreases in winter [6, 25]. In Calcutta, India, cholera cases tend to peak in April, May, and June. In South America, where the disease has become endemic since its reemergence in 1991, cases are concentrated in the austral summer months (January and February) [6]. The geographic limits of endemicity are also consistent and suggest a biogeographical pattern of endemic disease that follows the environmental habitat of the causative organism. Although inland countries report cases of cholera (WHO) and freshwater plankton species harbor *Vibrios*, case frequency and location of regions of endemicity increase near coastlines. Along the Indian subcontinent, there is a distinct boundary approximately 150 miles inland and to the northern part of the peninsula where the frequency of cholera cases rapidly diminishes [11, 25]. In South America, outbreaks start along the coast, and epidemiological studies in Mexico demonstrate that people living in coastal states are at a high risk of contracting the disease [27]. *Vibrio cholerae* clones often show distinct regional signatures indicating effect of environmental and geographical factors modulating micro evolution of the bacterium [28, 29]. *V. cholerae* O1 is also endemic in regions where there are now only sporadic cases of cholera. Such endemic clones are found along the U.S. Gulf Coast and in Australia and are unique, although closely related, based on phage typing and gene sequence analyses [14, 30].

Dhaka University Institutional Repository

Environmental survival and persistence:

The primary reservoir for *V. cholerae* is the aquatic and estuarine environment. Thus, contaminated water and food serve as the primary vehicles for transmitting the bacterium to humans [1, 31]. Although *V. cholerae* is known to adapt and persist in a variety of aquatic niches, the basis for its persistence in those ecological niches has not yet been elucidated. In endemic areas, such as Bangladesh and Peru, an epidemic occurs seasonally and *V. cholerae* appears during the epidemics and disappears during inter-epidemic periods [32, 33]. The underlying basis for the appearance and disappearance of *V. cholerae* during and after cholera epidemics is not clear. Until the early 1980s, *V. cholerae* was believed to be human host-adapted and incapable of surviving for more than a few hours in natural waters. Consequently, cholera epidemics were believed to be impossible if contamination of the water by human feces did not occur [34]. However, it is now well established that *V. cholerae* can persist indefinitely in aquatic environments, although the basis for this survival, particularly during inter-epidemic periods in cholera endemic areas, is not well understood. Franco *et al.* [34] have reported that CT- positive strains of *V. cholerae* (i) appeared in the aquatic environment 2-3 months before the onset of cholera epidemics in Peru, which suggests amplification of *V. cholerae* in water in response to an as-yet-unknown environmental stimulus, and (ii) were not detected during the inter-epidemic period (winter season), which indicates that epidemic strains of *V. cholerae* survive during the inter-epidemic period by some as-yet-unknown mechanism.

Colwell *et al.* [34] have demonstrated that *V. cholerae* can enter a viable but non-culturable (VBNC) state in response to adverse environmental conditions (e.g., nutrient deprivation, elevated salinity, and/or reduced temperature), while retaining full virulence [34-37]. However, attempts to resuscitate VBNC *V. cholerae* have been unsuccessful so far. Furthermore, the importance of the VBNC state in the epidemiology of cholera remains unclear. In studies investigating the environmental survival mechanisms of *V. cholerae*, Colwell and other investigators [1, 33, 38] also reported that *V. cholerae* binds to phytoplankton and zooplankton in the environment, and that cholera epidemics follow plankton blooms, suggesting some association with

Dhaka University Institutional Repository

plankton is involved in the endemicity of *V. cholerae*. Although epidemic *V. cholerae* strains are known to persist for many years in the marine environment while regularly causing significant cholera epidemics [39-41], the specific mechanism underlying their adaptation to long-term persistence is still unknown at the present time. Several studies have proposed that the persistence of *V. cholerae* in the environment is due to a viable but nonculturable (VBNC) state elicited by starvation and cold [35, 37, 42, 43], however, it is not known whether the VBNC phenotype promotes persistence in endemic regions and under conditions other than starvation and cold. Alternatively, another hypothesis is that production of EPS and the rugose phenotype might promote the survival and persistence of *V. cholerae* in aquatic environments.

Factors Influencing Epidemic Cholera:

The persistence and geographical expansion of cholera outbreaks are concordant with the suspicion that spread of a well-known infectious disease may intensify with climate change [44]. Ecological studies strongly suggest that cholera will intensify in the future with more cases and possibly wider and more frequent outbreaks since global climate warming has been predicted [45-47]. The level of risk for populations in tropical areas being exposed to cholera-causing bacteria is predicted to increase, based on the fact of enhanced survival and growth of cholera-causing bacteria in warmer temperatures [48], and shifts in the availability and use of water resources with sea-level rise and increased brackish water. Predicted global warming also suggests that human populations in subtropical or template regions will be newly exposed to the risk of cholera by transformation of their environment to one that will be favorable for proliferation of *V. cholerae*. Therefore, the increasing threat of cholera for human populations globally is a genuine human health risk and a problem that must be solved, with respect to prediction and prevention, based on improved understanding of the mechanisms of disease outbreaks.

Temperature, salinity, rainfall and plankton have proven to be important factors in the ecology of *V. cholerae*, influencing the transmission of the disease in those regions of the world where the human population relies on untreated water as a source of drinking water. Moreover, The results of environmental sampling and

Dhaka University Institutional Repository

analysis of the environmental and clinical data revealed significant correlations of water temperature, water depth, rainfall, conductivity, and copepod counts with the occurrence of *V. cholerae* [49]. In consequence to various external factors, the abundance and diversity of the *Vibrio* population is known to exhibit a complex seasonal distribution driven by fluctuations in physiochemical parameters including temperature and salinity and shifts in *Vibrio* reservoirs [50]. It is recognized that *V. cholerae* is a component of coastal and estuarine microbial ecosystems, with the copepod species of zooplankton that comprise the aquatic fauna of rivers, bays, estuaries and the open ocean serving as host for the bacterium [6, 51]. *V. cholerae* can be found attached to the carapace and in the gut of copepods in large numbers, the copepod essentially serving as a vector for this human pathogen [1, 52]. A single copepod, for example, can contain as many as 10^3–10^5 *V. cholerae* cells [53]. Because a concentration of 10^9 ml^{-1} *V. cholerae* comprises an infective dose, ingestion of untreated water containing a relatively small number of copepods carrying *V. cholerae* can initiate the disease [54]. Therefore, conditions favorable for multiplication of copepods and related chitinous zooplankton species for which *V. cholerae* is commensal or symbiotic will result in an increase in the number of *V. cholerae* to influence the seasonal shaping of cholera epidemics. Other than these, precipitation sea surface temperature and chlorophyll were also found to be crucial for maintaining the seasonal cycle of cholera epidemics [55].

Evolutionary complexity contributed by biotype switching, serotype conversion, and genetic reassortment:

For centuries cholera has been one of the most feared diseases and *Vibrio cholerae*, a water-borne pathogen is the causative agent of this severe watery diarrheal disease. It is the only enteric pathogen that has the potential to produce pandemic disease and, thus, is of immense public health importance. On the basis of variable somatic O antigen composition, >200 serogroups of *V. cholerae* have been recognized, of them toxigenic strains of Vibrio cholerae belong to the O1 or O139 serogroup. *V. cholerae* O1 is further classified into two biotypes, classical and El Tor, and into two major serotypes, Inaba and Ogawa. Classical and El Tor are two well-established biotypes within the *V. cholerae* O1 serogroup, both of which are believed to have evolved from separate lineages [56, 57], distinguished by differences in their

Dhaka University Institutional Repository

biochemical reactions or phenotypic traits [58]. *V. cholerae* O1 strains are known to undergo serotype conversion or switching between the Inaba and Ogawa serotypes [59-61]. Observations of the epidemic that broke out in Latin America in 1991 have supported this notion. In addition, epidemic strains undergoing a serotype conversion several months prior to the epidemic is also well documented [62]. Since 1817, seven cholera pandemics have occurred around the world. Firm evidence indicates that the fifth and sixth cholera pandemics were caused by the classical biotype, whereas the most extensive and ongoing seventh pandemic is caused by El Tor biotype. Since the onset of El Tor dominance in 1961, classical strains have gradually been replaced by El Tor strains and are now believed to be "extinct" throughout the world, except in Bangladesh where selective distribution of *V cholerae* O1 biotypes, as evidenced by clustering of the classic biotype in the southern region and of the El Tor biotype in all other regions of Bangladesh, has been documented [63]. However, classical biotype strains reemerged as a predominant epidemic strain in parts of Bangladesh in 1982 [64, 65] and coexisted with the El Tor strains, causing disease until 1993 in Bangladesh and until 1997 in Mexico [28].

V. cholerae O1 is known to interconvert between two serotypes, Ogawa and Inaba [59, 61, 66]. The frequency of conversion of Ogawa to Inaba is approximately one in ten thousand [67], whereas conversion from Inaba to Ogawa is rare and apparently is strain dependent [62]. During the El Tor cholera pandemic caused by *V. cholerae* O1, biotype interplay between Inaba and Ogawa was found to be frequent and at least four documented cases of O serogroup conversions have occurred, that is, from non-O1/O139 to O1 (twice), O1 to O139, and O1 to O37 [68]. However, a new epidemic strain, carrying the O139 rather than the O1 antigen, emerged in southern Asia in 1992 [69, 70]. The O139 and El Tor O1 strains continue to cause epidemics of cholera and the Inaba serotype of *V. cholerae* O1 was displaced. Extensive outbreaks were reported in various regions of Bangladesh and India, Pakistan, Nepal, China, Thailand, Kazakhstan, Afghanistan, and Malaysia [71, 72], as well as sporadic cases in the United Kingdom and the United States [73]. The emergence of *V. cholerae* O139 during 1992–1993 was associated, initially, with a total displacement of the *V. cholerae* O1 strains. However, the O1 strains reemerged in 1994 and initiated a series of disappearance and reemergence of either of the two serogroups that was associated with temporal genetic and phenotypic changes sustained by the strains. Since the

Dhaka University Institutional Repository

initial emergence of *V. cholerae* O139, new variants of the pathogen derived from multiple progenitors have been isolated and rapid genetic re-assortment in *V. cholerae* O139 strains occurred along with the changing epidemiology of *V. cholerae* O1 and a rapid shift in antimicrobial resistance [74, 75]. Soon after the *V. cholerae* O139 episode, new varieties of *V. cholerae* O1 (hybrid *V. cholerae*), of the El Tor biotype re-emerged, with attributes of the classical biotype, even when the classical biotype of *V. cholerae* O1 was no longer detected in southern Bangladesh, the last of the niches where this biotype prevailed. These two biotypes (classical and El Tor) have traditionally been distinguished by phenotypic tests and by genetic differences in the major toxin-coregulated pilus (TCP) gene, the *tcpA* allele of the TCP cluster [76], the *rstR* region (regulatory region for phage lysogeny) of CTX phages [77], and the type of cholera toxin (CT) produced. Apart from these characteristics, there are also dissimilarities in the infection patterns of disease caused by the two biotypes. These include the occurrence of more asymptomatic than symptomatic infections caused by El Tor strains, the asymptomatic patients outnumbering active patients by a ratio of up to 50:1 [10], better survival of El Tor strains in the environment and in the human host, and more efficient host-to-host transmission of El Tor strains than classical strains[10]. However, three variants of the El Tor biotype have been described recently: Hybrid (MJ1236)[78], which could not be biotyped because they are a mixture of both classical and El Tor [79]; Mozambique variant (B33) in 2004–2005, which has a typical El Tor genome but a tandem repeat of the classical CTX prophage in the small chromosome [80]; and the altered El Tor type (a typical El Tor biotype and an El Tor CTX prophage that produces CT of the classical type) predominant in Bangladesh since 2001 [81]. Hybrid and altered El Tor vibrios have also been described in other regions of Asia and Africa [82]. From an epidemiological viewpoint, among the three variants discussed here, the altered El Tor *V. cholerae* O1 (CIRS101) is currently the "most successful" in that strains belonging to this type have virtually replaced the prototype El Tor in Asia and many parts of Africa, notably East Africa. The one feature of this genome that stands out, compared to the genomes of *V. cholerae* MJ-1236 and B33, is the absence of putative prophages or prophage-like elements beyond those elements associated with pathogenic clones of *V. cholerae*, i.e., CTXphi, VSP-I and -II, and VPI 1 and 2. The presence or absence of genomic islands such as kappa prophage, GI-12, GI-14, and GI-15, may have ecological significance for *V. cholerae* MJ-1236 and B33, perhaps in the form of

Dhaka University Institutional Repository

reduced fitness[83]. Although how the classical cholera toxin in El Tor strains would influence *V. cholerae* pathogenicity is unclear, but is apparently driven by competition and, thereby, attaining enhanced fitness in the natural environment and in the human gut.

Virulence-associated prophage and genomic islands in the context of genome evolution:

V. cholerae possesses several known virulence factors, of which cholera toxin (CT) and toxin-coregulated pilus (TCP) are considered most significant. Genes coding for CT (*ctx*AB) are part of a temperate filamentous bacteriophage CTXphi [24]that can be incorporated into both chromosomes of *V. cholerae* at specific positions. The CTXphi found in classical and El Tor biotypes are different in their repressor gene sequence, rstR, and classified into CTXphi^{Class} and CTXphi^{ElTor} according to the biotype of the host. Attachment sites of CTXphi have also been known to harbor other genetic elements, including toxin-linked cryptic (TLC), RS1 elements, f237, VSK (=pre-CTX) prophages [68, 78]. The VSK prophage, a CTXphi-like phage devoid of ctxAB, has been found to be widely distributed among both clinical and environmental *V. cholerae* [68, 78]. Among the non-toxigenic strains, *V. cholerae* O39 AM-19226 and O12 1587 have this prophage in their small chromosome. It is evident that the two attachment sites are "home" for various prophages and genetic elements, which serve as an engine deriving genetic diversity in *V. cholerae*. Chitin competence coupled with lytic phage was proposed as a mechanism of such an event in their natural environment, and may be responsible for the emergence of new variants [84]. The genes coding for toxin-coregulated pilus (TCP) are part of a genomic island, VPI-1. Interestingly, VPI-1 is a hot spot for other mobile GIs. Three different GIs, GI-30 for 12129(1), GI-43 for TM11079-80, RC385, and GI-73 (1587, MZO-2, MZO-3), were found instead of VPI-1 in these strains. Most of these newly identified GIs are likely prophages [68]. This cassette-like property of GI mobility was also found for other known pathogenicity islands, including VPI-2 and VSP-2.

Extensive lateral gene transfer in the evolution of *V. cholerae*:

A recent study [68] demonstrated that lateral gene transfer plays an important role in the evolution of *V. cholerae*. A total of 95 genomic islands (GIs) were identified in 23 sequenced *V. cholerae* genomes including GIs related to the O antigen biosynthesis, CTXphi, VPI-1, 2 and VSP-II, a total of 13 (eight in the large chromosome and five in the small chromosome) genomic regions were found to have a cassette-like property, in which different GIs occupy the same or similar region. Most GIs were singletons, though six were found in multiple copies within a given genome. Therefore, the genetic diversity of the species *V. cholerae* clearly is derived mainly from lateral gene transfer, of which several of the laterally transferred genes show a cassette-like property. The mobile nature of O sero-phenotype in *V. cholerae* is one important event in the evoulition of *Vibrio cholerae*. Using *in vivo* and *in vitro* experiments, Schoolnik and colleagues demonstrated serogroup conversion of the *V. cholerae* O1 recipient by the *V. cholerae* O139 donor within a biofilm on the same chitin surface immersed in seawater, which is the most likely place where such a transfer of genomic islands among *V. cholerae* will occur in nature [85, 86]. The finding of a *V. cholerae* strain 12129(1) that has a non-O1 genomic backbone but also the O1 biosynthesis gene cluster clearly demonstrated that the opposite event, that is transformation of non-O1 to O1, is also possible in natural environments. On the basis of evidence compiled by several previous studies and the comparative genomics for 21 strains, O serogrouping as well as other phenotypes and virulence factors, such as cholera toxin, are coded by transferable genetic elements, and, indeed, are mobile in nature. Several non-genomics based studies have suggested that such conversions take place in nature [85, 87, 88]. Recently, chitin-induced natural transformation has been proposed as a lateral gene transfer mechanism for *V. cholerae* [86] and *V. cholerae* O1 to O139 serogroup conversion by a single-step exchange of large fragments of DNA was demonstrated in a microcosm experiment [85]. These findings clearly support our observation that O serogroup conversion frequently occurs in nature.

It is clear from this overall picture that *V. cholerae* rapidly evolves using lateral gene transfer as major weaponry. "Guns", in this case the genetic elements coding for toxins, antibiotic resistance, virulence factors and a wide range of proteins with as yet unknown function, are stored in a natural genetic reservoir consisting of *V. cholerae*, or even other species such as *Vibrio parahaemolyticus*. These genetic

Dhaka University Institutional Repository

elements can be readily transferred and are capable of behaving in a cassette-like manner that can be exchanged at the same chromosomal locus, with the cholera toxin phage and O antigenic determinant biosynthesis gene clusters being the most important. The genetic makeup of the pandemic clones is far from stable, as we are now witnessing an emergence of new variants with different genetic characteristics on different continents. Moreover, we now know that the current (seventh) pandemic is caused by strains belonging to a single phyletic line, diversified mainly via lateral gene transfer from a *Vibrio* gene reservoir in the natural environment. On the other hand, *V. cholerae* non-O1/O139 strains have occasionally been isolated from humans with gastroenteritis, extraintestinal infections, and cholera-like illness and these vibrios are clearly indigenous in both epidemic and non-epidemic estuarine environments throughout the world. In summary, acquisition of the major *V. cholerae* virulence markers (CTXΦ, VPI, and O-antigen biosynthesis regions) in *V. cholerae* non-O1/O139 strains by horizontal gene transfer is well known. In addition, chitin induces natural competence in *V. cholerae*, thereby enhancing the rapid evolution of *V. cholerae*.

Metagenomics

Members of the *Vibrio* genus often comprise <1% of the total marine bacterioplankton; however, *Vibrio* species are among the most abundant culturable bacteria in estuarine and marine waters throughout the world [89]. Various studies were carried out to investigate how the changes in composition of the natural community contribute to the seasonality of the *Vibrio* population, either synergistically with or independent of different physical parameters including particulate organic matter (POM), phytoplankton, zooplankton, finfish and shellfish [51, 90-93]. As heterotrophic bacteria, the association with and degradation of chitin is an important aspect of *Vibrio* ecology and of nutrient cycling in the marine environment [94]. Thus, *Vibrio* are believed to play an important role in the mineralization of the chitinous exoskeletons of zooplankton and are known to be closely associated with copepods and other higher plankton [1, 95]. In consequence to these external factors, the abundance and diversity of the *Vibrio* population is known to exhibit a complex seasonal distribution driven by fluctuations in physiochemical parameters including temperature and salinity and shifts in *Vibrio* reservoirs [89].

Dhaka University Institutional Repository

Ideally, algal species and plankton taxa represent unique microhabitats for microbial colonization and as such, previous studies have shown that some algal species can harbor distinct microbial assemblages [96]. Beside these, competition for substrate and antagonistic interactions with other microbiota has been shown to limit the growth of some bacteria species [97]. Thus, bacterial-bacterial antagonism is one mechanism that contributes to the control and regulation of *V. cholerae* in the environment. Since the human infectious dose of pathogenic *Vibrio cholerae* has a particular range, concentration of the pathogen in the environment is often a necessary step in primary disease transmission [98] and concentration in the environment can occur when a biotic or abiotic source of nutrients enriches the pathogenic *Vibrio cholerae* at higher concentrations than the surrounding water column. Thus studying the total microbial population covering the temporal shift of the predominance and presence of different groups and species is important to predict the possible effect of compositional shifting within the microbial community on evolution of cholera bacteria. Various study outcomes suggest that specific consortia among the commensal population might be crucial responding to changes in the environmental parameters to trigger *V. cholerae* to overcome natural competition to shape epidemics.

Metagenomics is a powerful tool and recently has been introduced for analyses of the natural microbial community, which describes functional and sequence based analysis of the collective microbial genomes contained in an environmental sample. It is one of the most remarkable events in the field of microbial ecology in the past decade. Metagenomics is defined as the direct genetic analysis of genomes contained with an environmental sample. The field initially started with the cloning of environmental DNA, followed by functional expression screening [99], and was then quickly complemented by direct random shotgun sequencing of environmental DNA [100, 101]. Using metagenomics, it is possible to investigate the dynamics and shift in microbial community that can serve as an indicator for forthcoming cholera epidemics or what happens in the community when there is an outbreak going on.

Dhaka University Institutional Repository

References:

1. Colwell, R.R., *Global climate and infectious disease: the cholera paradigm.* Science, 1996. **274**(5295): p. 2025-31.
2. Weiss, R.A. and A.J. McMichael, *Social and environmental risk factors in the emergence of infectious diseases.* Nat Med, 2004. **10**(12 Suppl): p. S70-6.
3. Satcher, D., *Emerging infections: getting ahead of the curve.* Emerg Infect Dis, 1995. **1**(1): p. 1-6.
4. Hill, D.R., L. Ford, and D.G. Lalloo, *Oral cholera vaccines: use in clinical practice.* Lancet Infect Dis, 2006. 6(6): p. 361-73.
5. Glass, R.I., et al., *Endemic cholera in rural Bangladesh, 1966-1980.* Am J Epidemiol, 1982. **116**(6): p. 959-70.
6. Kaper, J.B., J.G. Morris, Jr., and M.M. Levine, *Cholera.* Clin Microbiol Rev, 1995. **8**(1): p. 48-86.
7. Faruque, S.M., et al., *Genetic diversity and virulence potential of environmental Vibrio cholerae population in a cholera-endemic area.* Proc Natl Acad Sci U S A, 2004. **101**(7): p. 2123-8.
8. Safa, A., G.B. Nair, and R.Y. Kong, *Evolution of new variants of Vibrio cholerae O1.* Trends Microbiol, 2010. **18**(1): p. 46-54.
9. O'Shea, Y.A., et al., *Evolutionary genetic analysis of the emergence of epidemic Vibrio cholerae isolates on the basis of comparative nucleotide sequence analysis and multilocus virulence gene profiles.* J Clin Microbiol, 2004. **42**(10): p. 4657-71.
10. Sack, D.A., et al., *Cholera.* Lancet, 2004. **363**(9404): p. 223-33.
11. Pollitzer, R., S. Swaroop, and W. Burrows, *Cholera.* Monogr Ser World Health Organ, 1959. **58**(43): p. 1001-19.
12. Bik, E.M., et al., *Genesis of the novel epidemic Vibrio cholerae O139 strain: evidence for horizontal transfer of genes involved in polysaccharide synthesis.* EMBO J, 1995. **14**(2): p. 209-16.
13. Beltran, P., et al., *Genetic diversity and population structure of Vibrio cholerae.* J Clin Microbiol, 1999. **37**(3): p. 581-90.
14. Byun, R., et al., *Evolutionary relationships of pathogenic clones of Vibrio cholerae by sequence analysis of four housekeeping genes.* Infect Immun, 1999. 67(3): p. 1116-24.
15. Jiang, S.C., et al., *Genetic diversity of Vibrio cholerae in Chesapeake Bay determined by amplified fragment length polymorphism fingerprinting.* Appl Environ Microbiol, 2000. **66**(1): p. 140-7.
16. Chakraborty, S., et al., *Virulence genes in environmental strains of Vibrio cholerae.* Appl Environ Microbiol, 2000. 66(9): p. 4022-8.
17. Islam, M.S., B.S. Drasar, and R.B. Sack, *The aquatic environment as a reservoir of Vibrio cholerae: a review.* J Diarrhoeal Dis Res, 1993. **11**(4): p. 197-206.
18. Lipp, E.K., A. Huq, and R.R. Colwell, *Effects of global climate on infectious disease: the cholera model.* Clin Microbiol Rev, 2002. **15**(4): p. 757-70.
19. Mourino-Perez, R.R., A.Z. Worden, and F. Azam, *Growth of Vibrio cholerae O1 in red tide waters off California.* Appl Environ Microbiol, 2003. **69**(11): p. 6923-31.
20. Watnick, P.I. and R. Kolter, *Steps in the development of a Vibrio cholerae El Tor biofilm.* Mol Microbiol, 1999. **34**(3): p. 586-95.
21. Chiavelli, D.A., J.W. Marsh, and R.K. Taylor, *The mannose-sensitive hemagglutinin of Vibrio cholerae promotes adherence to zooplankton.* Appl Environ Microbiol, 2001. **67**(7): p. 3220-5.
22. Islam, M.S., et al., *Involvement of the hap gene (mucinase) in the survival of Vibrio cholerae O1 in association with the blue-green alga, Anabaena sp.* Can J Microbiol, 2002. **48**(9): p. 793-800.

23. Karaolis, D.K., et al., *A Vibrio cholerae pathogenicity island associated with epidemic and pandemic strains.* Proc Natl Acad Sci U S A, 1998. **95**(6): p. 3134-9.
24. Waldor, M.K. and J.J. Mekalanos, *Lysogenic conversion by a filamentous phage encoding cholera toxin.* Science, 1996. **272**(5270): p. 1910-4.
25. Cockburn, T.A. and J.G. Cassanos, *Epidemiology of endemic cholera.* Public Health Rep, 1960. **75**: p. 791-803.
26. Lobitz, B., et al., *Climate and infectious disease: use of remote sensing for detection of Vibrio cholerae by indirect measurement.* Proc Natl Acad Sci U S A, 2000. **97**(4): p. 1438-43.
27. Lizarraga-Partida, M.L. and M.L. Quilici, *Molecular analyses of Vibrio cholerae O1 clinical strains, including new nontoxigenic variants isolated in Mexico during the Cholera epidemic years between 1991 and 2000.* J Clin Microbiol, 2009. **47**(5): p. 1364-71.
28. Alam, M., et al., *Vibrio cholerae classical biotype strains reveal distinct signatures in Mexico.* J Clin Microbiol, 2012. **50**(7): p. 2212-6.
29. Nusrin, S., et al., *Peruvian Vibrio cholerae O1 El Tor strains possess a distinct region in the Vibrio seventh pandemic island-II that differentiates them from the prototype seventh pandemic El Tor strains.* J Med Microbiol, 2009. **58**(Pt 3): p. 342-54.
30. Karaolis, D.K., et al., *Comparison of Vibrio cholerae pathogenicity islands in sixth and seventh pandemic strains.* Infect Immun, 2001. **69**(3): p. 1947-52.
31. Huq, M.I., R.I. Glass, and B.J. Stoll, *Epidemiology of cholera.* N Engl J Med, 1980. **303**(11): p. 643-4.
32. Franco, A.A., et al., *Cholera in Lima, Peru, correlates with prior isolation of Vibrio cholerae from the environment.* Am J Epidemiol, 1997. **146**(12): p. 1067-75.
33. Huq, A., et al., *Detection of Vibrio cholerae O1 in the aquatic environment by fluorescent-monoclonal antibody and culture methods.* Appl Environ Microbiol, 1990. **56**(8): p. 2370-3.
34. Colwell, R.R., *Vibrios in the environment: viable but non-culturable Vibrio cholerae.* Vibrio cholerae and cholera: molecular to global perspectives, 1994: p. 117-133.
35. Colwell, R., et al., *Viable but non-culturable Vibrio cholerae and related pathogens in the environment: implications for release of genetically engineered microorganisms.* Nature Biotechnology, 1985. **3**(9): p. 817-820.
36. Johnson, J.A., et al., *Vibrio cholerae O139 synonym bengal is closely related to Vibrio cholerae El Tor but has important differences.* Infect Immun, 1994. **62**(5): p. 2108-10.
37. Xu, H.S., et al., *Survival and viability of nonculturableEscherichia coli andVibrio choierae in the estuarine and marine environment.* Microbial Ecology, 1982. 8(4): p. 313-323.
38. Islam, M.S., et al., *Isolation of Vibrio cholerae O139 Bengal from water in Bangladesh.* Lancet, 1993. **342**(8868): p. 430.
39. Barua, D., *The global epidemiology of cholera in recent years.* Proc R Soc Med, 1972. **65**(5): p. 423-8.
40. Goodgame, R.W. and W.B. Greenough, *Cholera in Africa: a message for the West.* Ann Intern Med, 1975. **82**(1): p. 101-6.
41. Ramamurthy, T., et al., *Emergence of novel strain of Vibrio cholerae with epidemic potential in southern and eastern india.* Lancet, 1993. **341**(8846): p. 703-4.
42. Roszak, D.B., D.J. Grimes, and R.R. Colwell, *Viable but nonrecoverable stage of Salmonella enteritidis in aquatic systems.* Can J Microbiol, 1984. **30**(3): p. 334-8.
43. Shiba, T., et al., *Decrease in culturability of Vibrio cholerae caused by glucose.* Appl Environ Microbiol, 1995. **61**(7): p. 2583-8.
44. Haines, A. and J.A. Patz, *Health effects of climate change.* JAMA, 2004. **291**(1): p. 99-103.

Dhaka University Institutional Repository

45. Pascual, M., et al., *Cholera dynamics and El Nino-Southern Oscillation.* Science, 2000. **289**(5485): p. 1766-9.
46. Rodo, X., et al., *ENSO and cholera: a nonstationary link related to climate change?* Proc Natl Acad Sci U S A, 2002. **99**(20): p. 12901-6.
47. Speelmon, E.C., et al., *Cholera incidence and El Nino-related higher ambient temperature.* JAMA, 2000. **283**(23): p. 3072-4.
48. Louis, V.R., et al., *Predictability of Vibrio cholerae in Chesapeake Bay.* Appl Environ Microbiol, 2003. **69**(5): p. 2773-85.
49. Huq, A., et al., *Critical factors influencing the occurrence of Vibrio cholerae in the environment of Bangladesh.* Appl Environ Microbiol, 2005. **71**(8): p. 4645-54.
50. Thompson, F.L., et al., *Genomic diversity of clinical and environmental Vibrio cholerae strains isolated in Brazil between 1991 and 2001 as revealed by fluorescent amplified fragment length polymorphism analysis.* J Clin Microbiol, 2003. **41**(5): p. 1946-50.
51. Huq, A., et al., *Ecological relationships between Vibrio cholerae and planktonic crustacean copepods.* Appl Environ Microbiol, 1983. **45**(1): p. 275-83.
52. Rawlings, T.K., G.M. Ruiz, and R.R. Colwell, *Association of Vibrio cholerae O1 El Tor and O139 Bengal with the Copepods Acartia tonsa and Eurytemora affinis.* Appl Environ Microbiol, 2007. **73**(24): p. 7926-33.
53. Colwell, R.R., et al., *Development and evaluation of a rapid, simple, sensitive, monoclonal antibody-based co-agglutination test for direct detection of Vibrio cholerae 01.* FEMS Microbiol Lett, 1992. **76**(3): p. 215-9.
54. Cash, R.A., et al., *Response of man to infection with Vibrio cholerae. I. Clinical, serologic, and bacteriologic responses to a known inoculum.* J Infect Dis, 1974. **129**(1): p. 45-52.
55. Mendelsohn, J. and T. Dawson, *Climate and cholera in KwaZulu-Natal, South Africa: the role of environmental factors and implications for epidemic preparedness.* Int J Hyg Environ Health, 2008. **211**(1-2): p. 156-62.
56. Kaper, J.B., et al., *Molecular epidemiology of Vibrio cholerae in the U.S. Gulf Coast.* J Clin Microbiol, 1982. **16**(1): p. 129-34.
57. Karaolis, D.K., R. Lan, and P.R. Reeves, *The sixth and seventh cholera pandemics are due to independent clones separately derived from environmental, nontoxigenic, non-O1 Vibrio cholerae.* J Bacteriol, 1995. **177**(11): p. 3191-8.
58. Dziejman, M., et al., *Comparative genomic analysis of Vibrio cholerae: genes that correlate with cholera endemic and pandemic disease.* Proc Natl Acad Sci U S A, 2002. **99**(3): p. 1556-61.
59. Colwell, R.R., et al., *Serogroup conversion of Vibrio cholerae.* Can J Microbiol, 1995. **41**(10): p. 946-50.
60. Fallarino, A., et al., *Identification of additional genes required for O-antigen biosynthesis in Vibrio cholerae O1.* J Bacteriol, 1997. **179**(7): p. 2147-53.
61. Stroeher, U.H., et al., *Serotype conversion in Vibrio cholerae O1.* Proc Natl Acad Sci U S A, 1992. **89**(7): p. 2566-70.
62. Manning, P.A., U.H. Stroeher, and R. Morona, *Molecular basis for O-antigen biosynthesis in Vibrio cholerae O1: Ogawa-Inaba switching.* Vibrio cholerae and cholera: molecular to global perspectives. ASM Press, Washington, DC, 1994: p. 77-94.
63. Siddique, A.K., et al., *Survival of classic cholera in Bangladesh.* Lancet, 1991. **337**(8750): p. 1125-7.
64. Clemens, J.D., et al., *Field trial of oral cholera vaccines in Bangladesh: results of one year of follow-up.* J Infect Dis, 1988. **158**(1): p. 60-9.

Dhaka University Institutional Repository

65. Samadi, A.R., et al., *Classical Vibrio cholerae biotype displaces EL tor in Bangladesh.* Lancet, 1983. **1**(8328): p. 805-7.
66. Gangarosa, E.J., et al., *Multiple serotypes of vibrio cholerae isolated from a case of cholera. Evidence suggesting in-vivo mutation.* Lancet, 1967. **1**(7491): p. 646-8.
67. Bhaskaran, K. and R.H. Gorrill, *A study of antigenic variation in Vibrio cholerae.* J Gen Microbiol, 1957. **16**(3): p. 721-9.
68. Chun, J., et al., *Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic Vibrio cholerae.* Proc Natl Acad Sci U S A, 2009. **106**(36): p. 15442-7.
69. *Large epidemic of cholera-like disease in Bangladesh caused by Vibrio cholerae O139 synonym Bengal. Cholera Working Group, International Centre for Diarrhoeal Diseases Research, Bangladesh.* Lancet, 1993. **342**(8868): p. 387-90.
70. Nair, G.B., et al., *Spread of Vibrio cholerae O139 Bengal in India.* J Infect Dis, 1994. **169**(5): p. 1029-34.
71. Hochhut, B., et al., *Molecular analysis of antibiotic resistance gene clusters in vibrio cholerae O139 and O1 SXT constins.* Antimicrob Agents Chemother, 2001. **45**(11): p. 2991-3000.
72. Faruque, S.M., et al., *Genomic diversity among Vibrio cholerae O139 strains isolated in Bangladesh and India between 1992 and 1998.* FEMS Microbiol Lett, 2000. **184**(2): p. 279-84.
73. Swerdlow, D.L. and A.A. Ries, *Vibrio cholerae non-O1--the eighth pandemic?* Lancet, 1993. **342**(8868): p. 382-3.
74. Faruque, S.M., M.J. Albert, and J.J. Mekalanos, *Epidemiology, genetics, and ecology of toxigenic Vibrio cholerae.* Microbiol Mol Biol Rev, 1998. **62**(4): p. 1301-14.
75. Mukhopadhyay, A.K., et al., *Temporal shifts in traits of Vibrio cholerae strains isolated from hospitalized patients in Calcutta: a 3-year (1993 to 1995) analysis.* J Clin Microbiol, 1996. **34**(10): p. 2537-43.
76. Keasler, S.P. and R.H. Hall, *Detecting and biotyping Vibrio cholerae O1 with multiplex polymerase chain reaction.* Lancet, 1993. **341**(8861): p. 1661.
77. Nusrin, S., et al., *Diverse CTX phages among toxigenic Vibrio cholerae O1 and O139 strains isolated between 1994 and 2002 in an area where cholera is endemic in Bangladesh.* J Clin Microbiol, 2004. **42**(12): p. 5854-6.
78. Nair, G.B., et al., *New variants of Vibrio cholerae O1 biotype El Tor with attributes of the classical biotype from hospitalized patients with acute diarrhea in Bangladesh.* J Clin Microbiol, 2002. **40**(9): p. 3296-9.
79. Safa, A., et al., *Genomic relatedness of the new Matlab variants of Vibrio cholerae O1 to the classical and El Tor biotypes as determined by pulsed-field gel electrophoresis.* J Clin Microbiol, 2005. **43**(3): p. 1401-4.
80. Faruque, S.M., et al., *Genomic analysis of the Mozambique strain of Vibrio cholerae O1 reveals the origin of El Tor strains carrying classical CTX prophage.* Proc Natl Acad Sci U S A, 2007. **104**(12): p. 5151-6.
81. Nair, G.B., et al., *Cholera due to altered El Tor strains of Vibrio cholerae O1 in Bangladesh.* J Clin Microbiol, 2006. **44**(11): p. 4211-3.
82. Safa, A., et al., *Vibrio cholerae O1 hybrid El Tor strains, Asia and Africa.* Emerg Infect Dis, 2008. **14**(6): p. 987-8.
83. Grim, C.J., et al., *Genome sequence of hybrid Vibrio cholerae O1 MJ-1236, B-33, and CIRS101 and comparative genomics with V. cholerae.* J Bacteriol, 2010. **192**(13): p. 3524-33.
84. Udden, S.M., et al., *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 U S A, 2008. **105**(33): p. 11951-6.

85. Blokesch, M. and G.K. Schoolnik, *Serogroup conversion of Vibrio cholerae in aquatic reservoirs.* PLoS Pathog, 2007. **3**(6): p. e81.
86. Meibom, K.L., et al., *Chitin induces natural competence in Vibrio cholerae.* Science, 2005. **310**(5755): p. 1824-7.
87. Bik, E.M., R.D. Gouw, and F.R. Mooi, *DNA fingerprinting of Vibrio cholerae strains with a novel insertion sequence element: a tool to identify epidemic strains.* J Clin Microbiol, 1996. **34**(6): p. 1453-61.
88. Li, M., et al., *Evidence for the emergence of non-O1 and non-O139 Vibrio cholerae strains with pathogenic potential by exchange of O-antigen biosynthesis regions.* Infect Immun, 2002. **70**(5): p. 2441-53.
89. Thompson, F.L., T. Iida, and J. Swings, *Biodiversity of vibrios.* Microbiol Mol Biol Rev, 2004. **68**(3): p. 403-31, table of contents.
90. Tamplin, M.L., et al., *Attachment of Vibrio cholerae serogroup O1 to zooplankton and phytoplankton of Bangladesh waters.* Appl Environ Microbiol, 1990. **56**(6): p. 1977-80.
91. Maugeri, T.L., et al., *Distribution of potentially pathogenic bacteria as free living and plankton associated in a marine coastal zone.* J Appl Microbiol, 2004. **97**(2): p. 354-61.
92. Baffone, W., et al., *Detection of free-living and plankton-bound vibrios in coastal waters of the Adriatic Sea (Italy) and study of their pathogenicity-associated properties.* Environ Microbiol, 2006. **8**(7): p. 1299-305.
93. Turner, J.W., et al., *Plankton composition and environmental factors contribute to Vibrio seasonality.* ISME J, 2009. **3**(9): p. 1082-92.
94. Meibom, K.L., et al., *The Vibrio cholerae chitin utilization program.* Proc Natl Acad Sci U S A, 2004. **101**(8): p. 2524-9.
95. Carman, K.R. and F.C. Dobbs, *Epibiotic microorganisms on copepods and other marine crustaceans.* Microsc Res Tech, 1997. **37**(2): p. 116-35.
96. Grossart, H.P., et al., *Marine diatom species harbour distinct bacterial communities.* Environ Microbiol, 2005. **7**(6): p. 860-73.
97. Long, R.A. and F. Azam, *Antagonistic interactions among marine pelagic bacteria.* Appl Environ Microbiol, 2001. **67**(11): p. 4975-83.
98. Huq, A. and R.R. Colwell, *Environmental factors associated with emergence of disease with special reference to cholera.* Eastern Mediterraneal Health Journal, 1996. **2**: p. 37-45.
99. Handelsman, J., et al., *Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products.* Chem Biol, 1998. 5(10): p. R245-9.
100. Tyson, G.W., et al., *Community structure and metabolism through reconstruction of microbial genomes from the environment.* Nature, 2004. **428**(6978): p. 37-43.
101. Venter, J.C., et al., *Environmental genome shotgun sequencing of the Sargasso Sea.* Science, 2004. **304**(5667): p. 66-74.

Dhaka University Institutional Repository

Vision Statement and Hypothesis:

- Much of the characteristic epidemiologic features of *V. cholerae*, as a rapidly evolving, serial re-emerging pathogen can be attributed to measurable and predictable environmental factors associated with its aquatic habitat.
- *V. cholerae* seasonality is a function of complex changes within microbial communities of the aquatic environment, triggered by specific environmental factors, specific to each cholera-endemic area, to initiate local outbreaks, further developing into an epidemic via fecal contamination of water and subsequent human-to-human transmission.
- The major virulence factors in epidemic *V. cholerae* strains are encoded by clusters of genes that move horizontally or vertically among strains leading to the origination of new pathogenic strains from nonpathogenic progenitors.
- Ecology drives change and exchange of vertically inherited and horizontally transferred gene pools to generate an endemic population structure.

Dhaka University Institutional Repository

CHAPTER 2:

SEASONALITY OF CHOLERA IN THE COASTAL AQUATIC ENVIRONMENT OF BANGLADESH USING IMPROVED METHODS OF DETECTION, ISOLATION, AND IDENTIFICATION OF *VIBRIO CHOLERAE*

Dhaka University Institutional Repository

Cholera, an epidemic and life-threatening diarrhea, is transmitted within households generally by the fecal-oral route. The disease is caused by *Vibrio cholerae*, which is a native of the aquatic environment. Therefore, water plays an important role in transmission and epidemiology of cholera. In the Ganges delta, cholera is endemic, with epidemics occurring twice yearly [1]. The organism has the ability to cause outbreaks of [2] epidemic and pandemic proportions. Therefore, cholera remains a serious public health problem for most developing as well as developed countries, severely affecting their economies. In 1977, Colwell *et al.* [2] first hypothesized that coastal water is an important natural reservoir of *V. cholerae*. Ever since *V. cholerae* has been detected in seawater and other environmental sources around the world, both in areas where cholera is endemic and non-endemic [3-6]. Despite its ubiquity, the ability to determine the presence of pathogenic strains of this bacterial species in the environment with a degree of efficiency has been hindered by the lack of sensitivity of traditional culture methods relied upon for detection. *V. cholerae* can only be isolated from less than 1% of water samples during epidemic periods and is rarely, if ever, isolated by culture in between epidemics [7].

In developing countries, exposure to *V. cholerae* typically occurs via consumption of natural surface water containing *V. cholerae* in sufficient numbers to cause disease, whereas in developed countries exposure to *V. cholerae* often occurs via consumption of raw or undercooked shellfish or travel to regions where cholera outbreaks occur frequently. There is no fully effective vaccine available as of today. Death and illness as a result of eating raw or undercooked seafood, or ingestion or exposure of seawater or natural surface water containing levels of *Vibrio* species sufficient to cause clinical symptoms are a major concern worldwide. However, these bacteria cannot be eradicated from their natural environment nor the optimum environmental conditions for their occurrence, survival, and multiplication been elucidated fully yet. Since these *Vibrio* species naturally occur in marine, estuarine and natural surface water environments, the only way to minimize or prevent infections is to minimize exposure to these bacteria when their concentrations are high. Epidemiological investigation of *V. cholerae*

Dhaka University Institutional Repository

abundance and distribution is, therefore, critical for preventing exposure and disease incidence.

Recent outbreaks of cholera in both developing and developed countries demonstrate the need to define a reliable method for detection of these organisms in their natural environments and any environmentally-derived food sources [8]. Both plankton and sediment play an important role in the occurrence and survival of *V. cholerae* in water and oysters. Sediment and plankton both can be suspended within the water column. Sediment is known to harbor *V. cholerae* (due to its polar electrical charge). Plankton is known to be natural reservoirs of *V. cholerae* and zooplankton blooms facilitate increases in *V. cholerae* density. Furthermore, an infectious dose of *V. cholerae* (6×10^6 cells) [9] can be present on the bodies of a few plankton [10, 11]. Presence of *V. cholerae* in water, sediment, and plankton will also result in the presence of *V. cholerae* in shellfish, as these organisms filter feed all three through their bodies and concentrate high numbers of *V. cholerae*. Traditional culturing methods yield only a small subset of culturable bacteria in the original sample (< 0.1% of total bacteria), and inherently miss those bacteria cells which cannot grow on media but do remain a significant public health threat as they are capable of causing disease [12].

To evaluate the presence of *V. cholerae* regardless of their culturability, a direct detection method should be employed. Direct molecular detection using the polymerase chain reaction (PCR) offers a useful alternative to culture-based methods to evaluate the presence of these organisms in environmental samples. It offers a better understanding of the ecology of *V. cholerae*, by providing a reliable estimates of the occurrence and distribution of both the toxigenic *V. cholerae* O1 and O139, and their non-toxigenic (non-O1/O139) subpopulations. This method is highly sensitive in detecting this organism but do not facilitate isolation of the organism nor does it indicate viability and accurate abundance of the organism.

To this end, microscopy methods, notably fluorescent antibody coupled with direct viable count (FA-DVC), or fluorescence in situ hybridization (FISH), a nucleic

Dhaka University Institutional Repository

acid marker based method of detection are reliable in estimating the accurate abundance of the target organism. FISH can be utilized to detect genomic DNA sequences specific to *V. cholerae* regardless of serogroup, while FA-DVC can be used to detect *V. cholerae* of serogroups O1 and O139, serogroups associated with pandemic cholera. The direct fluorescence antibody (DFA) detection of serogroups O1 and O139 is more reliable than attempting to detect these strains via culture methods as targeting one serogroup in the environment is rather difficult to do, considering the vast serodiversity of *V. cholerae* in the environment. Both methods allow the researcher to quantify the target if no pre-enrichment step is used.

PCR and oligonucleotide hybridization methods have shown promise in detecting and enumerating toxigenic *V. cholerae* but these methods typically rely on direct amplification of, or hybridization to, DNA sequences from environmental samples. These methods however do not facilitate isolation of the organism nor do they assess cell viability. These facts underscore the need for a method that quantifies the total number of viable *V. cholerae* present in a sample. Therefore, RNA colony blot hybridization protocol targeting 16S rDNA with probe Vchomim1276, previously described by Heidelberg [13-15], was developed and employed in this study.

Such coupling of methods can overcome limitations inherent in any one particular method and allow convergence on the precise answer(s). Even though use of multiple methods increase in cost and time commitment, it significantly reduces "not detected" results and enhances the accuracy and statistical power of the results, an extremely important and critical factor in studying the ecology of pathogens in the environment.

Dhaka University Institutional Repository

2.1: Seasonality and Distribution of Toxigenic *Vibrio cholerae*, Monitored by Direct PCR, in Coastal Villages in the Bay of Bengal

Abstract:

Vibrio cholerae is the causative agent of cholera, a severe diarrheal disease, and is indigenous to brackish waters. To advance our understanding of the ecology of this bacterium, we have employed a highly sensitive molecular method for direct detection of *V. cholerae* in two coastal regions of endemic cholera in Bangladesh over a period of 42 months, from March 2004 to September 2007. Toxigenic *V. cholerae* was detected by PCR targeting *V. cholerae* O1*rfb*, *V. cholerae* O139*rfb*, and *ctxA,* in parallel with traditional culture and enrichment. Of 899 samples collected and tested, 26% were positive for *V. cholerae* O1, O139 and/or ctxA in at least one of the three water fractions, particle-free water (<20 μm), 20-64 μm fraction, and >64 μm fraction. Of these, 48% contained *V. cholerae* O1 and 28% contained *V. cholerae* O139; *ctxA* was detected in 76% and 52% of the *V. cholerae* O1 and O139-positive samples respectively whereas 24% of the samples were positive only for *V. cholerae* toxin gene, *ctxA*. Toxigenic *V. cholerae* was more consistently detected from water samples (40%) than from >64 μm fraction (28%) and 20-64 μm fraction (32%) samples in Mathbaria whereas in Bakergonj they were detected from water (37%) and >64 μm fraction (38%) almost equally. Area-wise data analysis revealed that toxigenic *V. cholerae* was present in Mathbaria water throughout the year, while in Bakergonj, it appeared seasonally. Thus, PCR was successfully employed in detecting toxigenic *V. cholerae* directly in water and plankton samples and provides evidence for an environmental reservoir for this pathogen in the coastal water of Bay of Bengal.

Dhaka University Institutional Repository

Introduction:

Extensive studies have shown that *Vibrio cholerae* enters into the viable but nonculturable (VBNC) state when conditions are not conducive to growth and reproduction [16, 17] that is, the cells retain metabolic activity but fail to grow on conventional culture media [16, 18, 19]. This can result in significant underestimation of the total *V. cholerae* population [16]. Techniques employing microscopy can provide data on the occurrence of viable VBNC but it is obvious that the labeled antibody approach to detect all ~200 serogroups of *V. cholerae* is not feasible [20, 21]. Furthermore, *V. cholerae* non-O1/O139 strains can acquire toxin genes in the environment via lateral gene transfer. Currently, neither traditional culture methods nor the direct fluorescent-antibody assay (DFA) can detect the presence of cholera toxin directly in the field. While most environmental *V. cholerae* strains lack the genes required to produce cholera toxin [22], the possibility of genetic exchange in the environment and the potential emergence of new toxigenic clones highlight the importance of including *ctxA* in environmental screening.

PCR and other molecular detection methods offer a useful alternative to culture-based methods and microscopy, especially for environmental samples. We report here a PCR method for direct detection of *V. cholerae* environmental samples. Seawater and plankton samples were collected along the coast of Bangladesh for analysis. The objective of the research was to develop an understanding of the ecology of *V. cholerae*, as well as the occurrence and distribution of the toxigenic *V. cholerae* O1 and O139 subpopulations.

Methods:

Water and plankton samples were collected over a period of 42 months from March 2004 to September 2007 along two coastal regions of Mathbaria (biweekly, 6 ponds, 531 samples) and Bakergonj (monthly, 7 ponds and 1 river, 368 samples), in Bangladesh and analyzed employing methods described previously [23, 24]. All samples

Dhaka University Institutional Repository

were collected in sterile dark Nalgene bottles (Nalgene Nunc International, St. Louis, Mo.) employing aseptic techniques. The bottles were placed in an insulated plastic box at ambient air temperature (20-35° C), transported overnight from site of collection to the laboratory in Dhaka, Bangladesh [23, 24] and samples were processed within 20 hours of collection. Total DNA was extracted from water and from plankton grouped by size into two fractions (Phytoplankton, 20 μm to 64 μm and Zooplankton, >64 μm) according to the method described by Rivera *et al.* [22] using CTAB (cetyltrimethylammonium bromide) and phenol-chloroform-isopropyl alcohol. DNA was precipitated with a 0.6 volume of isopropyl alcohol, and the pellet was washed with 70% ethanol. Samples were assayed for *V. cholerae* O1 *rfb*, *V. cholerae* O139 *rfb*, and *ctxA* by a multiplex PCR assay, using the procedure described by Hoshino *et al.*, [25] in parallel to the convention culture and enrichment for detection and identification of total and toxigenic *V. cholerae*. Primer sequences for the genes, *V. cholerae* O1 *rfb*, *V. cholerae* O139 *rfb*, and *ctxA*, are shown in Table 2.1. Amplification with the three primer pairs was performed simultaneously in a total reaction volume of 30-μl, containing 10 to 20 ng of genomic DNA, 1X reaction buffer, 210 μM dNTPs, 0.5 μM O1-*rfb* (O1F2-1, O1R2-2) primers, 0.27 μM O139-*rfb* (O139F2, O139R2) primers, 0.17-μM *ctxA* (VCT1, VCT2) primers and 0.75-U Taq polymerase with the following cycle conditions: Initial denaturation: 5 min at 94°C, 35 cycles of 1 min at 94°C (denaturation), 1 min at 55°C (annealing) and 1 min at 72°C (extension) with the final extension for 7 min at 72°C. PCR product was run on a 2.0% agarose gel in 1X TAE for 1-2 hrs at 5-V/cm (CPMB 2.5A) followed by 15 min staining of the gel in 1 μg/ml ethidium bromide staining solution and 15 min destaining in distilled water. The products were visualized by viewing the gel under UV light. The O1-*rfb*, O139-*rfb*, and *ctx*A amplicons were 192, 449, and 308-bp in length, respectively. Positive and negative controls were run in parallel and a eubacterial 16S rDNA PCR reaction was included on each sample to test template quality.

Dhaka University Institutional Repository

Target	Primer	Sequence (5'-3')	Amplicon	Ref.
*ctx*A	VCT1	ACAGAGTGAGTACTTTGACC	308	Hoshino *et al.*, 1998
	VCT2	ATACCATCCATATATTTGGGAG		
O1-*rfb*	O1F2-1	GTTTCACTGAACAGATGGG	192	Hoshino *et al.*, 1998
	O1R2-2	GGTCATCTGTAAGTACAAC		
O139-*rfb*	O139F2	AGCCTCTTTATTACGGGTGG	449	Hoshino *et al.*, 1998
	O139R2	GTCAAACCCGATCGTAAAGG		

Table 2.1.1: Primer sequences for the genes, *V. cholerae* O1 *rfb*, *V. cholerae* O139 *rfb*, and *ctxA*.

Results and Discussion:

Total DNA extraction provided sufficient template for consistent analysis of the toxigenic strains of epidemic *V. cholerae*. By using the sample concentration, DNA extraction, and PCR protocols employed here toxigenic *V. cholerae* O1, O139, and cholera toxin gene were successfully detected even when they were present in very low numbers in water and plankton samples. While the presence of nonviable cells cannot be excluded with the use of PCR, this approach offers a rapid assessment for the presence of *V. cholerae* in environmental samples with a much higher level of sensitivity than traditional methods, notably culture. However, culture methods, with combination of direct and enrichment isolation techniques, provided valuable information about the seasonality and distribution of toxigenic *V. cholerae* O1 and O139 (Figure 2.1.1) and most importantly, it enables the year round isolation of non-toxigenic *V. cholerae* non-O1/O139 with high frequency of their occurrence immediately before and after the annual outbreaks of cholera and environmental detection of toxigenic *V. cholerae* O1 and O139 (Figure 2.1.1) in this region. This finding is in support of our previous report [26] demonstrating the increased occurrence of *V. cholerae* non-O1/O139 six to eight weeks prior to the seasonal emergence of toxigenic *V. cholerae*. The study, therefore, provides a critical insight into the temporal dynamics of commensal *V. cholerae* population in cholera endemic region. Abundance and distribution of commensal *V. cholerae* population followed a similar trends over four year period suggesting that this might serve as a critical variable in developing a predictive model to forecast an upcoming cholera outbreak and could be used as an early warning system.

Dhaka University Institutional Repository

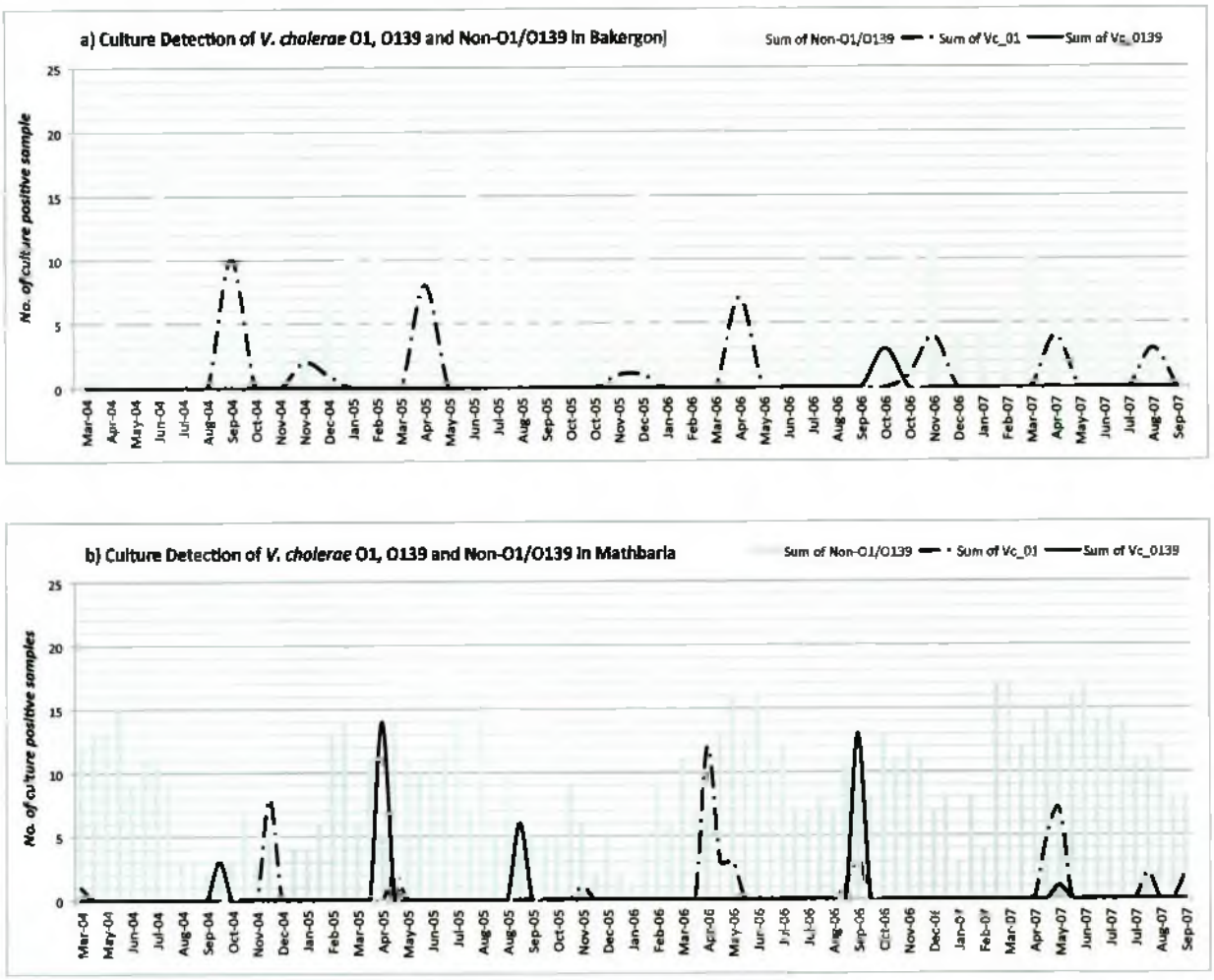


Figure 2.1.1: Temporal dynamics of *V. cholerae* isolation (O1, O139, and non-O1/O139) in two coastal regions of endemic cholera by culture method. Total number of samples (all three fractions, i.e., water, zoo- and phyto-plankton) positive for toxigenic *V. cholerae* O1 and O139 as well as non-toxigenic *V. cholerae* non-O1/O139 for each month were counted and cumulative number of samples from where isolation was made were plotted over time for both Bakergonj (a) and Mathbaria (b).

Results showed a higher level of detection of toxigenic *V. cholerae* O1 and O139 than the isolation rate using traditional detection methods, i.e., culture. Of 899 samples collected and tested, 26% (n = 234) were positive for *V. cholerae* O1, O139 and/or *ctxA* in at least one of the three fractions. Of these, 48% (n = 113) contained *V. cholerae* O1 and 28% (n = 65) contained *V. cholerae* O139 (Table 2.1.2). Cholera toxin gene, *ctxA*

Dhaka University Institutional Repository

was detected in 76% (n = 86) and 52% (n = 34) of the *V. cholerae* O1 and O139-positive samples, respectively, whereas 24% (n = 56) were positive only for the *V. cholerae* toxin gene, *ctxA*. The detection of *ctx*A genes in such a large number of samples negative for *V. cholerae* O1, and O139 indicate that cholera toxin gene is mobile and widespread within the environment. Moreover, cholera toxin gene is encoded by a temperate bacteriophage called the CTXφ bacteriophage; and is introduced into *V. cholerae* by horizontal gene transfer, therefore, potential of this gene to be carried by alternative hosts can not be readily ruled out and warrant significant implications in the area of human health. However, the occurrence of the cholera toxin gene in the environment temporarily coincided with the occurrence of *V. cholerae* O1, and O139 in both of the study areas (Figure 2.1.2 PCR), which is in contrast with the previous report in Dhaka, Bangladesh showing that bacterial viruses in the environment inversely correlate with the abundance of toxigenic *V. cholerae* in water samples [27, 28]. This contrasting finding is mostly likely attributable to the ecology of the Bay of Bengal which is significantly different from Dhaka City, the most heavily populated city in Bangladesh with nearby water bodies heavily contaminated with domestic sewage. In addition, the study reported here is a component of a longer program (4 years), involving fourteen inland and coastal environmental sites within two distinct cholera endemic areas in the Bay of Bengal [23] providing independent validation of our observations.

Area	Sample	*V. cholerae* O1	*V. cholerae* O139	*ctx*A	Total	%
Bakergonj	Zooplankton (20-64 µm)	*9*	*3*	*12*	24	26
	Plankton free Water (< 20 µm)	*14*	*3*	*17*	34	36
	Phytoplankton (> 64 µm)	*13*	*4*	*18*	35	38
	Total	***36***	***10***	***47***	**93**	
Mathbaria	Zooplankton (20-64 µm)	*25*	*18*	*40*	83	31
	Plankton free Water (< 20 µm)	*27*	*26*	*55*	109	41
	Phytoplankton (> 64 µm)	*25*	*11*	*37*	73	28
	Total	***77***	***55***	***132***	**265**	

Table 2.1.2: Occurrence of *V. cholerae* O1, *V. cholerae* O139, and cholera toxin gene, *ctxA* each of the fractionated samples.

Dhaka University Institutional Repository

Toxigenic *V. cholerae* was more consistently detected from water samples (41%) than from the >64 µm (28%) and 20-64 µm (32%) plankton fractions collected from coastal Mathbaria whereas in Bakergonj they were detected from water (37%) and the >64 µm (38%) plankton fractions almost equally, with the frequency of detection varying between months (Table 2.1.3). Analysis by geographic regions revealed, toxigenic *V. cholerae* was present in Mathbaria water throughout the year and at a higher frequency (34%), with distinct seasonal peaks during the pre and post Monsoon seasons (Table 2.1.3). Whereas, in Bakergonj, toxigenic *V. cholerae* was detected seasonally (pre and post Monsoon) and at much lower frequency (15%). In Bangladesh, cholera outbreaks usually occur twice during a year, with the highest number of cases just after the monsoon and a somewhat smaller number of cholera cases during the spring (pre monsoon). Interestingly, our environmental surveillance in the two coastal cities demonstrated an opposite phenomenon, although cholera also occurs twice a year in these regions, highest number of isolation was observed in the pre-monsoon season (66%), as opposed to post monsoon season (27%). These observations suggest that pre-monsoon cholera might heavily been influenced by the regional environmental drivers whereas post monsoon cholera may have been driven by environmental drivers as well as other local-level variables (e.g., water and sanitation). Despite these difference, the other general trends of seasonal cholera emergence in both areas followed typical historical patterns.

Dhaka University Institutional Repository

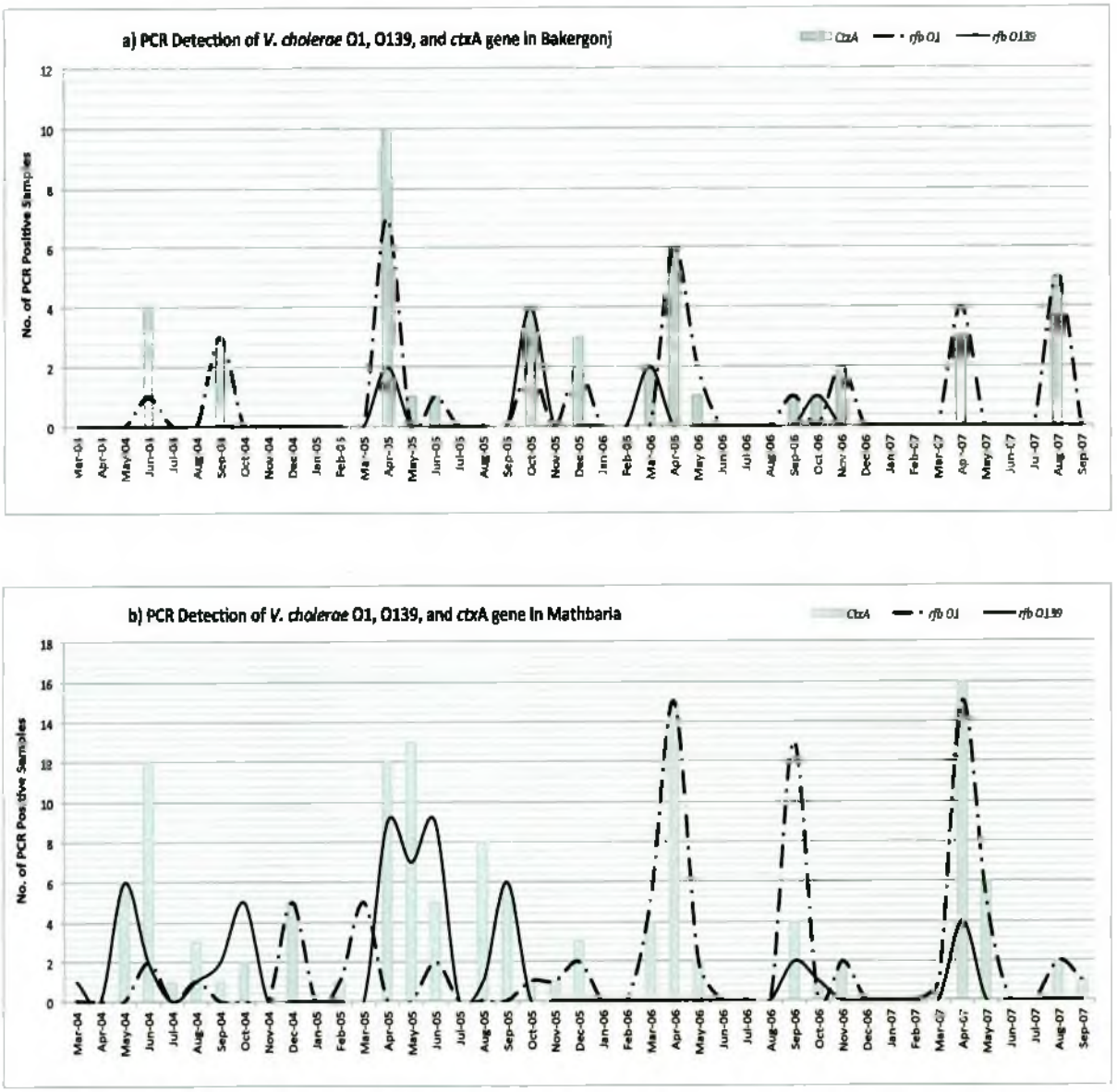


Figure 2.1.2: Temporal dynamics of *V. cholerae* O1, O139, and cholera toxin gene (*ctx*A) detection in two coastal regions of endemic cholera by direct multiplex PCR. Total number of samples (all three fractions, i.e., water, zoo- and phyto-plankton) positive for either *V. cholerae* O1, O139 or *ctx*A genes in each month were counted and cumulative number of PCR positive samples were plotted over time for both Bakergonj (a) and Mathbaria (b).

Dhaka University Institutional Repository

Month/Year	2004	2005	2006	2007	Total
January	ND	0	0	0	0
February	ND	2	0	0	2
March	1	5	13	2	21
April	0	40	42	42	124
May	12	21	7	11	51
June	21	18	0	0	39
July	1	0	0	0	1
August	5	9	0	14	28
September	9	12	25	2	48
October	7	12	2	ND	21
November	0	2	4	ND	6
December	10	5	0	ND	15

*ND: Not Done

Table 2.1.3: Seasonal detection of *V. cholerae* O1, *V. cholerae* O139, and cholera toxin gene, *ctx*A in Bakergonj and Mathbaria, Bangladesh.

Cholerae cases recorded in Bakerganj showed a similar seasonal pattern of cholera cases as observed in Matlab and Dhaka, Bangladesh [29] with a small outbreak in pre monsoon season (spring) and a bigger post monsoon peak to cause Fall outbreak (Fig 2.1.3). Mathbaria, on the other hand, displayed a different seasonal pattern, with only one big outbreak in spring (Fig 2.1.3). However, there was a break in the dynamics during winter months, particularly between February and March. These seasonal patterns of cholera cases are strongly in agreement with our environmental *V. cholerae* detection demonstrating highest number of isolation in the pre-monsoon season (66%), as opposed to post monsoon season (27%). Pattern of detection of all combined *V. cholerae* O1, O139 and *ctx*A in the environment of the two areas were similar and bi-modal, with the trend more pronounced in Mathbaria (Figure 2.1.4). *V. cholerae* O1 and *ctxA* were more frequently detected in fall in Bakerganj in comparison to Mathbaria (Figure 2.1.4). This may explain the uni-modal cycle of cholera seasonality in Mathbaria. However, detection of *V. cholerae* O1 and *ctx*A is strongly associated with cholera in the populations.

Dhaka University Institutional Repository

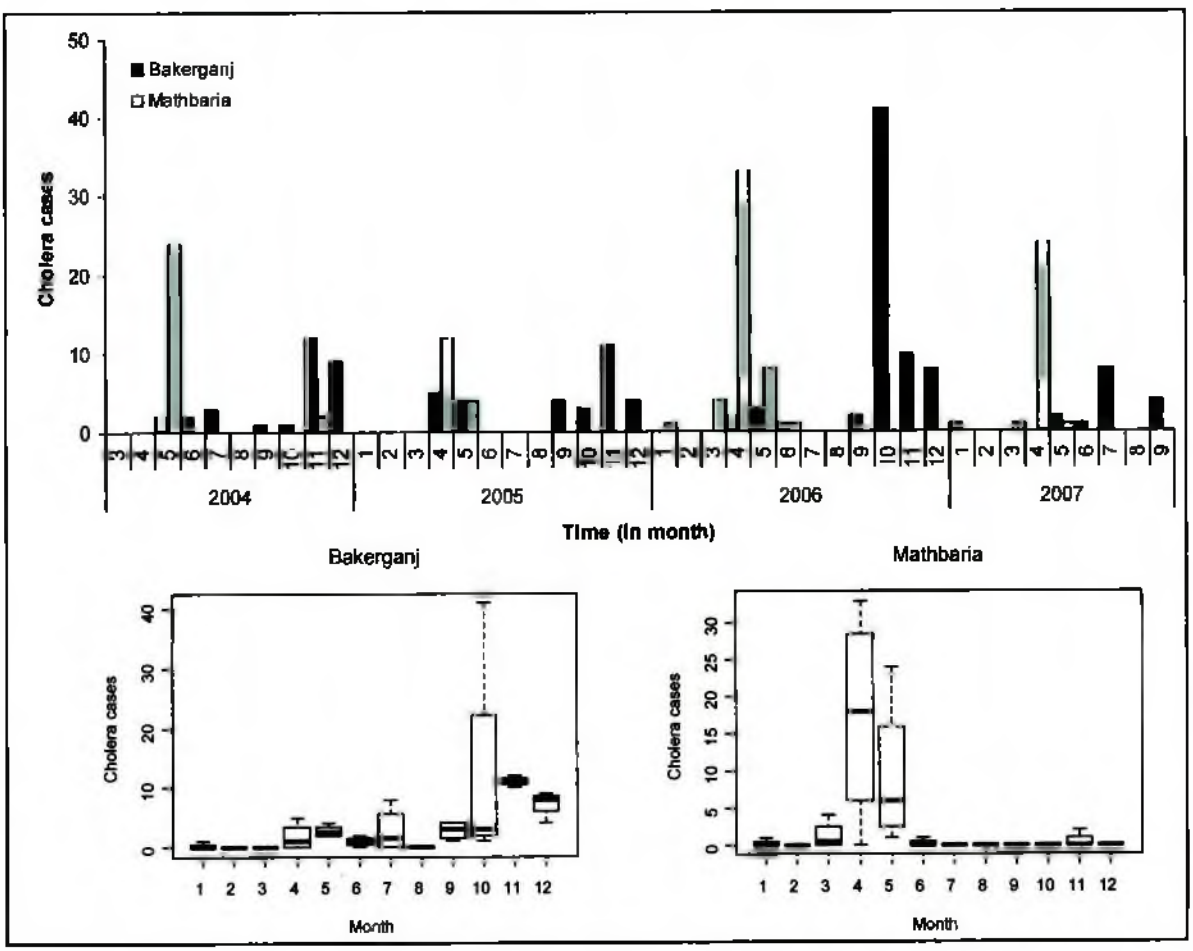
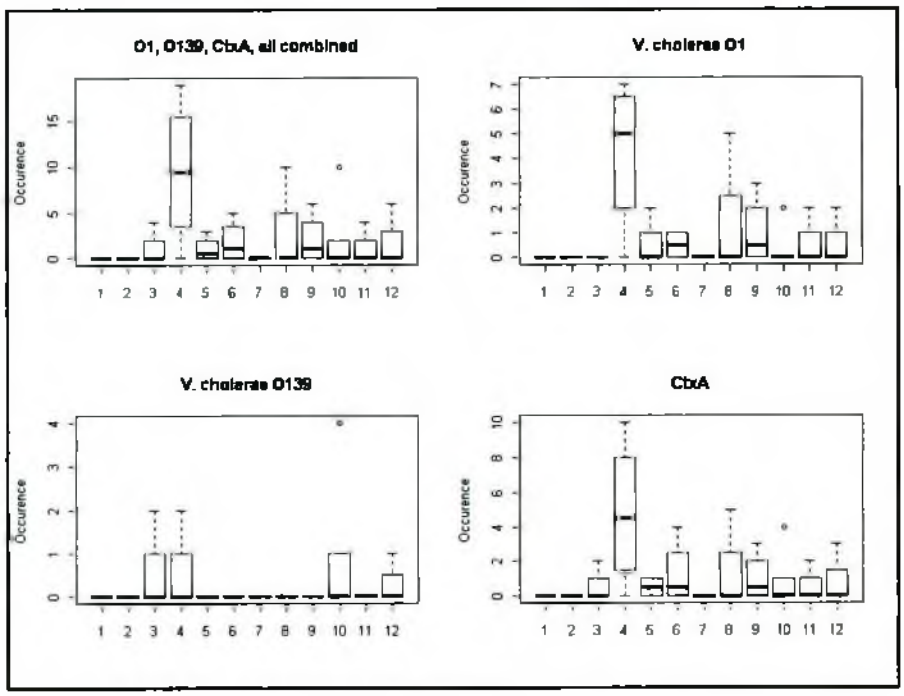
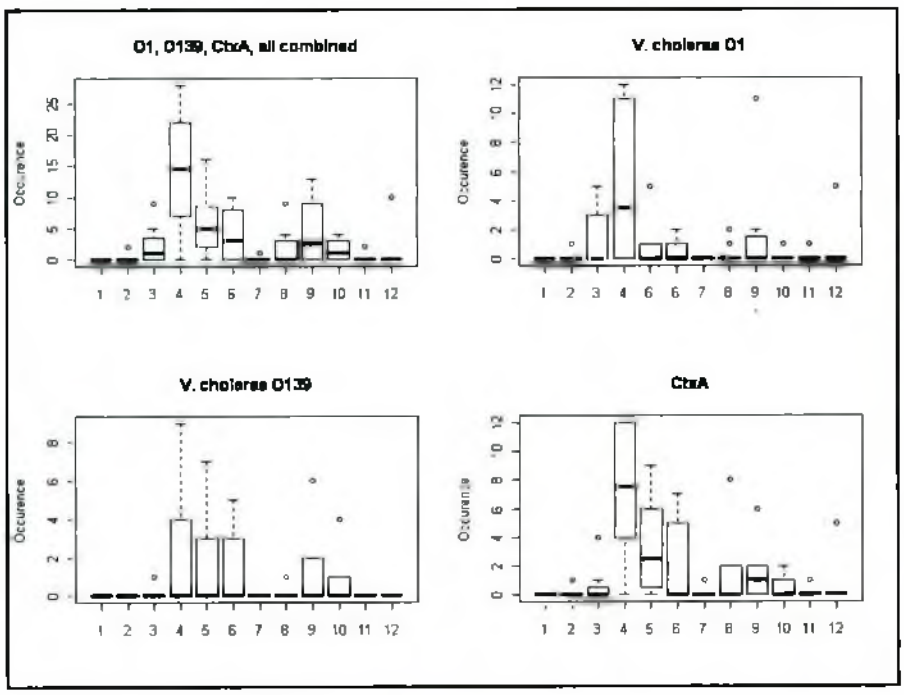


Figure 2.1.3: Seasonal occurrence of cholerae cases in Bakergonj and Mathbaria collected through hospital surveillance. The boxplots indicate the seasonality of cholera cases observed in these two localities.

Dhaka University Institutional Repository



a) Bakergonj



b) Mathbaria

Figure 2.1.4: Seasonality of detection of *V. cholerae* O1, O139, and cholera toxin gene (*ctxA*) in Bakergonj (a) and Mathbaria (b)

Dhaka University Institutional Repository

Conclusion:

In conclusion, PCR was successfully employed to detect the toxigenic *V. cholerae* directly in water and plankton samples and we have shown that direct DNA extraction and application of multiplex PCR assays are important tools for rapid assessment of environmental samples for toxigenic *V. cholerae* and provide a higher level of sensitivity than traditional culture. The results obtained using these methods provide unambiguous evident that *V. cholerae* and its major virulence factor, the cholera toxin gene, is present in the coastal waters of the Bay of Bengal throughout the year, with numbers of cells increasing during March to May and September to November, just prior to the annual outbreaks of cholera in the coastal areas. *V. cholerae* O1 and O139 cells, and cholera toxin gene, *ctx*A are prevalent in both water and plankton samples. The data demonstrate that the coastal water and plankton populations of the Bay of Bengal harbor toxigenic *V. cholerae* as a component of the natural microbial flora. The study also demonstrated the increased detection of commensal *V. cholerae* non-O1/O139 preceding the seasonal emergence of toxigenic *V. cholerae* O1 and O139, a critical finding in understanding the temporal dynamics of *V. cholerae* in the environment, and hence, offers a tremendous value to included as an important variable for a cholera prediction model.

Finally, the data presented here demonstrate unequivocally that *V. cholerae* is autochthonous in the coastal, brackish, and riverine water in this region and it is reasonable to examine those environmental conditions giving rise to both the plankton blooms and the increases in *V. cholerae* concentrations. This information will help to develop a predictive model for cholera epidemics and provide an early warning system for conditions conducive to cholera outbreaks.

Dhaka University Institutional Repository

2.2: Rapid Enumeration of *Vibrio cholerae* in Water and Plankton Samples by Fluorescence *In Situ* Hybridization (FISH) and RNA Colony Blot Hybridization

Abstract:

We employed a species-specific RNA colony blot hybridization protocol for enumeration of *V. cholerae* in environmental water samples using a fluorescently-labeled 16S rRNA probe to detect complimentary RNA signatures from lysed cells immobilized on membranes. Results for water samples from six ponds in Bangladesh demonstrated a peak in *V. cholerae* abundance 1-2 months prior to an observed increase in pathogenic *V. cholerae* isolation and arrival of clinical cases to a nearby health clinic, suggesting a change in the local niche of *V. cholerae*, from free-swimming to aggregated or biofilm phase, such as attached to zooplankton. This method provides a highly specific and sensitive tool for monitoring and enumerating *V. cholerae* in the environment. However, results of Fluorescent *in situ* hybridization (FISH), Acridine Orange Direct Count (AODC), DAPI (4',6' diamino-2-phenylindole-2HCl) staining and Heterotrophic plate count (HPC) carried on a subset of samples provided contrasting distribution of *V. cholerae* and other bacterial cells demonstrating relatively higher abundance of *V. cholerae* and other bacterial cells in plankton free water samples than that of each plankton fractionates and this trend was consistently observed among all 8 rounds of sampling. Moreover, it also indicated that AODC counts were almost 2 orders of magnitude higher than that of HPC, reaffirming the great plant count anomalies. Interestingly, marginal temporal variation observed in total bacterioplankton abundance in contrast to strong temporal variation of the *V. cholerae* abundance indicates that, there might be a significant shift within the composition of each community which not necessarily cause any shift significant shift in total community structure.

Dhaka University Institutional Repository

Introduction:

Vibrio cholerae is ubiquitous in aquatic environments [30], while *V. cholerae* serogroups O1 and O139 have been shown to be associated with epidemic cholera, serogroups non-O1/O139 have been recognized as the causative agents of sporadic cholera-like disease and outbreaks worldwide [31-34]. *V. cholerae* usually live in association with crustacean zooplankton, mainly copepods, where they build biofilms on external and internal surfaces [35, 36]. It has been hypothesised that attachment to surfaces serves as a survival strategy of *V. cholerae* to escape from heterotrophic nanoflagellate grazing [37]. Due to its ability to degrade chitin, these surfaces are also used as a nutrient source [38, 39]. Moreover, chitin has been shown to induce DNA uptake by *V. cholerae* [40, 41] enhancing the environmental fitness of this bacterium. The environmental nontoxigenic, non-O1/O139 strains also play an important role in the evolution of toxigenic *V. cholerae* [42]. However, in contrast to the vast knowledge of the ecology of *V. cholerae* O1 and O139, much less is known about the environmental behaviour of *V. cholerae* non-O1/ O139 strains. Cholera surveillance in most countries, is primarily based on detection of *V. cholerae* O1 and O139 and determining the presence of cholera toxin (CT) by using biological and molecular methods [43]. However, virulence-associated factors in *V. cholerae* isolates from environmental sources are of concern [43]. The emergence of serogroup O139 as a second etiologic agent of cholera epidemics, along with the possible conversion of non-O1/O139 to O1 serotype [44] and the emergence of *V. cholerae* O10, O12, O31, O37, O53, and O75 as bacterial serogroups associated with cholera-like epidemics [34, 45-50], has caused further interest in the non-O1/O139 *V. cholerae* strains.

Rapid and precise methods for detection and quantification of pathogenic bacteria are fundamental for monitoring and risk assessment in regard to the use of water as well as for studying their ecology in the environment. Detection of *V. cholerae* present in natural waters still relies mostly on an isolation method originally developed for clinical diagnosis, namely enrichment in alkaline peptone water, followed by subculture on selective media, and confirmation using selected biochemical, molecular, and

Dhaka University Institutional Repository

immunological tests. The alkaline nature of the enrichment broth allows differential multiplication of *Vibrio* species, but renders this method inappropriate for enumeration. In addition, they cannot be applied when *Vibrio* cells turn into the VBNC state. Due to this phenomenon, whereby cells are still present in the environment but cannot be detected by culture, culture-independent methods are needed. Traditional PCR is a common molecular method alternative to cultivation. Nevertheless, conventional PCR requires further product characterization by gel electrophoresis, and, even more important, it is a non-quantitative method. Real-time PCR is a strong improvement of traditional PCR and can be effectively used for quantification of microorganisms [51] Variable DNA extraction efficiencies as well as PCR inhibition, especially in environmental samples, are still a problem and need to be further improved [51].

Fluorescence *in situ* hybridization (FISH) is a direct cell-based detection cultivation-independent method to quantify *V. cholerae* in the environment [52-54]. FISH with rRNA-targeted oligonucleotide probes, combined with epifluorescence microscopy (EFM), has become a very strong tool in observing microbial communities in their environment and is therefore indispensable in the field of microbial ecology. FISH allows visual quantification of all *V. cholerae* cells, regardless of their culturability, in a sample. One problem with this method is observation of auto-fluorescing constituents found in environmental water samples. All "positive" samples should be documented by digital or film photography, and confirmation by conventional PCR and/or Real-Time PCR of a parallel sample (frozen, but not fixed) is advisable.

These facts underline the need for a method of detection of the total number of *V. cholerae* present. 16S rDNA is an appropriate target and probe Vchomim1276, previously described by Heidelberg [13-15], sensitivity and specificity of this probe was tested using environmental and clinical isolates, therefore, employed in an RNA colony blot hybridization protocol which was previously evaluated using environmental water and plankton samples seeded with *V. cholerae* [26]. In this study, we used both FISH and RNA colony blot hybridization methods for a faster enumeration of total (both

Dhaka University Institutional Repository

culturable and non-culturable *V. cholerae*) and culturable *V. cholerae* from fresh water ponds in cholera endemic regions, respectively.

Methods:

Water samples were collected from 6 ponds along coastal regions of Mathbaria, Bangladesh from July 2006 to March 2007 and analyzed employing methods described previously [23, 24]. During the study period, each of the ponds were sampled biweekly, and all together 17 visits were made over a period of 9 months. For RNA colony blot hybridization, whole water samples were plated on Heterotrophic plate count (HPC) & Luria-Bertani (LB) Agar plates in triplicate using 100 µl each of a three-fold dilution series and incubated at 30°C and 37°C, respectively, for appropriate growth, i.e., to achieve a well separated colonies (about 200-300) with size no larger than 5 mm in diameter. For fluorescence *in situ* hybridization, a discrete set of samples was analyzed by randomly choosing 8 visits (round) worth of samples (8 visits, 6 ponds, 3 fractionates, totaling 144 fractionates). From each pond 100 L of samples were concentrated and fractionated employing methods described previously [23, 24] and 1 mL of each concentrated fractions (Filtrated Water, Zoo- & Phyto plankton) is harvested by centrifugation at 13000 rpm for 5 min and fixed with 4% paraformaldehyde solution at room temperature for 1-2 hr. Cells are pelleted by centrifugation at 13000 rpm for 5 min, washed twice in Phosphate buffered saline (PBS), and concentrated by centrifugation followed by re-suspension in 100 µl PBS-ethanol solution and stored at –20°C until used.

The 16S rDNA probe, Vchomim1276, specific for *V. cholerae* and *V. mimicus* [55, 56] was employed. The probe, ACT TTG TGA GAT TCG CTC CAC CTC G (T_m = 72°C), was 5' end-labeled with either FITC, fluorescein or Cy3 at the time of manufacture (Sigma-Genosys). RNA colony blot hybridization with fluorescein or Cy3-labelled probe were employed to enumerate *V. cholerae* from 102 whole water samples whereas fluorescence *in situ* hybridization with FITC-labeled probe were employed to a smaller subset of samples (144 fractionated samples i.e., filtered water, Zooplankton and phytoplankton fractions) derived from 48 whole water samples.

Dhaka University Institutional Repository

RNA colony lift. A previously reported RNA colony blot lift method [57] was employed to liberate, denature, and immobilize nucleic acid onto nylon or nitrocellulose membranes (GE Osmonics). Briefly, colony containing agar plates were overlaid with membranes and the transfer was allowed to proceed for at least 15 min. Membranes were transferred colony side up onto filter paper wetted with 10% SDS and incubated for 5-10 minutes. Membranes were then transferred to filter paper wetted with 3X SSC, pre-warmed to 65°C, and allowed to incubate for 10-15 min. Membranes were dried at 37°C for 10 min., and then baked at 70°C for 15 min. Methylene blue was used to qualitatively examine RNA and/or DNA immobilized onto membranes.

RNA colony blot hybridization. All solutions used for RNA colony blot lift and hybridization were treated with diethyl pyrocarbonate (DEPC) to eliminate RNase contamination. Colony blots were washed in 3X SSC, 0.1% SDS, three times for 10 minutes each, at room temperature, and once at 65°C for 2 hrs, to remove cellular debris that could interfere with hybridization. Membranes were prehybridized at 60°C for 30 minutes in hybridization solution base (10 ml per 100 cm^2 membrane), 0.9 M NaCl, 50 mM sodium phosphate/pH 7.0, 5 mM EDTA, 0.5% SDS in DEPC-treated water. Hybridizations were performed overnight (approximately 16 hours) at 60°C, using hybridization solution base (10 ml per 100 cm^2 membrane) containing fluorescently labeled Vchomim1276 probe (80 ng/ml). After hybridization, membranes were washed in washing solution, 1% SDS, 1X SSC in DEPC-treated water, at hybridization temperature for 30 minutes. Membranes were wrapped in plastic wrap or placed in a hybrization bag and visualized using a Typhoon (GE Healthcare, Piscataway, NJ) variable mode imager. Alternatively, colony blots hybridized with Cy3- or fluorescein-labeled Vchomim1276 were visualized with a Dark Reader hand lamp (Clare Chemical).

Fluorescence *in situ* hybridization (FISH). FITC-labeled Vchomim1276 probe were diluted at a working concentration of 50 ng/µl in sterile water for a final concentration of 5 ng/µl. Hybridization was performed according to the protocol described by Kirschner et al. [58] with some modifications. Briefly, samples were fixed

Dhaka University Institutional Repository

with 4% PFA, 10 µl aliquots of fixed cells are spotted onto multiwell slides and air-dried. Hybridization chambers were setup as follows: hybridization solution-soaked filter paper strips were inserted into centrifuge tubes and sealed tightly. Tubes were placed in 45°C incubator or water bath for 30 min to equilibrate and than washed for 3 min each in successive ethanol solutions (50, 80 and 96-100%) and air-dried thereafter. Samples were hybridized with FITC-labeled probe at 45C for 24 hrs, than washed in pre-warmed washing buffer at 45°C and in sterile deionized water and air-dried at room temperature. Anti-fade agent was added to each well and a coverslip is applied. Microscope slides can be stored at −20°C for at least 6 months. Slides are viewed with epifluorescent microscope (Carl Zeiss), by examining the entire well.

A limited number of samples were also been assessed by two other microscopic methods, i.e., Acridine Orange Direct Count (AODC) and DAPI (4',6' diamino-2-phenylindole-2HCl) staining using standard protocol [59, 60].

Results and Discussion:

In this study, we employed an RNA colony blot hybridization method using a fluorescently labeled oligonucleotide probe specific for *V. cholerae*. The use of a commercially available fluorochrome label was preferred over other labels because of laboratory personnel safety, operational regulations, availability, and/or economic reasons. A previously described *V. cholerae/mimicus*-specific 16S rDNA probe [13] was employed for fluorescent label/RNA colony blot hybridization. This method takes advantage of the fact that during early growth (16 hours to 2 days) on complex or rich agar medium, growth is vigorous. Under these conditions, bacterial colonies have a high mRNA content. Therefore, by targeting RNA, instead of DNA, the fluorochrome signal was boosted sufficiently for reliable detection. Figure 2.2.1 shows results for two of the positive blots prepared using water collected from the Mathbaria province, Bangladesh (site 1 and 2).

Dhaka University Institutional Repository

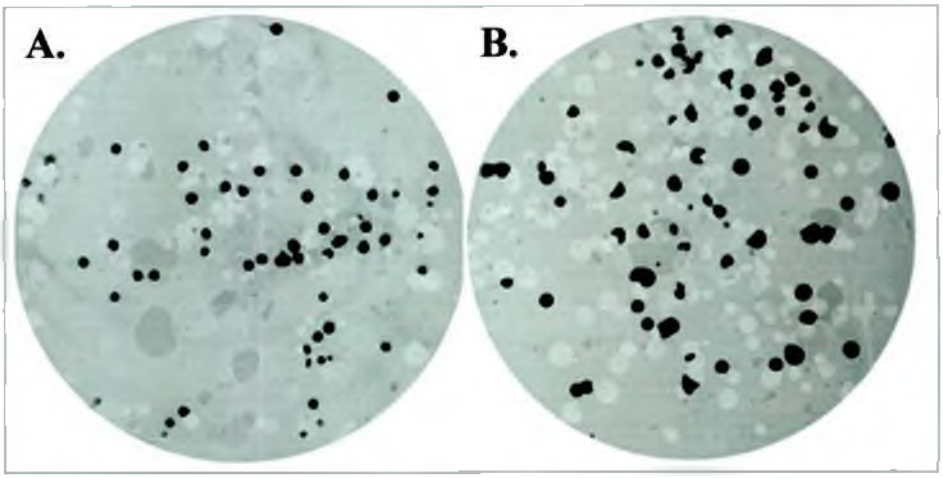


Figure 2.2.1. Representative RNA colony blots hybridized with probe Vchomim1276. Blots were prepared from plates being spread with 100µL of direct water samples and/or it's suitable dilutions. Blots shown here were prepared from 100µL of direct water sample from pond 3 at July 2006, Round # 60 (A) and pond 1 at Jan 2007, Round#73 (B) of Mathbaria, Bangladesh.

Temporal variation in *V. cholerae* population in Bangladesh. To assess utility of the RNA colony blot lift and hybridization method, 102 water samples collected from a cholera endemic area were tested and 41% of the samples yielded *V. cholerae* signal with maximum count of $1.0x10^4$ CFU/ml. When the abundance of total culturable *V. cholerae* counts, as determined by this method, were cumulated month by month, an interesting temporal variation of planktonic *V. cholerae* abundance was observed (Figure 2.2.2). Survey of *V. cholerae* in freshwater ponds in Bangladesh over a nine month period demonstrated an abundance of planktonic *V. cholerae* cells 1-2 months prior to an environmental peak in pathogenic *V. cholerae* isolation and onset of clinical cases (Figure 2.2.2). This finding is strongly in agreement with increased isolation of *V. cholerae* non-O1/O139 using culture method as described in previous section.

Dhaka University Institutional Repository

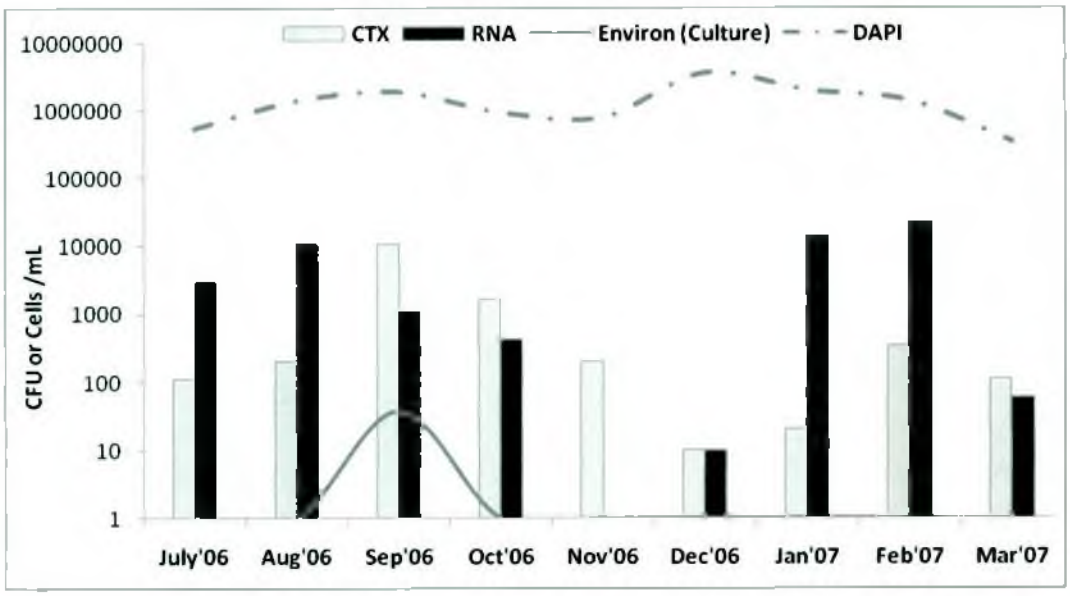


Figure 2.2.2. Seasonality and abundance of total and pathogenic *V. cholerae* as determined by conventional culture and molecular methods (targeting *ctxA* and 16s Intergenic spacer) in compared to total bacterial counts in coastal areas of Mathbaria, Bangladesh.

Results of the field trials demonstrate that the RNA colony blot hybridization method is robust and can be used to detect and enumerate *V. cholerae* in different types of environmental water samples. In its simplest form, this method can be used to detect and enumerate total *V. cholerae*. The method can also be adapted for use in isolation of the organism. If the colony blot lift is made from a spread-plate, then the master plate can be stored at 15°C, until the hybridization is complete, at which time the positive signals can be correlated to colonies and subsequently subcultured. If the colony blot is made by direct filtering of a water sample onto the membrane (and then overlaying onto an agar plate), then the subsequent growth on the membrane can be replica-plated onto a fresh agar plate prior to colony lysing. Again, positive signals can then be correlated to colonies and then subcultured and confirmed. However, it is important to mention that the procedure is limited because it relies on culture. Thus, for routine monitoring, this approach is inferior to PCR-based methods, which are also usually more sensitive. The method is applicable for *V. cholerae* population studies in which quantification and isolation of the bacterium followed by characterization is needed. For this purpose, this

Dhaka University Institutional Repository

method is preferred over traditional enrichment/selective medium isolation methods, because the necessity for extensive and laborious biochemical tests can be avoided.

Fluorescent *in situ* hybridization (FISH) carried on a subset of samples to understand the distribution of *V. cholerae* cells into three fractionates of the whole water sample, i.e., water, phyto- & zoo-plankton fractions, demonstrated that highest abundance of *V. cholerae* cells were associated with plankton free water, whereas their abundance were quite comparable in both of plankton fractionates (Figure 2.2.3) and this trend was consistently observed among all 8 rounds of sampling. However, when total (culturable, VBNC, and dead) bacterioplankton densities among these 3 fractionates measured by two other microscopic methods, i.e., AODC and DAPI were compared across 6 study sites, similar distribution cell densities were observed with a very tight concordance of density estimates between the two methods (Figure 2.2.4). Next, we compared the total (culturable, VBNC, and dead) bacterioplankton densities measured by AODC with the total culturable bacterioplankton densities as observed by HPC count (Figure 2.2.5), which also reaffirm the greater abundance of bacterioplankton in the water fractions compared to the plankton fractions. Moreover, it also indicated that AODC counts were almost 2 orders of magnitude higher than that of HPC, which can be explained by the great plant count anomaly, and indicated that environmental samples frequently inhabit VBNC and dead cells, and, in addition, some stressed bacterioplankton were also be there which can easily escape detection when relied on culture based detection and enumeration methods. Interestingly, we observed marginal temporal variation in total bacterioplankton abundance when measured by AODC, DAPI or HPC, whereas RNA colony blots hybridization and/or FISH, which is specific to *V. cholerae* cells only, demonstrated strong temporal variation of the *V. cholerae* abundance similar to what we have observed by *V. cholerae* specific direct PCR and culture method. This indicates that, even in absence of any significant shift in total community structure, there might be a significant shift within the composition of each community, perhaps due to the results of certain environmental drivers that may favor the growth of one community member over the others. Such distortion in the natural community composition might have a predictive

Dhaka University Institutional Repository

value, and may add new know in our understanding of seasonal emergence and reemergence of *V. cholerae* in their natural ecosystem.

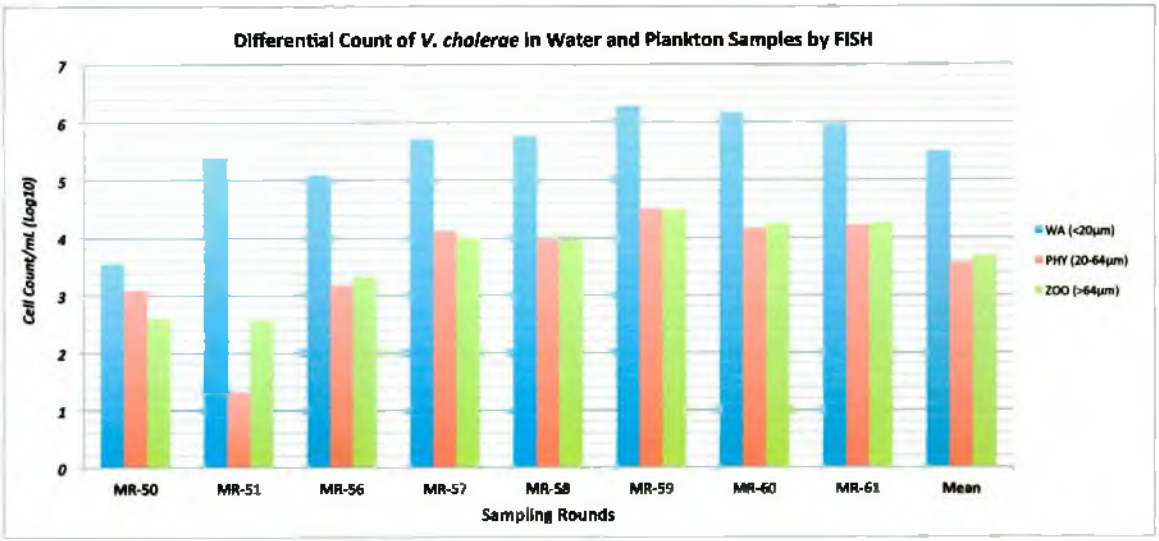


Figure 2.2.3: Detection and enumeration of total *V. cholerae* (both culturable and non-culturable *V. cholerae* cells) associated with plankton free water, Phytoplankton and Zooplankton.

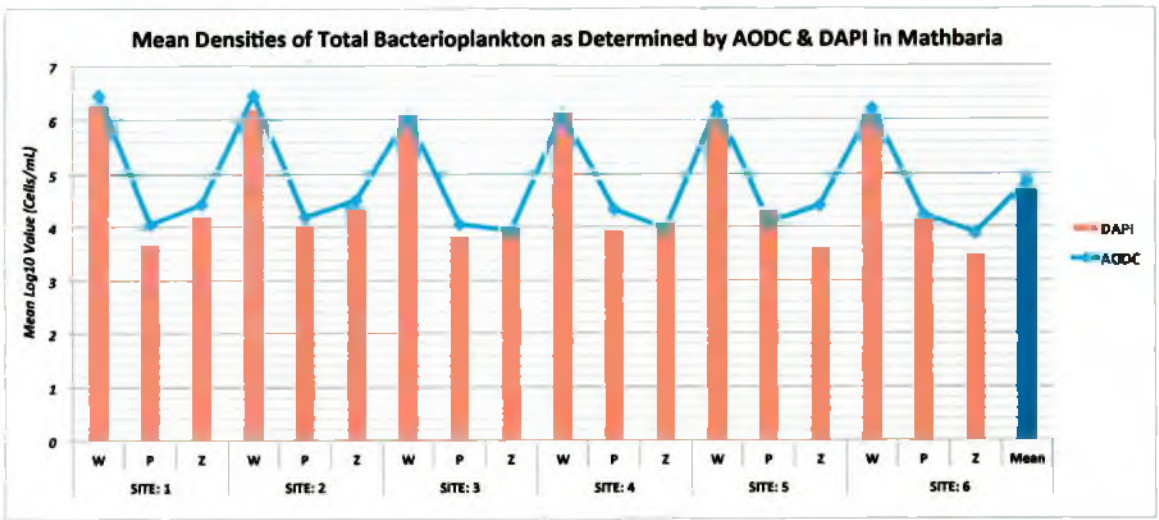


Figure 2.2.4: Enumeration of total (culturable, VBNC, and dead) bacterioplankton associated with plankton free Water, Phytoplankton and Zooplankton samples as measured by Acridine Orange Direct Count (AODC) and DAPI (4',6' diamino-2-phenylindole-2HCl) staining.

Dhaka University Institutional Repository

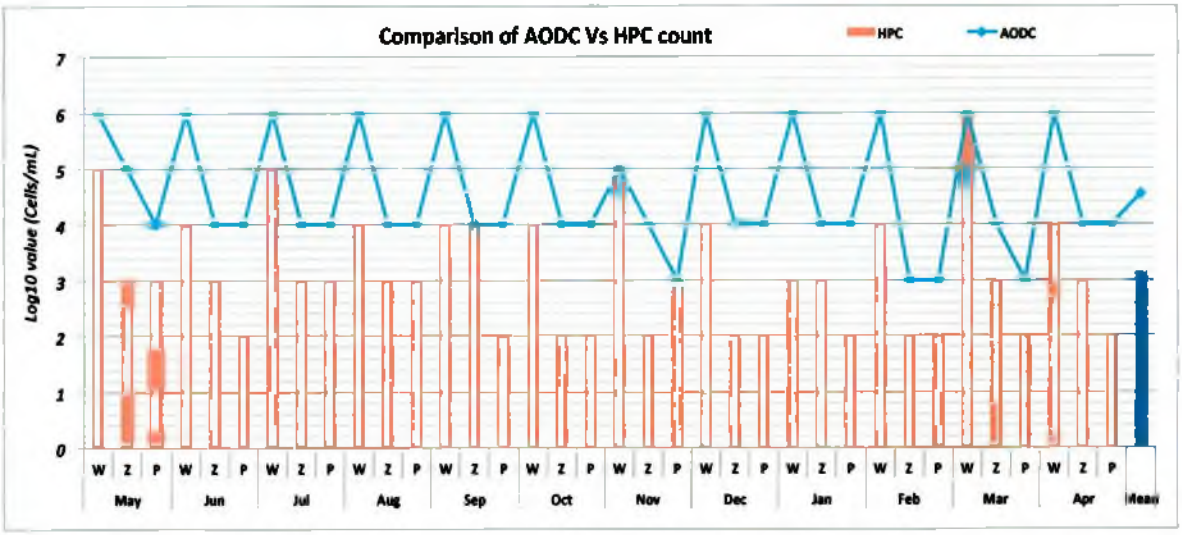


Figure 2.2.5: Comparison of total bacterioplankton count as opposed to total culturable heterotrophic bacterial count associated with plankton free Water, Phytoplankton and Zooplankton.

Conclusion:

In conclusion, the colony blot methods employed in this study have many advantages in detection and isolation of *V. cholerae*. Detection and isolation is less tedious than traditional culture methods and results are achieved in less time. Enumeration is also possible, whereas traditional culture methods involve enrichment. The method has high specificity, as is the case with other molecular methods, such as PCR. Environmental microbiological methods generally have low sensitivity, but this shortcoming can be addressed by including sample concentration by filtration. Use of FISH, AODC, DAPI, and HPC culture counts provided some important insides about the bacterial community structure and their shift in composition over time, and demonstrated that there might be some temporal distortion in the bacterial community, as a result of selective growth advantages gained by certain community members over others perhaps due to seasonal shift of some environmental parameters.

Dhaka University Institutional Repository

References:

1. Sack, R.B., et al., *A 4-year study of the epidemiology of Vibrio cholerae in four rural areas of Bangladesh.* J Infect Dis, 2003. **187**(1): p. 96-101.
2. Colwell, R.R., J. Kaper, and S.W. Joseph, *Vibrio cholerae, Vibrio parahaemolyticus, and other vibrios: occurrence and distribution in Chesapeake Bay.* Science, 1977. **198**(4315): p. 394-6.
3. Colwell, R.R., et al., *Occurrence of Vibrio cholerae serotype O1 in Maryland and Louisiana estuaries.* Appl Environ Microbiol, 1981. 41(2): p. 555-8.
4. Huq, S., *Climate change and Bangladesh.* Science, 2001. **294**(5547): p. 1617.
5. Bose, A., J.K. Philip, and M. Jesudason, *Neonatal septicemia caused by Vibrio cholerae O:139.* Pediatr Infect Dis J, 2000. **19**(2): p. 166.
6. Jiang, S.C. and W. Fu, *Seasonal Abundance and Distribution of Vibrio cholerae in Coastal Waters Quantified by a 16S-23S Intergenic Spacer Probe.* Microb Ecol, 2001. **42**(4): p. 540-548.
7. Huq, A., et al., *Detection of Vibrio cholerae O1 in the aquatic environment by fluorescent-monoclonal antibody and culture methods.* Appl Environ Microbiol, 1990. **56**(8): p. 2370-3.
8. Onifade, T.J., et al., *Toxin producing Vibrio choierae O75 outbreak, United States, March to April 2011.* Euro Surveill, 2011. **16**(20): p. 19870.
9. Cash, R., et al., *Response of man to infection with Vibrio cholerae I. Clinical, serologic, and bacteriologic responses to a known inoculum.* J. Infect. Dis., 1974. **129**: p. 45-52.
10. Huq, A., et al., *The role of plantonic copepods in the survival and multiplication of Vibrio cholerae in the aquatic environment.*, in *Vibrios in the Environment*. 1984, John Wiley & Sons: New York, New York. p. 521-534.
11. Rawlings, T.K., G.M. Ruiz, and R.R. Colwell, *Association of Vibrio cholerae O1 El Tor and O139 Bengal with the Copepods Acartia tonsa and Eurytemora affinis.* Appl Environ Micbrobiol, 2007. **73**(24): p. 7926-7933.
12. Colwell, R., *Viable but non-culturable bacteria in the aquatic environment.* Culture, 1991. **12**: p. 1-4.
13. Heidelberg, J., *Seasonal abundance of bacterioplankton populations of Bacteria, gamma proteobacteria, Vibrio/Photobacterium, Vibrio vulnificus, Vibrio cholerae/Vibrio mimicus, and Vibrio cincinnatiensis associated with zooplankton in the Choptank River, Maryland. Ph.D. thesis.*, 1997, University of Maryland College Park.
14. Heidelberg, J., K. Heidelberg, and R. Colwell, *Seasonality of Chesapeake Bay bacterioplankton species.* Appl Environ Microbiol, 2002. **68**(11): p. 5488-5497.
15. Heidelberg, J., K. Heidelberg, and R. Colwell, *Bacteria of the gamma-subclass Proteobacteria associated with zooplankton in Chesapeake Bay.* Appl Environ Microbiol, 2002. **68**(11): p. 5498-5507.
16. Xu, H.S., et al., *Survival and viability of nonculturableEscherichia coli andVibrio cholerae in the estuarine and marine environment.* Microbial Ecology, 1982. **8**(4): p. 313-323.

17. Colwell, R.R. and A. Huq, *Environmental reservoir of Vibrio cholerae. The causative agent of cholera.* Ann N Y Acad Sci, 1994. **740**: p. 44-54.

18. Roszak, D.B., D.J. Grimes, and R.R. Colwell, *Viable but nonrecoverable stage of Salmonella enteritidis in aquatic systems.* Can J Microbiol, 1984. **30**(3): p. 334-8.

19. Rollins, D. and R. Colwell, *Viable but nonculturable stage of Campylobacter jejuni and its role in survival in the natural aquatic environment.* Applied and Environmental Microbiology, 1986. **52**(3): p. 531-538.

20. Yamai, S., et al., *[Studies of Vibrio cholerae O140 (serogroup Hakata) isolated from river water].* Kansenshogaku Zasshi, 1997. **71**(9): p. 928-34.

21. Yamai, S., et al., *[Distribution of serogroups of Vibrio cholerae non-O1 non-O139 with specific reference to their ability to produce cholera toxin, and addition of novel serogroups].* Kansenshogaku Zasshi, 1997. **71**(10): p. 1037-45.

22. Rivera, I.N., et al., *Method of DNA extraction and application of multiplex polymerase chain reaction to detect toxigenic Vibrio cholerae O1 and O139 from aquatic ecosystems.* Environ Microbiol, 2003. **5**(7): p. 599-606.

23. Alam, M., et al., *Seasonal cholera caused by Vibrio cholerae serogroups O1 and O139 in the coastal aquatic environment of Bangladesh.* Appl Environ Microbiol, 2006. **72**(6): p. 4096-104.

24. Alam, M., et al., *Toxigenic Vibrio cholerae in the aquatic environment of Mathbaria, Bangladesh.* Appl Environ Microbiol, 2006. **72**(4): p. 2849-55.

25. Hoshino, K., et al., *Development and evaluation of a multiplex PCR assay for rapid detection of toxigenic Vibrio cholerae O1 and O139.* FEMS Immunol Med Microbiol, 1998. **20**(3): p. 201-7.

26. Grim, C.J., et al., *RNA colony blot hybridization method for enumeration of culturable Vibrio cholerae and Vibrio mimicus bacteria.* Appl Environ Microbiol, 2009. **75**(17): p. 5439-44.

27. Faruque, S.M., et al., *Self-limiting nature of seasonal cholera epidemics: Role of host-mediated amplification of phage.* Proceedings of the National Academy of Sciences, 2005. **102**(17): p. 6119-6124.

28. Faruque, S.M., et al., *Seasonal epidemics of cholera inversely correlate with the prevalence of environmental cholera phages.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(5): p. 1702-1707.

29. Huq, A., et al., *Critical factors influencing the occurrence of Vibrio cholerae in the environment of Bangladesh.* Appl Environ Microbiol, 2005. **71**(8): p. 4645-54.

30. Rhodes, J.B., H.L. Smith, Jr., and J.E. Ogg, *Isolation of non-O1 Vibrio cholerae serovars from surface waters in western Colorado.* Appl Environ Microbiol, 1986. **51**(6): p. 1216-9.

31. Bhattacharya, S.K., et al., *Acute secretory travellers' diarrhoea caused by Vibrio cholerae non-01 which does not produce cholera-like or heat-stable enterotoxins.* J Diarrhoeal Dis Res, 1992. **10**(3): p. 161-3.

32. Ramamurthy, T., et al., *Virulence patterns of Vibrio cholerae non-O1 strains isolated from hospitalised patients with acute diarrhoea in Calcutta, India.* J Med Microbiol, 1993. **39**(4): p. 310-7.

Dhaka University Institutional Repository

33. Russell, R.G., B.D. Tall, and J.G. Morris, Jr., *Non-O1 Vibrio cholerae intestinal pathology and invasion in the removable intestinal tie adult rabbit diarrhea model.* Infect Immun, 1992. **60**(2): p. 435-42.

34. Sharma, C., et al., *Molecular analysis of non-O1, non-O139 Vibrio cholerae associated with an unusual upsurge in the incidence of cholera-like disease in Calcutta, India.* J Clin Microbiol, 1998. **36**(3): p. 756-63.

35. Constantin de Magny, G., et al., *Environmental signatures associated with cholera epidemics.* Proc Natl Acad Sci U S A, 2008. **105**(46): p. 17676-81.

36. Rawlings, T.K., G.M. Ruiz, and R.R. Colwell, *Association of Vibrio cholerae O1 El Tor and O139 Bengal with the Copepods Acartia tonsa and Eurytemora affinis.* Appl Environ Microbiol, 2007. **73**(24): p. 7926-33.

37. Matz, C., et al., *Biofilm formation and phenotypic variation enhance predation-driven persistence of Vibrio cholerae.* Proc Natl Acad Sci U S A, 2005. **102**(46): p. 16819-24.

38. Nalin, D.R., *Cholera, copepods and chitin.* The Lancet, 1976. **308**(7992): p. 958.

39. Pruzzo, C., L. Vezzulli, and R.R. Colwell, *Global impact of Vibrio cholerae interactions with chitin.* Environ Microbiol, 2008. **10**(6): p. 1400-10.

40. Bartlett, D.H. and F. Azam, *Microbiology. Chitin, cholera, and competence.* Science, 2005. **310**(5755): p. 1775-7.

41. Meibom, K.L., et al., *Chitin induces natural competence in Vibrio cholerae.* Science, 2005. **310**(5755): p. 1824-7.

42. Karaolis, D.K., R. Lan, and P.R. Reeves, *The sixth and seventh cholera pandemics are due to independent clones separately derived from environmental, nontoxigenic, non-O1 Vibrio cholerae.* J Bacteriol, 1995. **177**(11): p. 3191-8.

43. Rivera, I.N., et al., *Genotypes associated with virulence in environmental isolates of Vibrio cholerae.* Appl Environ Microbiol, 2001. **67**(6): p. 2421-9.

44. Colwell, R.R., et al., *Serogroup conversion of Vibrio cholerae.* Can J Microbiol, 1995. **41**(10): p. 946-50.

45. Bagchi, K., et al., *Epidemic of diarrhea caused by Vibrio cholerae non-O1 that produced heat-stable toxin among Khmers in a camp in Thailand.* J Clin Microbiol, 1993. **31**(5): p. 1315-7.

46. Dakin, W.P., Howell, D.J., Sutton, R.G., O'Keefe, M.F., Thomas, P., *Gastroenteritis due to non-agglutinable (non-cholera) vibrios.* Med J Aust, 1974. **2**(13): p. 487-490.

47. Dalsgaard A, A.M., Taylor DN, Shimada T, Meza R, Serichantalergs O, Echeverria P., *Characterization of Vibrio cholerae non-O1 serogroups obtained from an outbreak of diarrhea in Lima, Peru.* J Clin Microbiol, 1995. **33**(10): p. 2715-2722.

48. Morris, J.G., Jr., et al., *Experimental non-O group 1 Vibrio cholerae gastroenteritis in humans.* J Clin Invest, 1990. **85**(3): p. 697-705.

49. Mukhopadhyay, A.K., et al., *Distribution and virulence of Vibrio cholerae belonging to serogroups other than O1 and O139: a nationwide survey.* Epidemiol Infect, 1995. **114**(1): p. 65-70.

50. Rudra, S., et al., *Cluster of cases of clinical cholera due to Vibrio cholerae O10 in east Delhi.* Indian J Med Res, 1996. **103**: p. 71-3.

Dhaka University Institutional Repository

51. Lyon, W.J., *TaqMan PCR for detection of Vibrio cholerae O1, O139, non-O1, and non-O139 in pure cultures, raw oysters, and synthetic seawater.* Appl Environ Microbiol, 2001. **67**(10): p. 4685-93.

52. Amann, R., B.M. Fuchs, and S. Behrens, *The identification of microorganisms by fluorescence in situ hybridisation.* Curr Opin Biotechnol, 2001. **12**(3): p. 231-6.

53. Amann, R.I., W. Ludwig, and K.H. Schleifer, *Phylogenetic identification and in situ detection of individual microbial cells without cultivation.* Microbiological reviews, 1995. **59**(1): p. 143-169.

54. Moter, A. and U.B. Gobel, *Fluorescence in situ hybridization (FISH) for direct visualization of microorganisms.* Journal of Microbiological Methods, 2000. **41**(2): p. 85-112.

55. Heidelberg, J.F., K.B. Heidelberg, and R.R. Colwell, *Seasonality of Chesapeake Bay bacterioplankton species.* Appl Environ Microbiol, 2002. **68**(11): p. 5488-97.

56. Heidelberg, J.F., K.B. Heidelberg, and R.R. Colwell, *Bacteria of the gamma-subclass Proteobacteria associated with zooplankton in Chesapeake Bay.* Appl Environ Microbiol, 2002. **68**(11): p. 5498-507.

57. Ivanov, I. and L. Gigova, *RNA colony hybridization method.* Gene, 1986. **46**(2): p. 287-290.

58. Kirschner, A.K., et al., *Interaction of Vibrio cholerae non-O1/non-O139 with copepods, cladocerans and competing bacteria in the large alkaline lake Neusiedler See, Austria.* Microb Ecol, 2011. **61**(3): p. 496-506.

59. Fischer, U.R. and B. Velimirov, *Comparative study of the abundance of various bacterial morphotypes in an eutrophic freshwater environment determined by AODC and TEM.* Journal of microbiological methods, 2000. **39**(3): p. 213-224.

60. Sanborn, W.R., et al., *Fluorescence microscopy for disease diagnosis and environmental monitoring*. 2005: World Health Organization, Regional Office for the Eastern Mediterranean.

Dhaka University Institutional Repository

CHAPTER 3:

VIRULENCE GENE PROFILE AND ANTIBIOTIC RESPONSES OF CLINICAL AND ENVIRONMENTAL *VIBRIO CHOLERAE* ISOLATED FROM COASTAL ENDEMIC FOCI OF BAY OF BENGAL

Dhaka University Institutional Repository

Abstract:

Vibrio cholerae is an autochthonous inhabitant of riverine and estuarine environments and is also a facultative pathogen for humans. We investigated the distribution of the eighteen virulence genetic markers, among 794 clinical and environmental *V. cholerae* samples isolated in Bangladesh between March 2004 and January 2006. While distribution of certain virulence genes, i.e., *hly*A, *hap*, *tox*R, *rtx*A, *vas*A, *vas*K, and *vas*H were found to be similar among the most *V. cholerae* O1, O139 and non-O1/O139, distribution other virulence marker genes markedly differed among toxigenic (O1 and O139), and non-toxigenic (non-O1/O139) *V. cholerae* strains. Most of the toxigenic *V. cholerae* strains contained *ctxA, zot, tcpI*, and *tcpA* genes, whereas only 0- 2% of non- toxigenic *V. cholerae* harbored them, in contrast, none of the toxigenic *V cholerae* harbored either *nag-ST, stn/sto*, or *luxA*, but they were encoded by 3, 2, and 10% of Non-O1/O139 strains respectively. Moreover, none of the *V. cholerae* O1 and O139 contained all four Type 3 secretion system (T3SS) genes, whereas these were encoded by at least 7% of the non-O1/O139 isolates tested. Besides virulence gene profiling, we also examined the prevailing patterns of antibiotic resistance of *V. cholerae* O1 and O139 isolated from this study. Of the six antimicrobial agents tested, 98% of *V. cholerae* O1 were resistant to sulfomethoxazole-trimethoprim (SXT) and furazolidone, while all were sensitive to gentamicin and ciprofloxacin. Isolates of *V. cholerae* O139 tested were uniformly susceptible to all drugs, except 23% of environmental strains resistant to SXT. Between March 2004 and May 2006, resistance to tetracycline and erythromycin increased sharply from 0 to 91% and from 0 to 64% in environmental and clinical strains, respectively. Molecular analysis of tetracycline and erythromycin resistant strains revealed that resurgence of Ogawa biotype were correlated with such increasing prevalence of multi-drug resistant (MDR) strains. Of significance also was the demonstration of similar fingerprinting patterns of MDR strains derived from clinical and environmental sources indicating their environmental origin.

Dhaka University Institutional Repository

Introduction:

Cholera has been one of the worst natural furies of human from time immemorial. The severe intestinal infection in cholera leads to rapidly progressing dehydration, which, without appropriate rehydration, kills half of the affected individuals. Electrolyte replacement therapy through oral and/or intravenous rehydration proves magical in recovering patients from the deadly episode of cholera. In addition, clinically typical cholera patients [1] are routinely treated with 1-3 days course of antibiotic to shorten the illness, and to reduce the rapid water loss due to purging [2, 3] and also the period of vibrio excretion. The antibiotic therapy is very useful not only because it helps the patients to recover early, but it also reduces or delays the immediate necessity of intravenous fluid replacement, which in many cases is not immediately accessible.

More than 200 serogroups have so far been described on the basis of different somatic O antigens, but only serogroups O1 and O139 are able to cause epidemic cholera worldwide [4]. The pathogenicity of serogroups O1 and O139 mostly depends on the combination of two major factors: the expression of the cholera toxin CT encoded by the genome of a filamentous bacteriophage CTXΦ [5], and the ability to adhere to the intestine through the type IV pilus TCP that functions as colonization factor and is encoded by the pathogenicity island VPI [6]. Non-O1/O139 *V. cholerae* strains are part of the normal bacterial flora of estuarine and coastal waters and their microbiological meaning have been neglected for a long time. However, in the past three decades they have been associated with septicemia [7], peritonitis [8], and gastroenteritis [9, 10], through the ingestion of contaminated food or exposure to aquatic environment. Similarly, non-toxigenic *Vibrio cholerae* O1 strains have also been associated with gastroenteritis [11] and localized outbreaks of cholera [12].

A wide range of putative virulence factors has been associated with these strains. These include hemolysin (*hlyA*) [13], non-O1 heat-stable enterotoxin (*nag-ST*) [14], ToxR (*toxR*) [15], zonula occludens toxin (*zot*) [16], actin cross-linking repeats in toxin (*rtxA*) [17], haemagglutinin protease (*hap*) [18], type VI secretion system (T6SS) (*vasA, vasH, vasK*) [19], and type III secretion system (T3SS) (*vcsC, vcsV, vcsN, vspD*) [20]. Occasionally, cholera toxin (*ctxA*) [16] and toxin-coregulated pilus

Dhaka University Institutional Repository

associated genes (*tcpA, tcpI*) [21] can be found also in *V. cholerae* non-O1/O139 strains. Therefore, in order to understand the ecology of the *V. cholerae* serogroups associated with cholera, as well as for a proper outbreak management, it is critical to determine and understand the dynamics of virulence genes carriage among circulating populations of *V. cholerae*.

Transmission of *V. cholerae* is associated with consumption of contaminated water and foods [22, 23]. In spite of the importance and success of rehydration therapy for cholera, antimicrobial therapy still plays a critical role because it can reduce the volume of diarrhea as well as shorten the duration of symptoms and the excretion of Vibrios in the stool [24]. The occurrence of antibiotic-resistant strains of *V. cholerae* is being reported with increasing frequency [23, 25, 26]. However, the antibiotic susceptibilities of organisms spatially and temporally fluctuate. These susceptibilities have to be examined in order to achieve a better understanding of the epidemiological features.

Tetracycline or doxycycline has long been the antibiotic of choice in treating severe cholera in Bangladesh except for young children and pregnant women because only one dose is required [1, 27]. Sulfomethoxazole-trimethoprim (SXT) is the antibiotic of choice for children. Furazolidone is the antibiotic of choice for pregnant women. Erythromycin, gentamycin and chloramphenicol may be used when other antibiotics are not available, or the organism is resistant to them [28]. Most cases of diarrhoea caused by *V. cholerae* can be treated adequately by giving a solution of oral rehydration salts (the WHO/UNICEF standard sachet). During an epidemic, 80-90% of diarrhoea patients can be treated by oral rehydration alone, but patients who become severely dehydrated must be given intravenous fluids. In severe cases, an effective antibiotic can reduce the volume and duration of diarrhoea and the period of vibrio excretion.

For management of cholera, tetracycline is used as the first line drug [29]; however due to emergence of resistance, tetracycline was replaced with erythromycin in many cases. The antibiotic of choice must always be based on the updated culture and sensitivity report. Since the emergence of such resistance amongst *V. cholerae* against antibiotics may significantly influence the future strategies of controlling

Dhaka University Institutional Repository

cholera, continuous monitoring of the epidemic strains is thus crucial. Most of the surveillances for antimicrobial resistance in Bangladesh has been conducted in Dhaka and Matlab region, while the vast majority of the areas especially the coastal periphery of the country remain unexplored. In this study, we evaluated the presence of virulence-associated factors in a collection of 794 *V. cholerae* strains isolated during an epidemiological and ecological surveillance that took place in the provinces of Bakerganj and Mathbaria between 2004 and 2006, as previously described [30]. In addition, a subset (179 isolates) of *V. cholerae* O1 and O139 strains isolated between March 2004 and May 2006 were also subject to antibiotic resistance profiling.

Methods:

A total of 794 *V. cholerae* strains (Table 3.1) isolated during an epidemiological and ecological surveillance that took place in the provinces of Bakerganj and Mathbaria between 2004 and 2006, as previously described [30]. *V. cholerae* O1 strains comprised 102 clinical and 50 environmental isolates, as well as 10 and 31 *V. cholerae* O139 isolates, respectively. Six hundred and one environmental *V. cholerae* non-O1/O139 isolates were also included in the study. In addition, 179 isolates, isolated between March 2004 and May 2006, were included in antibiotic resistance profiling. Sixty-one of these isolates were isolated from patients with acute cholera-like diarrhea admitted to the respective local Thana Health Complexes and the remaining ones (n=118) were collected from the aquatic environments in Mathbaria and Bakergonj, Bangladesh. Hospital surveillance and environmental sampling have been described elsewhere [31].

Virulence gene profiling. Chromosomal DNA was extracted as previously described [32], and PCR reaction was performed in 50 µl reaction mix containing 1 U of GoTaq DNA polymerase (Biorad), according to manufacturer's instructions. Screening of the strains for the presence of the virulence associated factors was carried out by PCR, using positive and negative controls and primers previously described: *ctxA*, *hlyA*, *zot/toxR*, *tcpA*, *tcpI*, *stn/sto* by Rivera et al. [33] , *nag-ST*, *rtxA*, and *hap* by Haley et al [34], bioluminescence gene *luxA* by Grim et al. [35], and *vcsC*, *vcsV*, *vcsN*, *vspD* by Huq et al. [36]. Specific primers for *vasA*, *vasH*, and *vasK* were designed: vasA-104F (GTACGACCGATCCTGACGTT), vasA-446R

Dhaka University Institutional Repository

(ATCTGAATGGTCGTGGCTTC), vasH-857F (GTGGCACGCTATTTCTGGAT), vasH-1242R (TTTCAGCTCACGCACATTTC), vasK-1851F (GCGTCAAATTCAGGAAGAGC), and vasK-2250R (CTGTCCCAGAACCCAACTGT). After amplification, 6 µl of each reaction mixture was subjected to electrophoresis on a 1% agarose gel using an horizontal electrophoresis apparatus (Horizon 11.14, Life Technologies, Gibco BRL). Gels containing amplified DNA were stained with ethidium bromide and visualized using a UV transilluminator. Images were digitized using a 1D Gel documentation system (Bio-Rad).

Areas of Isolation	Year	*V. cholerae* O1				*V. cholerae* O139		*V. cholerae* non-O1/O139[a]
		Clinical		Environmental		Clinical	Environmental	
		Inaba	Ogawa	Inaba	Ogawa			
Mathbaria	2004	19	6	12	0	0	5	162
	2005	6	10	3	2	10	25	220
	2006	1	0	0	0	0	0	8
Bakerganj	2004	21	9	18	1	0	0	72
	2005	10	17	12	2	0	1	131
	2006	1	2	0	0	0	0	8
Total (794)		*58*	*44*	*45*	*5*	*10*	*31*	*601*

[a]All *V. cholerae* non-O1/O139 strains were isolated from environmental samples.

Table 3.1. *V. cholerae* O1, O139 and non-O1/O139 isolated from clinical and environmental samples collected in Bakerganj and Mathbaria between March 2004 and January 2006.

Antibiotic resistance, and genotyping. All of the *V. cholerae* isolates were examined for resistance to vibrio static compound, VSC (150µg); a molecule naturally active against *V. cholerae* and used for identification, tetracycline, TE (30µg), gentamicin, CN (10µg), trimethoprim- sulfamethoxazole, SXT (25µg), furazolidone, FR (100µg), erythromycin, E (15µg), and ciprofloxacin, CIP (5µg) by an agar diffusion method (Bauer *et al.*, 1966) on Mueller-Hinton agar (MHA), using commercial discs (BBL) according to the interpretation criteria recommended by the World health Organization (WHO, 1993). An inoculum dose of about 1.5 x 10^8

Dhaka University Institutional Repository

colony forming units (cfu) per mL (0.5 McFarland standard) of the test strains was used for measuring drug resistance. The randomly selected strains of *V. cholerae* O1 and O139 were examined using a PCR based method for the presence of *ctxA* (gene encoding the B subunit of cholera toxin), *tcpA* (encodes for biosynthesis of the toxin co-regulated pili), *rstR* (encodes for the repressor gene for the CTX prophage) [37-39], SXT element (encodes the resistance markers for trimethoprim sulphomethoxazol, chloramphenicol and low level of streptomycin [29, 40], and *intI1* (encodes class 1 integrons) [29]. Oligonucleotide primers used in the PCR assays; their sequences and the amplicon sizes are shown in Table 3.2.

Pulsed filed gel electrophoresis (PFGE). Intact agarose-embedded chromosomal DNA from isolates of *V. cholerae* was prepared, and PFGE was performed using a contour-clamped homogeneous electric field (CHEFDRII) apparatus (Bio-Rad) by procedures described by [41]. The conditions used for separation were as follows: 2- to10-s pulses for 14 h, 21.79- to 38.38-s pulses for 5 h. An electric field of 6 V/cm was applied at an included field angle of 120°. Genomic DNA was digested with the *NotI* restriction enzyme (GIBCOBRL, Gaithersburg, Md.) at a concentration of 20 U/µl. The restriction fragments were separated in 1% pulsed-field-certi- fied agarose in 0.5XTris-borate-EDTA buffer. In the post electrophoresis-gel-treatment step, the gel was stained and destained. The DNA was visualized with a UV transilluminator, and images of the transilluminator were digitized with a 1D Gel documentation system (Bio-Rad). The images were then processed with Quantity One software (version 4.4.1; Bio- Rad). A known strain of *V. cholerae* (MG-117029) was used as a molecular size marker. The molecular weight of the DNA fragments was determined by using Quantity One software (version 4.4.1; Bio- Rad).

Dhaka University Institutional Repository

Table 3.2. Oligonucleotide primers, sequences, and amplicons used in PCR assays [a]

Primer No.	Target Gene	Primer sequence	Amplicon (bp)
1	*rfb* O1F	GTTTCACTGAACAGATGGG	192
	rfb O1R	GGTCATCTGTAAGTACAAC	
2	*rfb* O139F	AGCCTCTTTATTACGGGTGG	449
	rfb O139R	GTCAAACCCGATCGTAAAGG	
3	*ctx*AF	CTCAGACGGGATTTGTTAGGCACG	302
	*ctx*AR	TCTATCTCTGTAGCCCCTATTACG	
4	*tcpA*F	GAAGAAGTTTGTAAAAGAAGAACAC	472
	*tcpA*R	GAAAGCACCTTCTTTCACGTTG	
5	*RstR1*F	CTTCTCATCAGCAAAGCCTCCATC	500
	*RstA3*R	TCGAGTTGTAATTCATCAAGAGTG	
6	*RstR2*F	GCACCATGATTTAAGATGCTC	500
	*RstA3*R	TCGAGTTGTAATTCATCAAGAGTG	
7	*RstR3*F	CTGTAAATCTCTTCAATCCTAGG	~300
	*RstA3*R	TCGAGTTGTAATTCATCAAGAGTG	
8	*RstR4*F	GTTAACGCTTCAAGCCTG	400
	*RstA3*R	TCGAGTTGTAATTCATCAAGAGTG	
9	*intI1*F	GTTCGGTCAAGGTTCTG	923
	*intI1*F	GCCAACTTTCAGCACATG	
10	*sxt*F	CATCTACCACTTCATAGGCAGC	236
	*sxt*R	CAGCTTAACTCACCAAGGAC	
11	SXT1	ATGGCGTTATCAGTTAGCTGGC	1240
	SXT2	ATTGTCCGAAACGCCAGTTGAAC	
12	SXT3	TCGGGTATCGCCCAAGGGCA	946
	SXT4	GCGAAGATCATGCATAGACC	
13	SXT1	ATGGCGTTATCAGTTAGCTGGC	1035
	SXT4	GCGAAGATCATGCATAGACC	
14	SXT3	TCGGGTATCGCCCAAGGGCA	1151
	SXT2	ATTGTCCGAAACGCCAGTTGAAC	

[a] Primers were for the detection of pathogenic and antibiotic resistance-related genes in *V. cholerae* O1 and O139 strains isolated from two coastal foci of endemic cholera during the year 2004 and 2006 in Bangladesh

Dhaka University Institutional Repository

Results and Discussion:

The virulence profiles for each *V. cholerae* serogroup are summarized in Table 3.3. Out of 152 *V. cholerae* O1 strains, clinical ones showed the typical cholera associated virulence profile characterized by *ctxA*, *zot*, *tcpI*, *tcpA* (74-96%, Table 2), related to an almost ubiquitous presence of T6SS genes *vasA*, *vasK*, *vasH* (87-100%), and *hlyA* and *hap* (85-98%). They were all negative for *nag-ST* and *stn/sto*. Conversely, the virulence profile of the environmental O1 strains was more diversified with most of the strains not encoding *tcpA* and/or *toxR*. Of particular interest were two environmental O1 strains isolated in Bakerganj in 2005 (NIH 825 and NIH 826) that lacked the common cholera associated virulence genes (*ctxA*, *tcpI*, *tcpA*, *zot*) but encoded *hlyA*, *hap*, *rtxA*, as well as T6SS and T3SS associated genes.

All of the O139 strains tested were positive for *tcpI*, *hap*, and T6SS genes *vasA*, *vasK*, and *vasH*. Most of them were also positive for *hlyA*, *ctxA*, *zot*, *toxR*, and *rtxA* (80-100%, Table 2). They were all negative for *nag-ST* and *stn/sto*. Interestingly, three environmental O139 strains isolated in Mathbaria in 2005 encoded the T3SS genes *vcsN*, *vcsV*, *vcsC*, and *vspD*. In general, environmental O139 strains showed a higher variability in their virulence profiles, compared to the clinical O139 strains isolated in Mathbaria, all characterized by the same profile: *hlyA*, *tcpI*, *tcpA*, *ctxA*, *zot*, *toxR*, *hap*, *rtxA*, and T6SS. Interestingly, the only environmental O139 strain isolated in Bakerganj (NIH 599) encoded the same virulence genes as the clinical O139 strains from Mathbaria.

Dhaka University Institutional Repository

Table 3.3. Virulence gene profiles of *V. cholerae* isolates examined in this study.

V. cholerae (n)	*hlyA*	*hap*	*ctxA*	*zot*	*toxR*	*nag-ST*	*stn/sto*	*tcpI*	*tcpA*	*rtxA*	T6SS				T3SS			
											vasA	*vasK*	*vasH*	*luxA*	*vcsN*	*vcsV*	*vcsC*	*vspD*
O1 (152)	82/97[§]	56/57	141/152	146/152	119/152	0/97	0/41	54/57	113/152	148/152	132/152	151/152	152/152	0/41	0/152	2/81	0/57	2/152
	85%	*98%*	*93%*	*96%*	*78%*	*0%*	*0%*	*95%*	*74%*	*97%*	*87%*	*99%*	*100%*	*0%*	*0%*	*2%*	*0%*	*1%*
O139 (41)	34/41	18/18	34/41	33/41	37/41	0/41	0/24	18/18	22/41	39/41	41/41	41/41	41/41	0/24	3/41	2/36	2/18	1/41
	83%	*100%*	*83%*	*80%*	*90%*	*0%*	*0%*	*100%*	*54%*	*95%*	*100%*	*100%*	*100%*	*0%*	*7%*	*6%*	*11%*	*2%*
Non-O1/O139 (601)	478/582	246/251	15/601	13/601	538/601	17/600	6/334	6/257	1/599	527/548	573/601	594/601	593/601	34/334	41/601	54/405	35/270	51/601
	82%	*98%*	*2%*	*2%*	*90%*	*3%*	*2%*	*2%*	*0.20%*	*96%*	*95%*	*99%*	*99%*	*10%*	*7%*	*13%*	*13%*	*8%*
Total (794)	**83%**	**98%**	**24%**	**24%**	**87%**	**2%**	**2%**	**23%**	**17%**	**96%**	**94%**	**99%**	**99%**	**9%**	**6%**	**11%**	**11%**	**7%**

[§]Number of positive strains out of the total tested

% expresses the number of positive strains out of the total tested.

Dhaka University Institutional Repository

A total of 61 clinical isolate of *V. cholerae* O1 and O139 isolated during March 2004 to May 2006 in Mathbaria and Bakergonj, Bangladesh were analyzed. Of them, 84% (n=51) was *V. cholerae* serogroup O1 biogroup El Tor and remaining 16% (n=10) by serogroup O139 (Table 3.4). Among these clinical *V. cholerae* O1 El Tor isolates, 47% (n=24) belonged to Inaba, 49% (n=25) to Ogawa serotype and remaining 4% (n=2) were rough. Conversely, 134 environmental *V. cholerae* O1 and O139 were isolated form various water reservoirs (lake, pond, canal, river etc) of Mathbaria and Bakergonj, which are most frequently used by the nearby villagers for drinking and other domestic purposes due to the high salinity of ground water in this region and 118 isolates were included in this study for analysis. Of the 118 environmental isolates, 70% (n=83) were *V. cholerae* O1 (biogroup El Tor) and the rest 30% (n=35) belonged to serogroup O139 (Table 2) and among the O1 serogroups, 53% (n=44) belonged to serotype Inaba and remaining 47% (n=39) to Ogawa serotype.

Drug resistance was commonly observed among all of the clinical and environmental *V. cholerae* O1 isolates studied. All (n=83) environmental *V. cholerae* O1 strains were uniformly resistant to furazolidone, VSC, and SXT with an alarming 41% (n=34) resistant to tetracycline and erythromycin whereas 23% (n=8) of the *V. cholerae* O139 (n=35), isolated in this period, also showed resistance to both VSC, and SXT but remained completely sensitive to the rest of the other drugs tested. A similar pattern of resistance was also observed in the clinical isolates (n=51) tested. Most *V. cholerae* O1 of patient's origin showed resistance to furazolidone (98%), and SXT (90%) with 41% (n=21) resistant to tetracycline and 35% (n=18) to erythromycin whereas all *V. cholerae* O139 of clinical origin (n=10) were sensitive to all the drugs tested. Sensitivity to ciprofloxacin and chloramphenicol were uniformly observed for all *V. cholerae* O1 and O139 strains of both clinical and environmental origin. Furthermore, we found one strain of *V. cholerae* O1 (NHCM0025) collected from a patient's sample of Mathbaria to be sensitive to all the drugs tested.

Dhaka University Institutional Repository

The non-O1/O139 environmental strains represented the prevalent sample of this collection including 601 strains. However, the virulence associated profiles were not that diversified. Almost all of the strains lacked the typical cholera virulence genes, with only 26 strains showing the presence of *tcpI*, *tcpA*, *ctxA*, or *zot* (0.2-2%). Conversely, most of the strains showed the presence of *hlyA*, *rtxA* and *hap*, as well as T6SS related genes (82-99%). As already observed for the O139 strains, the presence of T3SS related genes (7-13%) was detected starting from 2005 in both Mathbaria and Bakerganj strains. Remarkably, the non-O1/O139 strains were the only ones encoding the non-O1 heat-stable enterotoxin Nag-ST (3%). This is consistent with the data reported so far, being Nag-ST mostly associated with non-O1/O139 *V. cholerae* and *V. mimicus* strains [14, 42, 43], although few reports of *V. cholerae* O1 strains encoding Nag-ST have also been reported [44, 45].

Overall, almost all the strains under analysis encoded *rtxA*, *hap*, and *hlyA* genes, as well as the newly described type VI secretion system [46]. RtxA has been previously correlated with the cytotoxic activity that causes mammalian cells to detach and round up [17], whereas *V. cholerae* hemolysin exhibits vacuolating activity on mammalian cells [47], and HapA acts on tight junction-associated proteins in the epithelial cells [48]. The combination of these virulence factors might be responsible for the occasional diarrhea caused by non-toxigenic non-O1/O139 strains. Furthermore, the association of these virulence factors with CT, as observed for few non-O1/O139 strains isolated in Mathbaria, might represent a warning for the emergence of new toxigenic clones in the environment. A remarkable example is given by strain NIH 831 isolated in Bakerganj in 2005 that encodes all the virulence factors under analysis but T3SS, virtually sharing the same pathogenicity profile usually associated with toxigenic O1 strains. The appearance of toxigenic non-O1/O139 *V. cholerae* strains is consistent with the identification of new pathogenic clones in both clinical and environmental samples worldwide [33, 49, 50], and further supports the epidemiological link between the environmental reservoir and human infection. New variants may lead to epidemic outbreaks of cholera in Bangladesh and, hence, continuous monitoring of *V. cholerae* is essential in the context of public health.

Dhaka University Institutional Repository

Table: 3.4. Antibiotic resistance of *V. cholerae* O1 biotype El Tor serotype Inaba and Ogawa isolated from environmental and clinical samples between March 2004 to May 2006 from Mathbaria and Bakergonj, in Bangladesh.

Area / Year	Source of Isolation	Bakergonj									Mathbaria								
		Vibrio cholerae		Resistant (%)							*Vibrio cholerae*		Resistant (%)						
		Biotype	Serotype	VSC	E	CN	SXT	TE	CIP	FR	Biotype	Serotype	VSC	E	CN	SXT	TE	CIP	FR
2004	Environment	O1 (n=16)		100	0	0	100	0	0	100	O1 (n=12)		100	0	0	100	0	0	100
			INET (n=16)	100	0	0	100	0	0	100		INET (n=12)	100	0	0	100	0	0	100
			OGET (n=0)	NIF								OGET (n=0)	NIF						
		O139 (n=0)		NIF							O139 (n=7)		0	0	0	0	0	0	0
	Clinical	O1 (n=11)		NT	0	0	91	0	0	91	O1 (n=8)		NT	0	0	87	0	0	87
			INET (n=6)	NT	0	0	55	0	0	55		INET (n=5)	NT	0	0	100	0	0	100
			OGET (n=5)	NT	0	0	80	0	0	80		OGET (n=3)	NT	0	0	67	0	0	67
		O139 (n=0)		NIF							O139 (n=0)		NIF						
2005	Environment	O1 (n=17)		100	12	0	100	12	0	100	O1 (n=5)		100	40	0	100	40	0	100
			INET (n=14)	100	0	0	100	0	0	100		INET (n=2)	100	0	0	100	0	0	100
			OGET (n=3)	100	67	0	100	67	0	100		OGET (n=3)	100	67	0	100	67	0	100
		O139 (n=1)		0	0	0	0	0	0	0	O139 (n=27)		30	0	0	30	0	0	0
	Clinical	O1 (n=14)		NT	60	0	86	57	0	100	O1 (n=4)		NT	0	0	100	0	0	100
			INET (n=7)	NT	14	0	71	14	0	100		INET (n=2)	NT	0	0	100	0	0	100
			OGET (n=7)	NT	86	0	100	100	0	100		OGET (n=2)	NT	0	0	100	0	0	100
		O139 (n=0)		NIF							O139 (n=3)		NT	0	0	0	0	0	0
2006	Environment	O1 (n=14)		100	100	0	100	100	0	100	O1 (n=19)		100	84	0	100	84	0	100
			INET (n=0)	NIF								INET (n=0)	NIF						
			OGET (n=14)	100	100	0	100	100	0	100		OGET (n=19)	100	84	0	100	84	0	100
		O139		NIF							O139		NIF						
	Clinical	O1 (n=4)		NT	0	0	50	25	0	100	O1 (n=10)		NT	90	0	100	100	0	100
			INET (n=3)	NT	0	0	33	0	0	100		INET (n=1)	NT	100	0	100	100	0	100
			OGET (n=1)	NT	0	0	100	100	0	100		OGET (n=9)	NT	89	0	100	100	0	100
		O139 (n=0)		NIF							O139 (n=0)		NIF						

Numbers in parenthesis indicate the number of strains isolated, NIF, No Isolates Found, NT, Not Tested, VSC, Vibrio Static Compound, E, Erythromycin, TE, Tetracycline, SXT, Sulfamethoxazole Trimethoprim, CN, Gentamycin, CIP, Ciprofloxacin, FR, Furazolidone.

Dhaka University Institutional Repository

A marked difference was observed among *V. cholerae* O1 in both clinical and environmental origins in terms of their year of isolation and resistance to tetracycline and erythromycin (Figure 3.1). In Bakergonj, all environmental strains isolated in 2004 were sensitive to tetracycline and erythromycin. Among the 2005 isolates, 12% demonstrated resistance to these two drugs, and in 2006, all isolates (100%) became resistant to tetracycline and erythromycin. Similar trends of resistance towards tetracycline, and erythromycin were also observed among the clinical isolates from Bakegonj (Table 3.4). However, the only O139 strain isolated from Bakergonj was sensitive to all the drugs tested. Compared to the 2004 and 2006 data among the environmental strains of Mathbaria, resistance of *V. cholerae* O1 towards tetracycline, and erythromycin had increased from 0 to 84% with 40% resistance in 2005. Interestingly resistance to tetracycline and erythromycin among the clinical strains in Mathbaria was only observed in 2006 and with alarming rates (100%, 90% respectively) whereas it has been observed among the environmental strain in 2005 (Table 3.4). Moreover, most of the *V. cholerae* O139 isolated from Mathbaria were sensitive to all the drugs except for eight (30%) out of 27 environmental isolates of 2005 that were resistant to SXT. Occurrence of Ogawa biotype has significantly increased in 2005 and 2006 in both clinical (42 to 44% and 44 to 64%) and environmental (0 to 27% and 27 to 100%) cases. However, it is evident that the resurgence of Ogawa biotype at higher frequency in both clinical cases and aquatic environments of Mathbaria and Bakergonj mostly contributed these increasing trends of resistance toward tetracycline and erythromycin (Figure 3.1). All of the strains resistant to tetracycline and erythromycin, except for two rough and two Inaba biotypes, belonged to Ogawa biotype in both clinical and environmental cases in Mathbaria and Bakergonj. On the other hand, all (n=7) of the Ogawa biotypes isolated in 2004 from clinical cases were susceptible to tetracycline and erythromycin. No Ogawa biotype was found from environmental samples in 2004 (Table 3.4).

Tetracycline is the drug of choice for the treatment of cholera and erythromycin has been used when other antibiotics are not available, or the organism is resistant to them. Thus, the prevalence of endemic cholera in this region and the consequent use of tetracycline and erythromycin may explain the emergence of resistance of *V. cholerae*

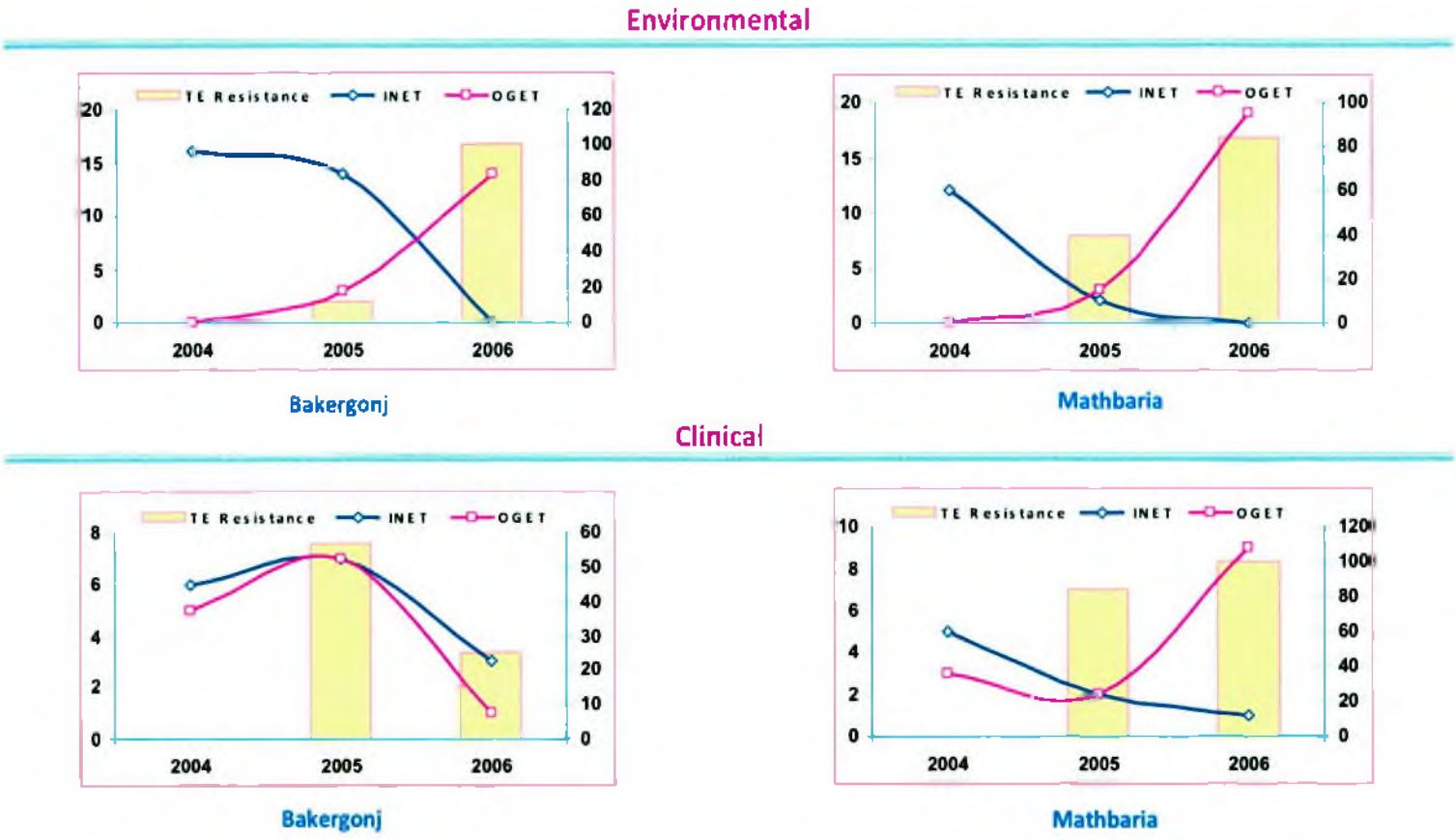
Dhaka University Institutional Repository

towards these antimicrobial agents. The extensive prophylactic use of tetracycline and erythromycin may have triggered the rapid emergence and spread of tetracycline and erythromycin resistant *V. cholerae* O1 strains in Bangladesh. Furthermore, the association between the development of resistance to tetracycline and co-trimoxazole with large-scale of antibiotics for treatment and prophylaxis of cholera is already well recognized [51-53]. Still our demonstration of multiple drug resistant *V. cholerae* O1 isolates showing resistance to all the antibiotics traditionally used to treat cholera is disturbing and has a direct impact on the treatment of current and future cholera in Bangladesh and other countries to which this isolates may spread. Therefore, the high degree of tetracycline and erythromycin resistance in Dhaka and Matlab, where the ICDDRB hospitals are located, could be due to easier access to drugs in the capital than in the other areas but their simultaneous emergence and spread in the remote areas of Bangladesh like Bakergonj and Mathbaria, which is far away from the capital (~ 300km) and in vicinity of the Bay of Bengal remains a potential threat for the antimicrobial therapy of such diseases.

PCR analysis of the selected environmental and clinical strains revealed that all had important epidemic markers, *rfb*O1, *ctx*A, *rstR*2, and *tcpA* El Tor with no differences between erythromycin and tetracycline resistance and sensitive strains in regards to these epidemic markers tested. PCR analysis of the *V. cholerae* O1 did not show any hybrids, i.e., emerging varieties of *V. cholerae* O1 El Tor biotype with attributes of the classical biotype, but rather the typical seventh pandemic El Tor strains.

Dhaka University Institutional Repository

Figure 3.1: Resistance to Tetracycline and Erythromycin linked to the resurgence of *V. cholerae* O1 biotype Ogawa El Tor (OGET) replacing Inaba El Tor (INET). Primary X-axis represents the number of isolates belonging to OGET and INET biotype, whereas secondary X-axis represents % of isolates resistant to Tetracycline and Erythromycin.



Dhaka University Institutional Repository

Table 3.5: Cumulative presentation of resistant pattern, plasmid analysis, and genetic screening of *V. cholerae* O1 and O139 isolated from environmental and clinical samples from two coastal foci of endemic cholera, in Bangladesh.

Year	*Source*	*Serotype*	*Biotypes[a] (n)*	*Susceptibility Pattern (n)*	*rstR*				*ctxA*	*tcpA*	*SXT fragment[b]*				*IntI1*
					I	II	III	IV			1	II	III	IV	
2004	Environment (n=10)	O1 (n=8)	OG (1), IN (7)	*SXT^R, FR^R (7), $SXT^RFR^RTE^RE^R$ (1)*	0	8	0	0	8	**El Tor**	5	8	7	7	0
		O139 (n=2)		*All sensitive (2)*	0	2	0	0	2		1	1	2	I	0
	Clinical (n=9)	O1 (n=9)	OG (5), IN (4)	*All sensitive (1), SXT^RFR^R (7), $SXT^RFR^RTE^RE^R$ (1)*	0	9	0	0	9	**El Tor**	8	8	8	8	0
		O139 (n=0)		-	0	0	0	0	0		0	0	0	0	0
2005	Environment (n=9)	O1 (n=5)	OG (4), IN (1)	*All sensitive (1), SXT^RFR^R (2), $SXT^RFR^RTE^RE^R$ (2)*	0	5	0	0	5	**El Tor**	4	5	3	4	0
		O139 (n=4)		*All sensitive (1), SXT^R (3)*	0	4	0	0	4		3	4	3	4	0
	Clinical (n=21)	O1 (n=11)	OG (9), IN (2)	*SXT^RFR^R (5), $SXT^RFR^RTE^RE^R$ (6)*	0	11	0	0	11	**El Tor**	10	11	11	11	0
		O139 (n=10)		*SXT^RFR^R (10)*	0	10	0	0	10		10	10	10	10	0
2006	Environment (n=7)	O1 (n=7)	OG (7), IN (0)	*SXT^RFR^R (1), $SXT^RFR^RTE^RE^R$ (6)*	0	7	0	0	7	**El Tor**	6	6	7	7	0
		O139 (n=0)		-	0	0	0	0	0		0	0	0	0	0
	Clinical (n=6)	O1 (n=6)	OG (2), IN (3) R (1)	*SXT^R (1), SXT^RFR^R (1), $SXT^RFR^RTE^R$ (1), $SXT^RFR^RTE^RE^R$ (3)*	0	6	0	0	6	**El Tor**	6	6	6	6	0
		O139 (n=0)		-	0	0	0	0	0		0	0	0	0	0

a: OG, Ogawa biotype, IN, Inaba biotype, R, Rough strains; b: SXT PCR has been done using four combinations of primers (Table 3.2)

Dhaka University Institutional Repository

A correlation between antimicrobial resistance and the presence of integrons was explored by PCR aiming to identify class I integron which are specific for the 5' conserved segment containing integrase gene (*intI*1) [29]. PCR within *Int*-U and *Int*-D primers did not yield amplicons for any of the 62 environmental and clinical isolates but yielded desired amplicons for the positive control (Table 3.5). Thus, the 62-kb self-transmissible transposons like SXT element described by Waldor *et al.* [54], does not seem to carry class I integrons. The *V. cholerae* O1 El Tor strains, which re-emerged in India and Bangladesh after the O139 epidemic, were resistant to sulfamethoxazole, trimethoprim, and streptomycin [25, 55]. Waldor *et al.* [54] showed that these O1 strains also contained the SXT element. So, it remained to be determined whether the O1 strains from this region contains class I integrons, and this study concluded that Class I integrons were absent in clinical and environmental *V. cholerae* isolates from Bangladesh during this study period. Next, we investigated the presence of SXT element, which has been reported to be responsible for resistance to chloramphenicol, streptomycin, sulfamethoxazole and trimethoprim in *V. cholerae* [56]. Recently Hochhut *et al.* [56] reported that deletion of the antibiotic resistance genes located on the SXT element rendered these strains sensitive to the antibiotics. Complete SXT element has been detected in most of the *V. cholerae* O139 isolates tested, whereas a few of the environmental strains failed to amplify all targeted regions although most of them were sensitive to SXT. Moreover, one environmental O139 strain (EM 188), which was sensitive to all the drugs, did not contain SXT element. Among the *V. cholerae* O1, 80% (n=37) of the strains contained the complete SXT element with some (15%) of them having some deletion and still others (4%), which were sensitive to all the drugs tested, did not contain any SXT element. We also noticed that one *V. cholerae* O1 (NHCB0714), harboring the SXT element, was sensitive to all antibiotics examined. The frequent presence of the SXT element in environmental isolates suggests the possibility that the transmission of SXT element takes place in the environmental water where *V. cholerae* resides as part of the microbiota. On the other hand, many of the isolates, mostly *V. cholerae* O139, were sensitive to these drugs, but we analyzed for the presence of SXT element in them, as they might contain an element devoid of an antibiotic gene cluster.

Dhaka University Institutional Repository

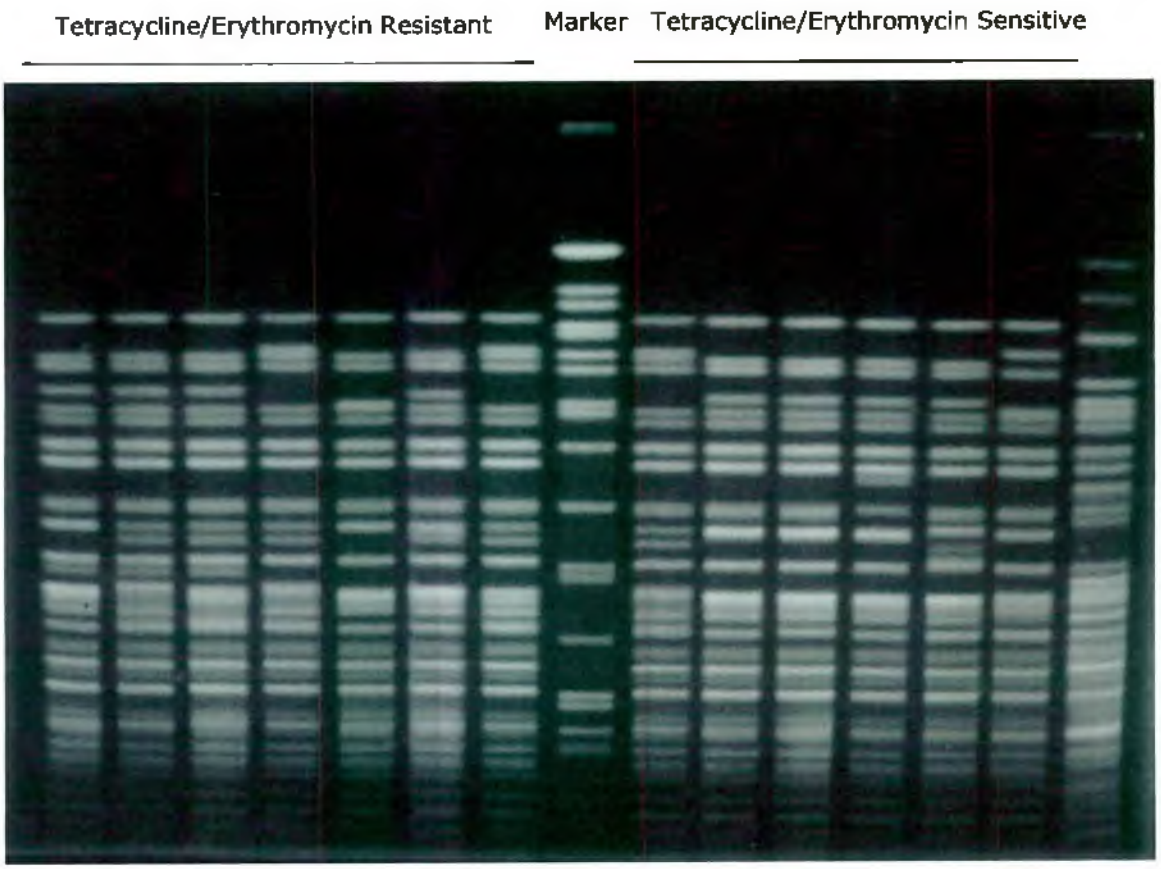


Figure 3.2: Pulsed Field Gel Electrophoresis (PFGE) of *Not*I-digested genomic DNA revealed the strains to be of similar clonal lineage.

Pulsed-field gel electrophoresis (PFGE) of *Not*I digested chromosomal DNA of the *V. cholerae* strains yielded 16 to 19 reproducible DNA fragments ranging in size approximately from <20 to <668.9 Kbp (Figure 3.2). DNA fingerprinting analysis revealed the strains to be of similar clonal lineage, although no genetic clue could be found between tetracycline and erythromycin resistance and sensitive strains.

Dhaka University Institutional Repository

Conclusion:

In this study, the occurrence and distribution of a large number of virulence-associated genes in environmental and clinical strains of *V. cholerae*, isolated in two coastal areas of Bay of Bengal, were demonstrated. The vast majority of environmental *V. cholerae* strains were non-O1/O139, which carry a discrete set of virulence genes; infrequently carried by the toxigenic *V. cholerae* strains, and occasionally carry the genes for the major virulence factors, typically carried by the toxigenic *V. cholerae* strains. This supports the hypothesis that some pathogenic bacteria have evolved from nonpathogenic strains of the same species via horizontal transfer of virulence genes [57] as virulence genes are dispersed among environmental strains of *V. cholerae* and may be ferried about, given the fact that most of the virulence genes are located on mobile elements. A remarkable example is given by strain NIH 831 encoding all the virulence factors under analysis but T3SS, virtually sharing the same pathogenicity profile usually associated with toxigenic O1 strains. Our findings indicate that horizontal transfer of genes has occurred not only between O1 biotypes but also between O1 and non-O1/O139 strains, and stress the need to monitor the emergence and prevalence of non-O1/O139 strains, which not only are a source of morbidity as agents of sporadic disease but also represent a threat, since they might become potentially epidemic. The study also demonstrates that, neither erythromycin nor tetracycline, the two first-line drugs used in Bangladesh, should not be recommended any longer for treating severe cases of cholera. However, clinicians noted less of a clinical response from ciprofloxacin than would have been expected from the findings of studies done some years ago [58, 59]. Therefore, a prudent national policy needs to be adapted to a) abandon any systematic chemoprophylaxis, b) advise only oral rehydration therapy for mild to- moderate cases, c) reserve antibiotic therapy only for severe illness [28] and d) replace tetracycline & erythromycin with other potential drug in treating severe cholera. These measures against the cholera epidemic should be accompanied by general reinforcement of microbiologic surveillance to monitor antibiotic resistance so that the country can respond effectively to any future bacterial epidemics.

Dhaka University Institutional Repository

References:

1. Sack, D.A., et al., *Cholera.* Lancet, 2004. **363**(9404): p. 223-33.
2. Sack, D.A., et al., *Single-dose doxycycline for cholera.* Antimicrob Agents Chemother, 1978. **14**(3): p. 462-4.
3. Lindenbaum, J., W.B. Greenough, and M.R. Islam, *Antibiotic therapy of cholera.* Bull World Health Organ, 1967. **36**(6): p. 871-83.
4. Kaper, J.B., J Glenn Morris JR, and M.M. Levine, *Cholera.* Clin. Microbiol. Rev., 1995. **8**(1): p. 48–86.
5. Waldor, M.K., H. Tschape, and J.J. Mekalanos, *A new type of conjugative transposon encodes resistance to sulfamethoxazole, trimethoprim, and streptomycin in Vibrio cholerae 0139.* J. Bacteriol., 1996. **178**(14): p. 4157-4165.
6. Karaolis, D.K.R., et al., *A Vibrio cholerae pathogenicity island associated with epidemic and pandemic strains.* Proc. Natl. Acad. Sci. U.S.A., 1998. **95**(6): p. 3134-3139.
7. Briceño, L.I., et al., *Non-toxigenic hemolytic Vibrio cholerae non-01 non-0139 fatal septicemia. Report of one case.* Rev Med Chil, 2009. **137**(9): p. 1193-1196.
8. Ichinose, Y., et al., *The characterization of Vibrio cholerae non-01 strain causing perforative pan-peritonitis.* Kansenshogaku Zasshi. The Journal of the Japanese Association for Infectious Diseases, 2001. **75**(8): p. 696-702.
9. Wilson, R., et al., *Non-O group 1 Vibrio cholerae gastroenteritis associated with eating raw oysters.* Am J Epidemiol, 1981. **114**(2): p. 293-298.
10. Morris, J.J., et al., *Experimental non-O group 1 Vibrio cholerae gastroenteritis in humans.* J Clin Invest, 1990. **85**(3): p. 697-705.
11. Morris, J.G., et al., *Isolation of nontoxigenic Vibrio cholerae O group 1 from a patient with severe gastrointestinal disease.* J Clin Microbiol, 1984. **19**(2): p. 296-297.
12. Saha, P.K., et al., *Nontoxigenic Vibrio cholerae 01 serotype Inaba biotype El Tor associated with a cluster of cases of cholera in southern India.* J Clin Microbiol, 1996. **34**(5): p. 1114-7.
13. Zhang, X.H. and B. Austin, *Haemolysins in Vibrio species.* J Appl Microbiol, 2005. **98**(5): p. 1011-1019.
14. Dalsgaard, A., et al., *Characterization of Vibrio cholerae non-01 serogroups obtained from an outbreak of diarrhea in Lima, Peru.* J Clin Microbiol, 1995. **33**(10): p. 2715-22.
15. Waldor, M.K. and J.J. Mekalanos, *ToxR regulates virulence gene expression in non-01 strains of Vibrio cholerae that cause epidemic cholera.* Infect. Immun., 1994. **62**(1): p. 72-78.
16. Jiang, S., W. Chu, and W. Fu, *Prevalence of cholera toxin genes (ctxA and zot) among non-01/0139 Vibrio cholerae strains from Newport Bay, California.* Appl Environ Microbiol, 2003. **69**(12): p. 7541-7544.
17. Lin, W., et al., *Identification of a Vibrio cholerae RTX toxin gene cluster that is tightly linked to the cholera toxin prophage.* Proc Natl Acad Sci U S A, 1999. **96**(3): p. 1071-1076.

Dhaka University Institutional Repository

18. Silva, A.J., et al., *Contribution of hemagglutinin/protease and motility to the pathogenesis of El Tor biotype cholera.* Infect Immun, 2006. **74**(4): p. 2072-2079.
19. Unterweger, D., et al., *Constitutive type VI secretion system expression gives Vibrio cholerae intra- and interspecific competitive advantages.* PLoS ONE, 2012. **7**(10): p. e48320.
20. Dziejman, M., et al., *Genomic characterization of non-O1, non-O139 Vibrio cholerae reveals genes for a type III secretion system.* Proc Natl Acad Sci U S A, 2005. **102**(9): p. 3465-3470.
21. Sarkar, A., et al., *Vibrio Pathogenicity Island and cholera toxin genetic element-associated virulence genes and their expression in non-O1 non-O139 strains of Vibrio cholerae.* Infect Immun, 2002. **70**(8): p. 4735-4742.
22. Pitrak, D.L. and J.D. Gindorf, *Bacteremic cellulitis caused by non-serogroup O1 Vibrio cholerae acquired in a freshwater inland lake.* J Clin Microbiol, 1989. **27**(12): p. 2874-6.
23. Radu, S., et al., *Molecular characterization of Vibrio cholerae O1 and non-O1 from human and environmental sources in Malaysia.* Epidemiol Infect, 1999. **123**(2): p. 225-32.
24. Mahalanabis, D., A. Molla, and D.A. Sack, *Clinical management of cholera.* Cholera. New York: Plenum, 1992: p. 253-83.
25. Bag, P.K., et al., *Rapid spread of the new clone of Vibrio cholerae O1 biotype El Tor in cholera endemic areas in India.* Epidemiol Infect, 1998. **121**(2): p. 245-51.
26. Dalsgaard, A., et al., *A high proportion of Vibrio cholerae strains isolated from children with diarrhoea in Bangkok, Thailand are multiple antibiotic resistant and belong to heterogenous non-O1, non-O139 O-serotypes.* Epidemiol Infect, 1999. **122**(2): p. 217-26.
27. Greenough, W.B., 3rd, et al., *Tetracycline in the Treatment of Cholera.* Lancet, 1964. **1**(7329): p. 355-7.
28. Control, W.H.O.G.T.F.o.C., *Guidelines for cholera control.* 1993: World Health Organization.
29. Dalsgaard, A., et al., *Vibrio cholerae O1 outbreak isolates in Mozambique and South Africa in 1998 are multiple-drug resistant, contain the SXT element and the aadA2 gene located on class 1 integrons.* J Antimicrob Chemother, 2001. **48**(6): p. 827-38.
30. Alam, M., et al., *Seasonal cholera caused by Vibrio cholerae serogroups O1 and O139 in the coastal aquatic environment of Bangladesh.* Appl. Environ. Microbiol., 2006. **72**(6): p. 4096-4104.
31. Alam, M., et al., *Seasonal cholera caused by Vibrio cholerae serogroups O1 and O139 in the coastal aquatic environment of Bangladesh.* Appl Environ Microbiol, 2006. **72**(6): p. 4096-104.
32. Ausubel, F.M., et al., *Current protocols in molecular biology*, ed. J.W.a. Sons. 1990, New York: Green Publishing Associates and Wiley.
33. Rivera, I.N.G., et al., *Genotypes associated with virulence in environmental isolates of Vibrio cholerae.* Appl Environ Microbiol, 2001. **67**(6): p. 2421-2429.

Dhaka University Institutional Repository

34. Haley, B.J., et al., *Vibrio cholerae in a historically cholera-free country.* Environ Microbiol Rep, 2012. **4**(4): p. 381-389.
35. Grim, C.J., et al., *Occurrence and expression of luminescence in Vibrio cholerae.* Appl Environ Microbiol, 2008. **74**(3): p. 708-715.
36. Huq, A., et al., *Detection, isolation, and identification of Vibrio cholerae from the environment,* in *Curr Protoc Microbiol.* 2012, John Wiley & Sons, Inc.
37. Nair, G.B., et al., *New variants of Vibrio cholerae O1 biotype El Tor with attributes of the classical biotype from hospitalized patients with acute diarrhea in Bangladesh.* J Clin Microbiol, 2002. **40**(9): p. 3296-9.
38. Nusrin, S., et al., *Peruvian Vibrio cholerae O1 El Tor strains possess a distinct region in the Vibrio seventh pandemic island-II that differentiates them from the prototype seventh pandemic El Tor strains.* J Med Microbiol, 2009. **58**(Pt 3): p. 342-54.
39. Hoshino, K., et al., *Development and evaluation of a multiplex PCR assay for rapid detection of toxigenic Vibrio cholerae O1 and O139.* FEMS Immunol Med Microbiol, 1998. **20**(3): p. 201-7.
40. Bhanumathi, R., et al., *Molecular characterization of Vibrio cholerae O139 bengal isolated from water and the aquatic plant Eichhornia crassipes in the River Ganga, Varanasi, India.* Appl Environ Microbiol, 2003. **69**(4): p. 2389-94.
41. Talukder, K.A., et al., *Phenotypic and genotypic characterization of provisional serotype Shigella flexneri 1c and clonal relationships with 1a and 1b strains isolated in Bangladesh.* J Clin Microbiol, 2003. **41**(1): p. 110-7.
42. Vicente, A.C.P., A.M. Coelho, and C.A. Salles, *Detection of Vibrio cholerae and V. mimicus heatstable toxin gene sequence by PCR.* J Med Microbiol, 1997. **46**(5): p. 398-402.
43. Sarkar, B., et al., *Preferential association of the heat-stable enterotoxin gene (stn) with environmental strains of Vibrio cholerae belonging to the O14 serogroup.* Epidemiol Infect, 2002. **129**(2): p. 245-251.
44. Takeda, T., et al., *Detection of heat-stable enterotoxin in a cholera toxin gene-positive strain of Vibrio cholerae O1.* FEMS Microbiol Lett, 1991. **64**(1): p. 23-27.
45. Mallard, K.E. and P.M. Desmarchelier, *Detection of heat-stable enterotoxin genes among Australian Vibrio cholerae O1 strains.* FEMS Microbiol Lett, 1995. **127**(1-2): p. 111-115.
46. Pukatzki, S., et al., *Identification of a conserved bacterial protein secretion system in Vibrio cholerae using the Dictyostelium host model system.* Proc Natl Acad Sci U S A, 2006. **103**(5): p. 1528-1533.
47. Figueroa-Arredondo, P., et al., *Cell vacuolation caused by Vibrio cholerae hemolysin.* Infect Immun, 2001. **69**(3): p. 1613-1624.
48. Wu, Z., P. Nybom, and K.-E. Magnusson, *Distinct effects of Vibrio cholerae haemagglutinin/protease on the structure and localization of the tight junction-associated proteins occludin and ZO-1.* Cell Microbiol, 2000. **2**(1): p. 11-17.

Dhaka University Institutional Repository

49. Bakhshi, B., et al., *A molecular survey on virulence associated genotypes of non-O1 non-O139 Vibrio cholerae in aquatic environment of Tehran, Iran.* Water Research, 2009. **43**(5): p. 1441-1447.
50. Bakhshi, B., et al., *Presence of CTX gene cluster in environmental non-O1/O139 Vibrio cholerae and its potential clinical significance.* Indian J Med Microbiol, 2012. **30**(3): p. 285-289.
51. Tabtieng, R., et al., *An epidemic of Vibrio cholerae el tor Inaba resistant to several antibiotics with a conjugative group C plasmid coding for type II dihydrofolate reductase in Thailand.* Am J Trop Med Hyg, 1989. **41**(6): p. 680-6.
52. Glass, R.I., et al., *Plasmid-borne multiple drug resistance in Vibrio cholerae serogroup O1, biotype El Tor: evidence for a point-source outbreak in Bangladesh.* J Infect Dis, 1983. **147**(2): p. 204-9.
53. Finch, M.J., et al., *Epidemiology of antimicrobial resistant cholera in Kenya and East Africa.* Am J Trop Med Hyg, 1988. **39**(5): p. 484-90.
54. Waldor, M.K., H. Tschape, and J.J. Mekalanos, *A new type of conjugative transposon encodes resistance to sulfamethoxazole, trimethoprim, and streptomycin in Vibrio cholerae O139.* J Bacteriol, 1996. **178**(14): p. 4157-65.
55. Mukhopadhyay, A.K., et al., *Biotype traits and antibiotic susceptibility of Vibrio cholerae serogroup O1 before, during and after the emergence of the O139 serogroup.* Epidemiol Infect, 1995. **115**(3): p. 427-34.
56. Hochhut, B., et al., *Molecular analysis of antibiotic resistance gene clusters in vibrio cholerae O139 and O1 SXT constins.* Antimicrob Agents Chemother, 2001. **45**(11): p. 2991-3000.
57. Covacci, A., et al., *Did the inheritance of a pathogenicity island modify the virulence of< i> Helicobacter pylori?</i>.* Trends in microbiology, 1997. **5**(5): p. 205-208.
58. Khan, W.A., et al., *Randomised controlled comparison of single-dose ciprofloxacin and doxycycline for cholera caused by Vibrio cholerae 01 or 0139.* Lancet, 1996. **348**(9023): p. 296-300.
59. Khan, A.M., et al., *Tetracycline in the treatment of cholera caused by Vibrio cholerae O1 resistant to the drug in vitro.* J Health Popul Nutr, 2003. **21**(1): p. 76-8.

Dhaka University Institutional Repository

CHAPTER 4:

INFLUENCE OF ENVIRONMENT FACTORS IN OCCURRENCE AND DISTRIBUTION OF *VIBRIO CHOLERAE*

Dhaka University Institutional Repository

Abstract:

For centuries cholera has been one of the most feared diseases and *Vibrio cholerae*, a water-borne pathogen that is the causative agent of this severe watery diarrheal disease. It is the only enteric pathogen that has the potential to produce pandemics of disease and a major public health threat for developing countries. Since *Vibrio cholerae* is autochthonous to the aquatic environment, it is not possible to eradicate the disease but prevention is an achievable goal. The study, therefore, sought to identify factors involved in in the occurrence and transmission of cholera in two costal foci of endemic cholera in Bangladesh during February 2005 to September 2007, during which both clinical and environmental samples were collected at biweekly or monthly intervals. The results of environmental sampling and analysis of the environmental and clinical data have revealed significant correlations of water temperature, water depth, salinity, and conductivity, with the occurrence of toxigenic *V. cholerae* in the environment as well as cases of cholera. In addition, significant association was observed between occurrence of cholera cases caused by *V. cholerae* O1 and presence of *V. cholerae* O1, *ctxA*, or a combination of *V. cholerae* O1 and *ctxA* in the environment without any time lag or with a lag of two to four weeks before appearance of cholera in Mathbaria and Bakerganj, respectively. In conclusion, the data gathered in this study further support the findings of numerous prior studies that environmental factors determine cholera epidemicity and that both quantitative and qualitative (genomic) data on *Vibrio* abundance in the environment are critical to estimate the risk of disease.

Introduction:

Infectious diseases remain the leading cause of death in the world. Many factors have contributed to the persistence and increase in the occurrence of infectious diseases, such as societal changes, health care, food production, human behaviour, environmental change, public health infrastructure and microbial adaptation. Many diseases related to environmental factors have recently emerged worldwide and are of a degree to raise serious concerns. Over the past few decades unprecedented rates of change in diversity of non-human biota have coincided with the emergence and reemergence of numerous infectious diseases around the world. Virtually all of the

Dhaka University Institutional Repository

world's terrestrial and aquatic communities and ecosystems have undergone dramatic changes in biodiversity and biocomplexity. Thus the coincidence of broad scale environmental changes and emergence of infectious diseases may point to underlying predictable ecological relationships. The role of biological diversity and environment structure has received a great deal of attention from ecologists in recent years as alterations to aquatic environments via natural or anthropogenic factors can directly or indirectly lead to infectious disease outbreaks. At present, little information is available on the molecular ecology and evolutionary dynamics of infectious agents in the environment. Thus understanding environmental factors involved in persistence, growth, and transmission of pathogens and the response of organisms to environmental changes will provide information about the emergence or re-emergence of infectious diseases. *Vibrio cholerae*, the causative agent of the Asiatic cholera, is an excellent model system to study these processes because *V. cholerae* causes periodic seasonal outbreaks in regions where it is an established member of the indigenous aquatic flora. The capacity of the organism to cause seasonal epidemics is linked to its survival in free-living and in viable but non-culturable or biofilm growth modes in aquatic environments between outbreaks. However, the processes governing the survival of *V. cholerae* and its adaptability to changes in habitat remain elusive.

Cholera remains a major public health threat in developing countries where water and sanitation facilities are inadequately available. The World Health Organization estimates that annually, about 3 to 5 million people are affected worldwide by cholera and over one hundred thousand cases result in death. *V. cholerae* is a natural inhabitant of the aquatic environment [1] which with evidence of new biotypes emerging from its environmental reservoir [2, 3], indicate that cholera bacteria cannot be eradicated. Therefore, cholera outbreaks are likely to occur over time and the most effective means of minimizing or preventing the disease is to minimize exposure to pathogenic strains and/or high concentration of cholera bacteria. Cholera is one of the most prevalent water-related infections in many regions of the world, specifically in South Asia, Sub-Saharan Africa, and Latin America. Observational records show that the vast majority of cholera outbreaks originate in coastal regions, indicating strong association between the environment and the disease [4-6]. Despite significant advances in our knowledge of the metabolism, pathogenic mechanism, and the genomics of *V. cholerae*, we still cannot predict precisely when

Dhaka University Institutional Repository

the next cholera epidemic will occur, its probability, timing, or location of the outbreak, all of which is essential to design and implement an effective intervention strategy. The recent outbreak of cholera in Haiti that has affected millions of people is a classic result of having a limited understanding of the disease dynamics and relationship of *V. cholerae* with its environment.

It is now recognized that *V. cholerae* is a component of coastal and estuarine microbial ecosystems, with the copepod species of zooplankton that comprise the aquatic fauna of rivers, bays, estuaries and the open ocean serving as host for the bacterium [7-10]. *V. cholerae* can be found attached to the carapace and in the gut of copepods in large numbers, the copepod essentially serving as a vector for this human pathogen [1, 11, 12]. The importance of copepods in cholera transmission was demonstrated in a study showing that the number of cholera cases in Bangladeshi villages was significantly reduced when a simple filtration method that effectively removed the plankton and particulate matter was used to treat drinking water [13, 14]. Another advance in understanding cholera epidemics was made with the discovery that the epidemics are significantly influenced by climatic factors [1, 5, 15, 16]. Several environmental drivers of the seasonal cycle of cholera in India and Bangladesh where cholera is endemic have been identified [1]. Studying the association between selected environmental variables and confirmed cholera cases in two separate human populations resident in Kolkata, India, and Matlab, Bangladesh, in the northern Bay of Bengal, a recent study [17] concluded that ocean and climate patterns are useful predictors of cholera epidemics, with the dynamics of endemic cholera being related to climate and/or changes in the aquatic ecosystem.

Microbiological and environmental controls of cholera epidemics are based on studies over the past twenty years that were conducted in the Bengal Delta, unfortunately referred to as the native homeland of the disease. The Bengal deltaic region has endemic (defined as regions where recurrence of the disease has continued for at least ten consecutive years) cholera. The challenge is how to translate knowledge of the ecology, microbiology, and pathology of *Vibrio cholerae* and the epidemiology of cholera into a predictive model applicable to other susceptible where cholera outbreaks are sporadic but not endemic. Three observations motivated us to seek an understanding of the relationship between the cholera and associated

Dhaka University Institutional Repository

geophysical processes in epidemic regions: (i) current knowledge of cholera and the effectiveness of interventions are based on surveillance and treatment data from endemic regions; (ii) the disease cannot be eradicated, since *Vibrio cholerae* is autochthonous to the aquatic environment; and (iii) outbreaks of cholera in endemic regions show strong hydroclimatological controls, an encouraging observation since such conditions can be monitored in epidemic regions.

Methods:

Bakerganj. Bakerganj, an administrative unit comprising a group of villages, is located in the southern district of Barisal, about 70 km north of the coast of the Bay of Bengal, and approximately 300 km southwest of Dhaka, the capital city of Bangladesh (Fig 1). An epidemiological and ecological surveillance [18-21] has reported significant correlation between selected environmental parameters and the ecology and epidemiology of *V. cholerae* in Bakerganj. In the present study, which is part of an on ongoing epidemiological and ecological surveillance, samples were collected biweekly from SDHC of Bakerganj, and monthly from 8 environmental sites (one river, five ponds, and two lakes). The SDHC is a government run community level rural health care facility, with hospital beds. The sampling sites comprised of the major river, Tulatali, which flows through Bakerganj, and man-made ponds and lakes that are located in the villages and are heavily used by the nearby villagers for drinking and other domestic purposes.

Mathbaria. Mathbaria is located adjacent to the Bay of Bengal, and approximately 400 km southwest of Dhaka. Mathbaria is also an administrative unit under the district of Pirojpur with a police station and an SDHC. This area has not been studied previously. The SDHC in Mathbaria is also a government run community level rural health care facility, with hospital beds. In this study, samples from Mathbaria were collected biweekly from patients attending the SDHC of Mathbaria, and also from 6 man-made ponds (one pond has link with the river) that are heavily used for drinking and other domestic purposes. A major river, Baleshwar, flows along the western boundary of Mathbaria, the other side of which is the tropical mangrove forest of 'Sundarbans'.

Dhaka University Institutional Repository

Environmental analysis. (a) Sample collection and processing. Water and plankton samples were collected monthly from seven ponds and one river in Bakerganj and biweekly from six ponds in Mathbaria. All samples were collected aseptically in sterile dark Nalgene bottles (Nalgene Nunc International, USA), placed in an insulated plastic box, and transported at ambient air temperature from the site of collection to the central laboratory at the ICDDR, B in Dhaka. All samples were processed the following morning, with approximately 20 hours elapsing between sample collection in the field and processing in the laboratory.

One hundred liter of water was filtered successively through 64 μm and 20 μm mesh nylon nets (Millipore Corp., Bedford, MA), which served as crude estimates of zooplankton and phytoplankton, respectively, and concentrated to 50 ml for each. Filtrates of the 20 μm mesh nylon net were designated as water samples, to be analyzed for planktonic (unattached, free swimming) bacteria. Both plankton and water samples were subjected to direct plating, as well as enrichment for *V. cholerae*. For plating, plankton samples were crushed, using a glass homogenizer (Elberbach Corp., Ann. Arbor. Mich.), to release attached bacteria. Water samples were concentrated by filtering through a 0.22 μm bacteriological membrane filter (Millipore Corp., Bedford, MA) and washed from the filter to Luria Bertani (LB) broth for enrichment in alkaline peptone water for *V. cholerae*.

(b) Physicochemical parameters. Air and water temperature, water conductivity, total dissolved solid (TDS) and salinity were measured using a portable meter (HACH model sens*ION*156 portable pH/Conductivity/Dissolved Oxygen meter). Dissolved oxygen (DO) and pH were also measured; by using the same sens*ION*156 portable meter connected with Platinum pH Electrode, and DO Probe, respectively. All the parameters were measured by inserting the probe directly into the pond water and in case of rain water from the pond is collected in a 5- L plastic beaker and measured therein. All three probes were calibrated prior to each sampling visit. The pH probe was calibrated by 3 pt. calibration using pH 4, 7 & 10, DO probe was calibrated using saturated air, and Conductivity probe was calibrated using 1000 μS/cm conductivity standard. A graduated rope with a heavy iron ring attached was used to measure water depth at a fixed point around the center of the pond.

Dhaka University Institutional Repository

c) Enrichment and plating. Samples were enriched in alkaline peptone water (APW) at 37°C for 6-8 hours before plating as described previously [22-24]. About 2-3 loops, using an inoculating loop, of enriched APW broth were streaked onto both thiosulfate–citrate-bile-salts-sucrose (TCBS) and taurocholate-tellurite-gelatin agar (TTGA), and incubated at 37°C for 18 to 24 h. Colonies with the characteristic appearance of *V. cholerae* were confirmed by biochemical and serological tests and, in the case of the latter by using polyvalent and monoclonal antibodies specific for *V. cholerae* O1 or O139 and, finally, by molecular methods [25].

(d) Direct detection of *Vibrio cholerae* in environmental samples. Total DNA was extracted from all three fractions of the sample, i.e., >64-µm, >20-µm to <64-µm, and filtrate from the 20-µm, following the method of Rivera et al. [26], using cetyltrimethylammonium bromide (CTAB) and phenol-chloroform-isopropyl alcohol. DNA was precipitated in 0.6 volume of isopropyl alcohol and centrifuged to form a pellet which was washed with 70% ethanol, resuspended in tris-ethylenediaminetetraacetic acid (TE) buffer (pH 8.0), and quantified by ultraviolet spectrophotometry. Molecular detection of toxigenic *V. cholerae* serogroups O1 and O139 and cholera toxin *(ctxA)* genes was carried out by multiplex polymerase chain reaction, using primers and procedures described by Hoshino et al. [27]. Unfiltered water samples inoculated directly into APW were similarly treated for DNA extraction.

Collection and analysis of clinical specimens. Rectal swabs were obtained every two weeks from patients admitted to the local Sub-District Health Complex of Bakerganj and Mathbaria. The rectal swabs from suspected cholera patients with watery stool collected for consecutive three days were immediately placed into Cary Blair medium (in grams per liter: sodium thioglycolate, 1.5 g; disodium phosphate, 1.1g; calcium chloride, 0.1 g; sodium chloride, 5 g; agar, 5 g; pH 8.4) and transported at ambient temperature to the central laboratory of the International Center for Diarrhoeal Disease Research, Bangladesh (icddr,b), in Dhaka. In the laboratory, the rectal swabs were placed in alkaline peptone water (APW; in grams per liter: Bacto Peptone, 10 g; sodium chloride, 10 g; pH 8.8) and incubated at 37°C for 4 h. The

Dhaka University Institutional Repository

rectal swabs as well as the 4-h-enriched broth were cultured by inoculation onto taurocholate tellurite gelatin agar (in grams per liter: trypticase, 10 g; sodium chloride, 10 g; sodium taurocholate, 5 g; sodium carbonate, 1 g; gelatin, 30 g; agar, 15 g; pH 8.5; after autoclaving, the agar is cooled to 50°C, 5 ml of 0.1% potassium tellurite solution is added and mixed well, and the agar is poured into plates). Suspected colonies resembling *V. cholerae* were tested by slide agglutination with polyvalent anti-O1 and anti-O139 sera.

Statistical analysis: Clinical cholera cases and environmental data collected through February 2005 to September 2007 were used to derive a relationship between the number of cholera cases in each studied area (Mathbaria and Bakergonj) and environmental variables (8 Variables). These variables were air and water temperature (°C), water depth (m), water conductivity (uS/cm), Total dissolved solids, TDS (mg/L), salinity (ppt), dissolved oxygen, DO (%), and pH. Data statistical analysis were done using STATA 12.0 statistical software package at 0.05 significant levels. Where, the statistical methods involved mainly: ***(i) Pearson correlation*** to diagnose the degree of linear relationship between all the environmental variables and cholerae. ***(ii) Descriptive statistics*** were used to estimate the means and standard deviations value for all the variables included in the model. In addition, a Monte Carlo approach was used to detect significant association between occurrence of cholera and the direct detection of molecular markers of *V. cholerae* serogroups O1 or O139 and *ctxA* in one of the three fractions of the sample, i.e., >64-µm, >20-µm to <64-µm, and filtrate from the 20-µm. The number of simultaneous occurrences of positive observations in the two variables was divided by the number of observations. The bootstrap procedure, employing 100,000 series generated randomly and preserving the same first-order autocorrelation structure as observed in the data, was established to compute a stable distribution of the statistic that can be expected to occur randomly. The p-value of the test corresponds to one minus the rank of the observed statistic over the 100,000 statistics obtained randomly, divided by 100,001. If the p-value was equal to or less than a significance level of 5% corrected by Bonferonni method for multiple comparison procedure, we considered the association to be significantly different than by chance. A time lag was considered between occurrence of positive observations in the two variables, equal to four weeks for Bakerganj, and two and four weeks for Mathbaria. We investigated the association

Dhaka University Institutional Repository

between occurrence of cholera and an earlier presence of molecular markers of *V. cholerae* O1 or O139 and *ctxA* in the water. In addition, we also analyzed the association between occurrence of cholera and a later occurrence of molecular markers of *V. cholerae* O1 or O139 and *ctxA*.

Results and Discussion:

Approximately 2,409 water (filtrate from the 20-μm), and plankton (>64-μm, and >20-μm to <64-μm) samples were collected during the course of this study (32 months: Feb 2005 to Sep 2007). Toxigenic *Vibrio cholerae* O1 and O139 were cultured from 3.24% and 1.7% of environmental samples respectively, whereas non-toxigenic *V. cholerae* non-O1/O139 was isolated in much greater frequency (45%). Interestingly, area wise isolation rate for *V. cholerae* O139 appeared to be quite different, with Mathbaria demonstrating approx. 2.4% (n=38) isolation rate as opposed to 0.37% (n=3) observed in Bakergonj. However, isolation rate for *V. cholerae* O1 was comparable in both areas. It is important to note that cholera caused by *V. cholerae* O139 has been declining in recent years, despite detection of a molecular marker for this serogroup and continued isolation of *V. cholerae* O139 by culture methods, showing its persistence in the Mathbaria environment and reaffirmed that coastal aquatic environment of Bay of Bengal is yet a reservoir for *V. cholerae* O139.

Area/ Variable	Bakerganj					Mathbaria				
	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
pH	32	7.88	0.36	7.03	9.02	32	7.77	0.30	7.11	8.48
Dissolved O_2 (mg/L)	32	7.35	2.15	4.90	15.91	32	7.02	0.83	5.28	8.91
Total Dissolved Solids (mg/L)	32	104.19	37.14	66.30	203.53	32	212.00	105.24	85.45	555.58
Conductivity (μS)	32	222.76	82.36	142.15	418.75	32	438.45	233.97	178.34	1247.53
Water Depth (m)	32	2.07	0.35	1.53	2.68	32	2.07	0.39	1.50	2.69
Air Temperature (°C)	32	31.42	3.30	23.99	36.63	32	33.10	4.29	26.98	50.78
Water Temperature (°C)	32	29.15	3.64	19.90	35.53	32	29.26	3.64	20.74	33.98
Salinity (ppt)	32	0.03	0.04	0.00	0.16	32	0.17	0.11	0.05	0.52
Environmental detection of										
V. cholerae non-O1/O139	32	8.47	3.88	1.00	19.00	32	20.88	7.57	7.00	40.00
V. cholerae O1 and O139	32	1.00	2.13	0.00	8.00	32	2.31	4.80	0.00	17.00

Table 4.1. Descriptive statistics of the environmental variables (monthly average) and *V. cholerae* isolation (average number of samples positive/month) used for statistical correlation.

Dhaka University Institutional Repository

Descriptive statistics for the environmental variables and *V. cholerae* isolations are shown in Table 4.1. Subsequent analysis indicated that most of the environmental variables i.e., pH, water and air temperature, conductivity, TDS, and salinity, were positively correlated with the isolation of toxigenic *V. cholerae* in the environment (Fig. 4.1). Interestingly, in both Mathbaria and Bakerganj, water depth demonstrated a negative correlation with toxigenic *V. cholerae* isolation in the environment, which might be explained by the fact that elevated air temperature, for which we observed a positive correlation with toxigenic *V. cholerae* isolation, might caused increased evaporation of the surface water and leading to greater human consumption of water thereby reducing the water depth, which in turn may increase the water salinity favoring the growth of toxigenic *V. cholerae*. In general, toxigenic *V. cholerae* occurrence in both Mathbaria and Bakergonj show relatively week correlation with the environmental factors (not more than 0.4). However, isolation of *V. cholerae* non-O1/O139 in Mathbaria shows strong correlation with TDS, conductivity, water temp and salinity (Fig. 4.1).

Dhaka University Institutional Repository

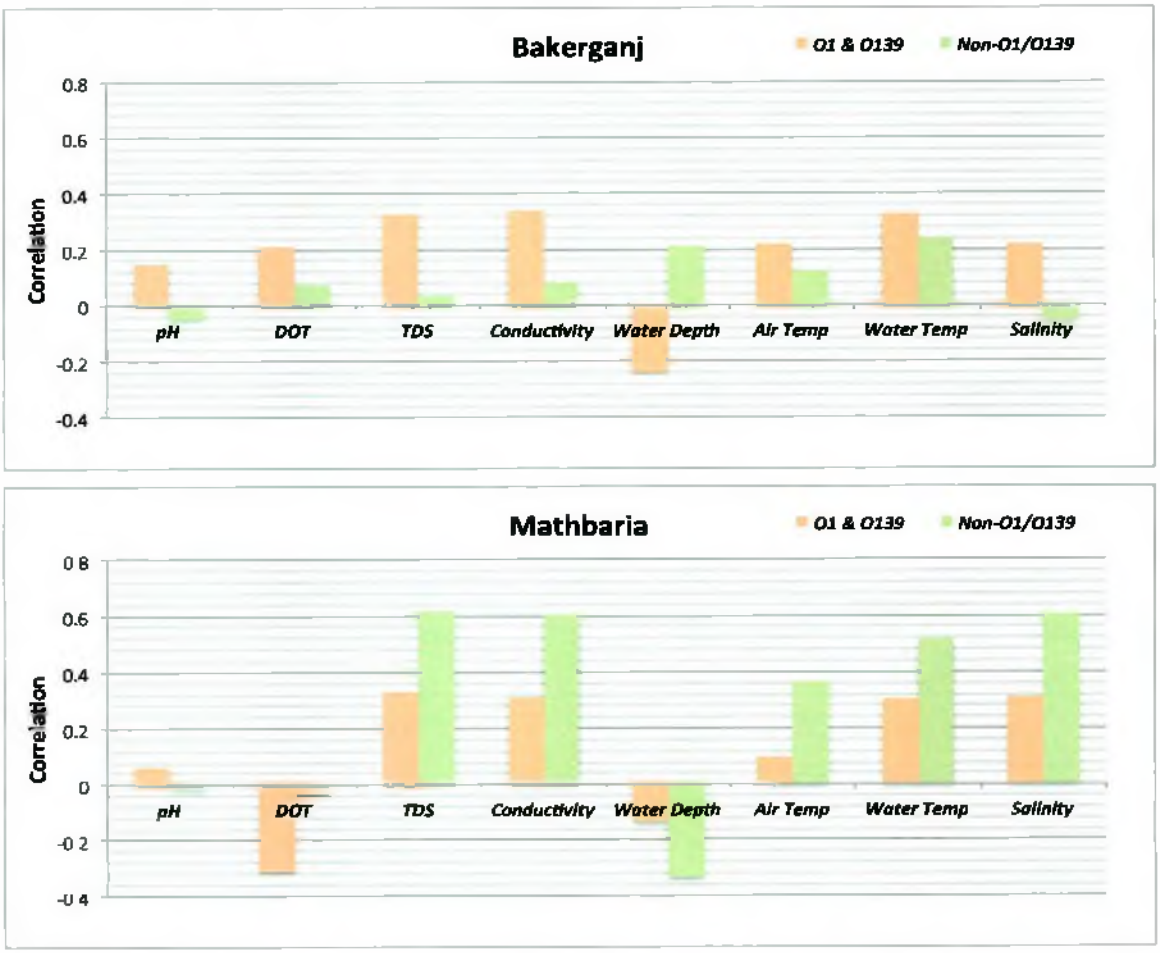


Figure 4.1 Statistical correlations of the environmental variables with isolation of toxigenic and non-toxigenic *V. cholerae*.

A close association was observed between TDS, conductivity, air and water temperature with cholera and diarrheal cases among villagers at Bakerganj and Mathbaria (Fig. 4.2 and 4.3, respectively). It was interesting that, during the study, there was a negative association between the number of cholera cases and water depth at Mathbaria, which was also observed for environmental isolation of toxigenic *V. cholerae* (Fig. 4.4); with decreased water depth, the number of cholera cases increased. Increase in salinity is most probably responsible, according to Huq et al., 2005 [20] and Louis et al. [30]. However, this observation doesn't appear to be true for Bakergonj clinical cholera case, which perhaps may be explained by fact that most of study sites at Bakerganj were directly connected with adjacent river to maintain adequate supply of water for the villagers as we also observed similar trend with Bakergonj environmental isolation of toxigenic *V. cholerae*. Figure 4.5 and 4.6 illustrates a moderate, correlation of pH and salinity with cholera cases, indicatlng

that increases in water pH and salinity are associated with simultaneous increases in cholera cases in these regions.

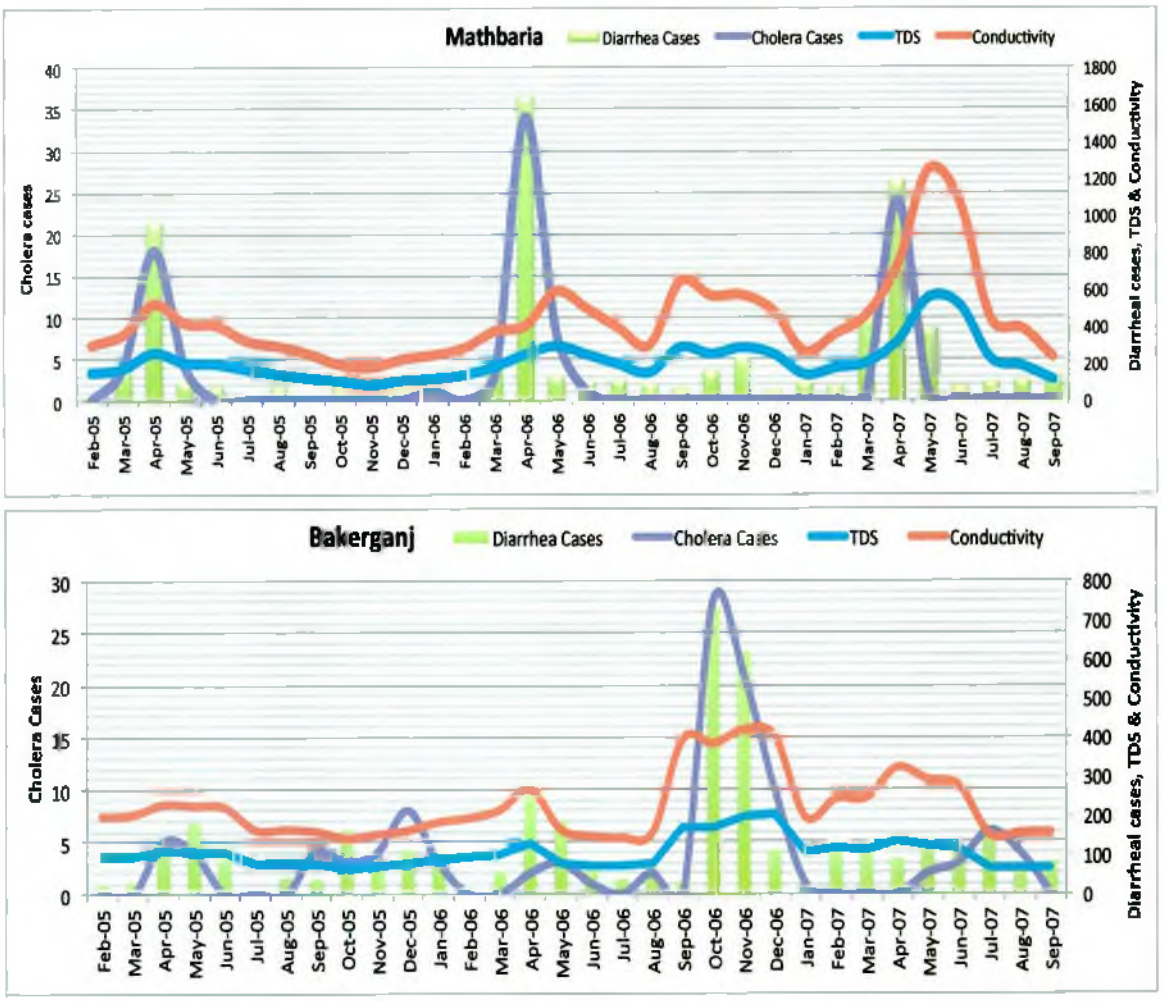


Figure 4.2. Temporal dynamics of Conductivity and Total Dissolved Solids (TDS) and their correlation with observed Cholera and Diarrheal cases in Mathbaria and Bakerganj, Bangladesh.

Dhaka University Institutional Repository

Dhaka University Institutional Repository

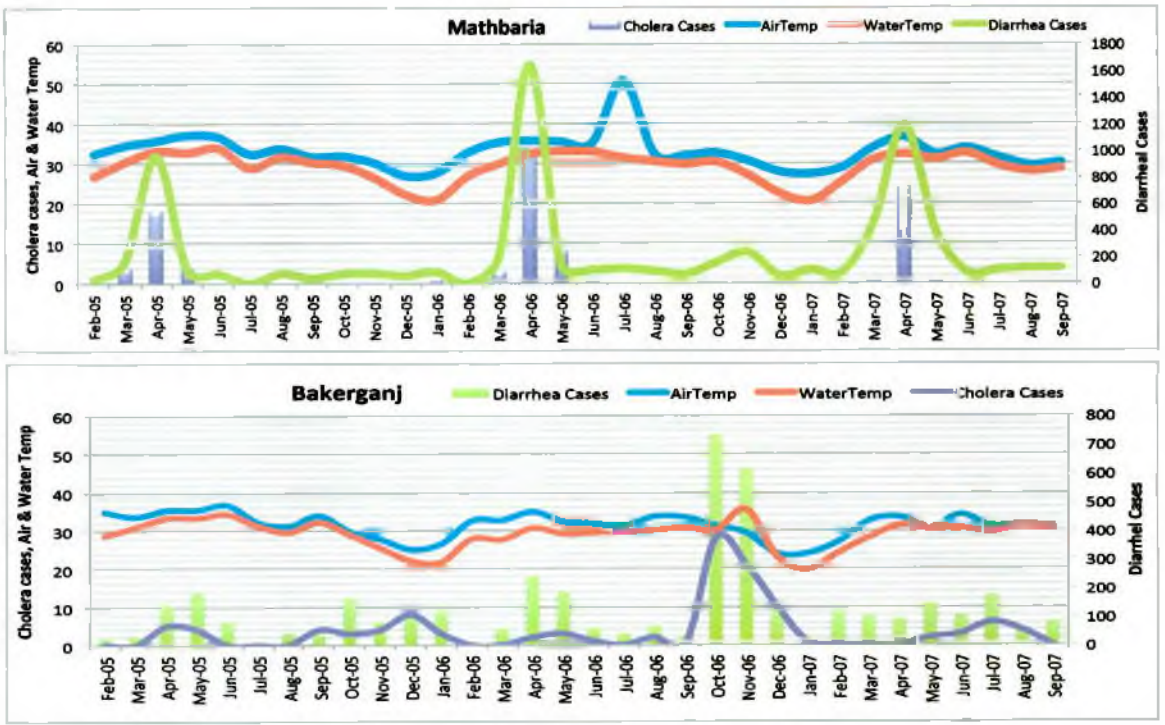


Figure 4.3. Temporal dynamics of air and water temperature and their correlation with observed Cholera and Diarrheal cases in Mathbaria and Bakerganj, Bangladesh.

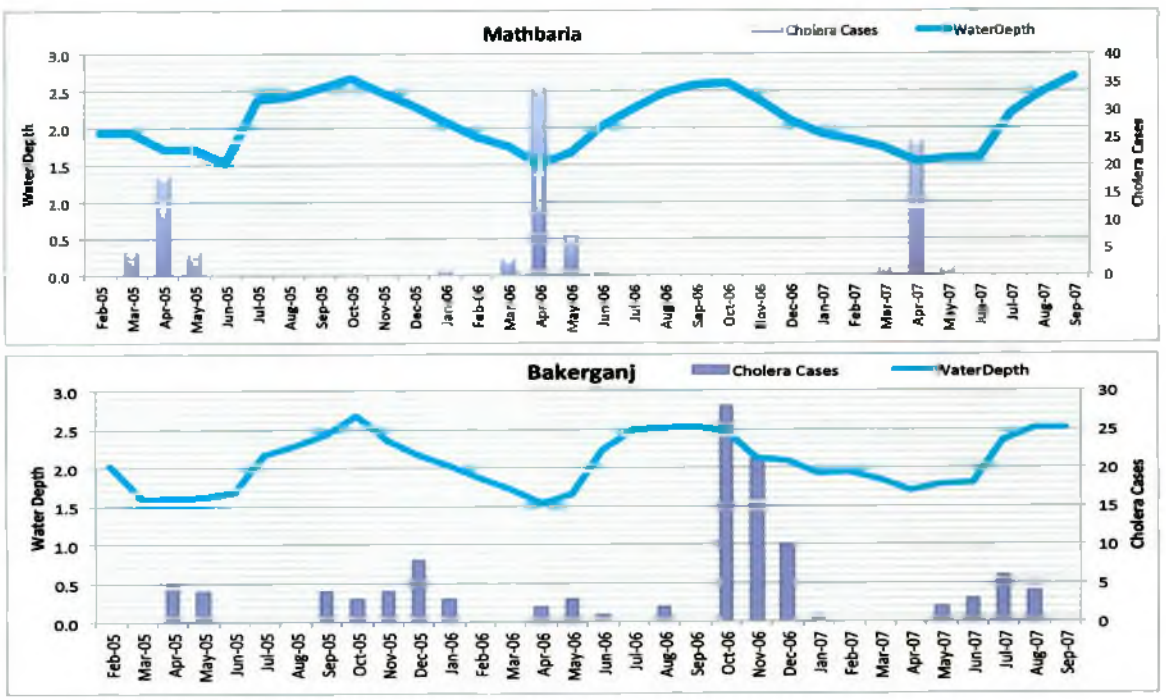


Figure 4.4. Temporal dynamics of water depth correlated with the observed cholera cases in Mathbaria and Bakerganj, Bangladesh.

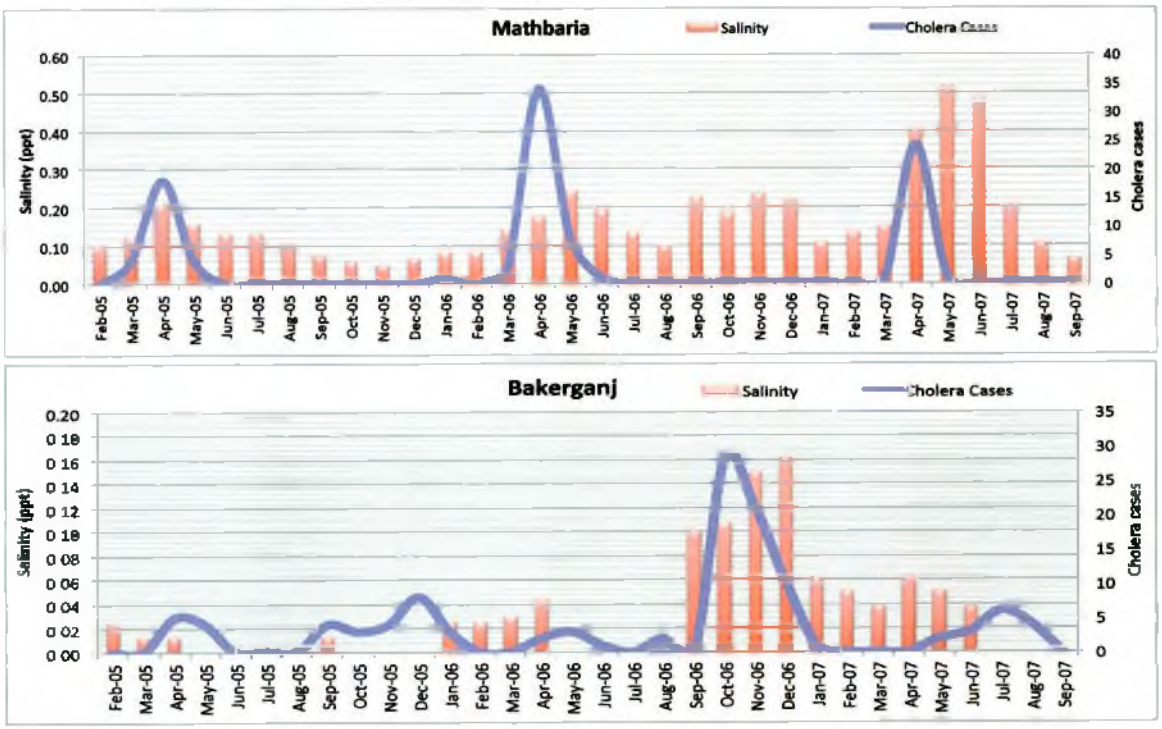


Figure 4.5. Correlation of water salinity with observed cholera cases in Mathbaria and Bakerganj, Bangladesh.

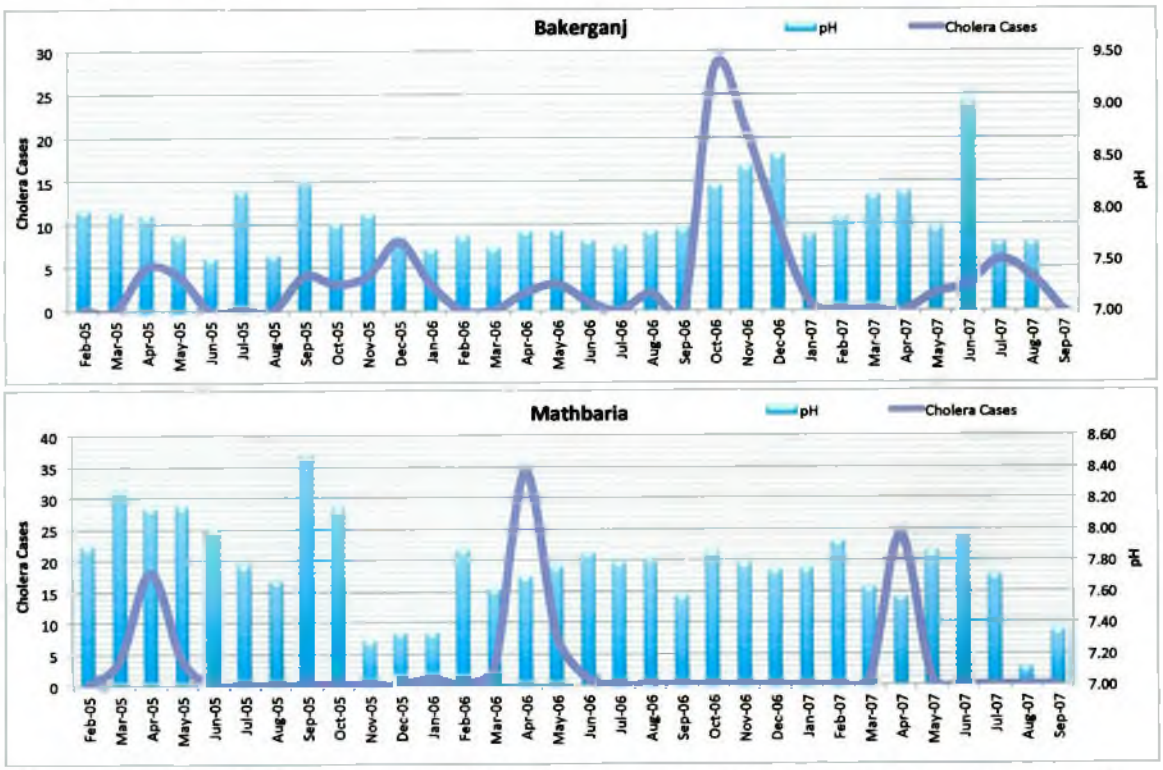


Figure 4.6. Correlation of water pH with observed cholera cases in Mathbaria and Bakerganj, Bangladesh.

Dhaka University Institutional Repository

Dhaka University Institutional Repository

Next we assessed the association between occurrence of cholera and *Vibrio cholerae* O1 and O139 (wbeT), and *ctxA*. In Bakerganj, significant association between occurrence of cholera caused by *V. cholerae* O1 and the presence of *V. cholerae* O1 (wbeT), detection of *ctxA,* or a combination of *V. cholerae* O1 (wbeT) and *ctxA* in the environment, without a time lag, unexpectedly was not observed. The association was significant, however, when detection of *V. cholerae* O1 (wbeT), *ctxA* or a combination of *V. cholerae* O1 (wbeT) and *ctxA* was present in the environment with a four week delay before appearance of cholera (p-values=0.044, 0.039, and 0.035, respectively). These results suggest that detection of *V. cholerae* O1 (wbeT) and *ctxA* in water samples provides at least a four week early warning of a cholera outbreak in the population of Bakerganj. A significant association was not observed for presence of *V. cholerae* O1 (wbeT), detection of *ctxA,* or a combination of *V. cholerae* O1 (wbeT) and *ctxA* detection in the environment four weeks after occurrence of cholera caused by *V. cholerae* O1, indicating that once a cholera outbreak occurs, other factors, such as person to person transmission become dominant. Because cholera caused by *V. cholerae* O139 had not occurred, association of *V. cholerae* O139 in environmental samples could not be analyzed.

In Mathbaria, highly significant association between occurrence of cholera caused by *V. cholerae* O1 and presence of either *V. cholerae* O1 (wbeT) (p-value = 0.009), *ctxA* (p-value = 0.001), or a combination of *V. cholerae* O1 (wbeT) and *ctxA* (p-value = 0.001) in the environmental samples, without any time lag, was observed. With respect to time lag, it was only the association between occurrence of cholera caused by *V. cholerae* O1 and presence of *V. cholerae* O1 (wbeT) (p-value = 0.034) detected one round (two weeks) earlier that proved to be significant. None of the associations were significant for a lag period of two rounds (four weeks earlier), most likely because person to person transmission and other factors were dominant. As was concluded from analysis of data for Bakerganj, early detection of *V. cholerae* O1 (wbeT) and *ctxA* in pond water can serve as warning of cholera in Mathbaria, but with an even shorter time for advance warning (one round, i.e., 2 weeks). In contrast to Bakerganj, highly significant association was found between occurrence of cholera caused by *V. cholerae* O1 and detection of *V. cholerae* O1 (wbeT), with a time lag of two to four weeks (p-value = 0.002 and 0.046, respectively), *ctxA* (p-value < 0.001 and 0.015, respectively), or a combination of

Dhaka University Institutional Repository

V. cholerae O1 (wbeT) and *ctxA* (p-value < 0.001 and 0.007, respectively). This different pattern may be related to epidemiological and/or ecological factors in Bakerganj and Mathbaria, as noted previously (see above) for zooplankton diversity. Bakerganj is more densely populated than Mathbaria, with a cholera dynamic of two peaks per year, April to May and September to December. Mathbaria, less populated, is located in a very low lying coastal area, and has only one cholera outbreak per year, March to May, with sporadic cases in November. The contrast between the two locations highlights local specificity, even though both are geographically close, indicating the need for fine scale data analysis to understand underlying processes. In any case, the results for Bakerganj clearly support the hypothesis of the environment as the source of cholera, as do the results for Mathbaria. The argument has been made that cholera outbreaks are local in their source [31] and DNA based methods are being applied to confirm this hypothesis, the results of which will be published elsewhere.

Significant association between occurrence of cholera cases caused by *V. cholerae* O139 and presence of *V. cholerae* O139 in the environment, detection of *ctxA*, or a combination of *V. cholerae* O139 and *ctxA* in the environment without a time lag, or a lag of two to four weeks was not observed in this study, a result most likely related to the limited number of recorded cholera cases caused by *V. cholerae* O139 in Mathbaria. Cholera caused by *V. cholerae* O139 has been declining in recent years, despite detection of a molecular marker for this serogroup and continued isolation of *V. cholerae* O139 by culture methods, showing its persistence in the Mathbaria environment (unpublished data).

Finally, we observed a good correlation when the numbers of species identified by 16S metagenomics were plotted against environmental isolation of toxigenic *V. cholerae* and cholera cases reported in Mathbaria (Fig. 4.7) during the study period. This prompted us to investigate whether the environmental variables investigated in the study play a role in the temporal variation of species diversity observed by metagenomics. Interestingly, we observed a positive correlation of certain environmental variables i.e., salinity, conductivity, and water temperature, with the temporal variation of numbers of species detected by 16S metagenomics, whereas water depth demonstrated an inverse correlation (Fig. 4.8). Such observation

Dhaka University Institutional Repository

indicated that increased salinity, conductivity water temp and water depth might stimulate the growth of entire microbial community being *V. cholerae* as a member of the community.

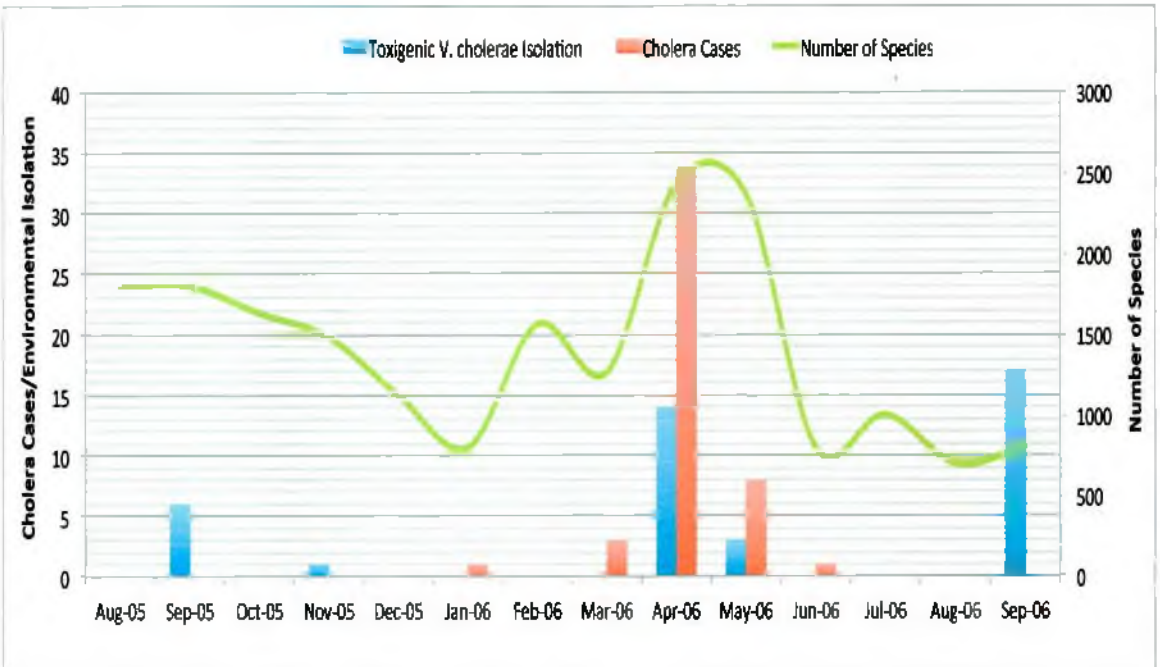
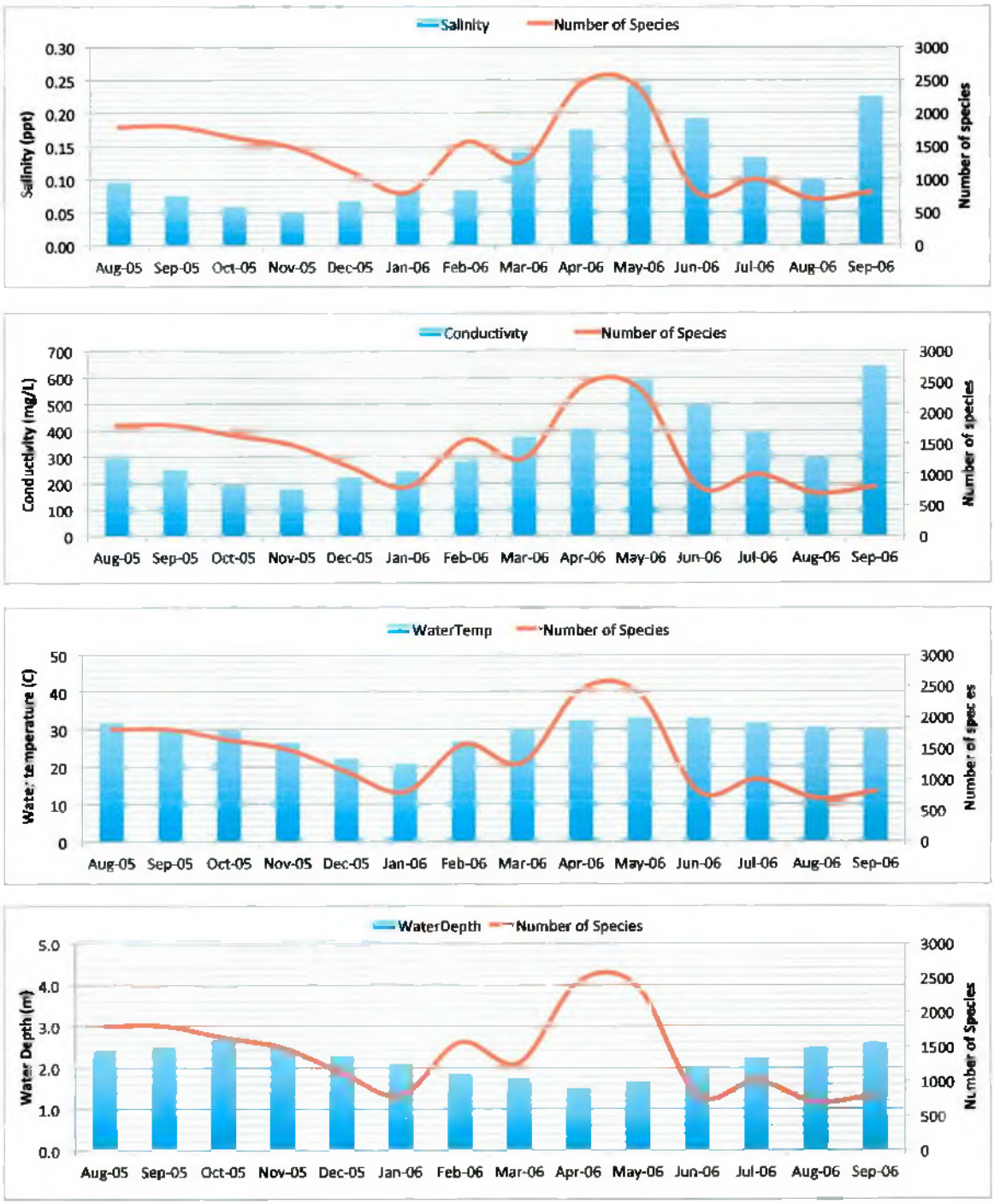


Figure 4.7. Correlation of cholera cases and environmental isolation of toxigenic *V. cholerae* with the metagenomic identification of microbial species over time in Mathbaria, Bangladesh.

Dhaka University Institutional Repository



466259

Figure 4.8. Temporal distribution of numbers of species identified by 16S metagenomics correlated with certain environmental variables.

ঢাকা বিশ্ববিদ্যালয় গ্রন্থাগার

Dhaka University Institutional Repository

Conclusion:

In this study, we examined the role of environmental factors in the occurrence and transmission of cholera to gain additional precision in the predictability of an upcoming cholera outbreak. Previous studies documented temperatures and pH as an important predictor of *Vibrio cholerae* occurrence in the Chesapeake Bay (20, 21, 25). Earlier studies in Bangladesh and Peru identified temperature as a key factor associated with increased counts of *V. cholerae* and cases of cholera (9, 13, 24) whereas a recent study (20) in Bangladesh suggested multiple variables i.e., temperature, salinity, and zooplankton. Interestingly, a nearly identical result was obtained for Bakerganj and Mathbaria, with water temp and salinity being positively correlated with the environmental isolation of *V. cholerae* and cases of cholera observed in these regions. In contrast, the data presented here highlight the importance of the multiple variables i.e., conductivity, air and water temp, water depth and salinity, as key to the development of a predictive model for cholera. A remarkable finding of this study was the close association of salinity, conductivity, water temp and water depth with the number of microbial species detected by 16S metagenomics, which indicated that these predictors might have predictive value for other water borne diseases, and therefore warrant further attention.

Previous studies (3, 17, 20) have reported the failure of conventional microbiology to provide useful information for prediction of cholera, mostly because highly selective nature of the conventional culture media might have provided an additional shock to the already stressed organism from the environment. Therefore, in addition to conventional culture isolation, we also employed a PCR based direct detection method, which provided a very important finding of the study, a significant association between occurrence of cholera cases caused by *V. cholerae* O1 and presence of *V. cholerae* O1, *ctxA*, or a combination of *V. cholerae* O1 and *ctxA* in the environment without any time lag or with a four week lag before appearance of cholera in Mathbaria and Bakerganj, respectively. Finally, from the result of this study, we were able to validate the effect of few previously proposed environmental variables on the occurrence of cholera, as well as propose additional variables, combinations of which might be useful in developing environmental models for the prediction of cholera epidemics.

Dhaka University Institutional Repository

References:

1. Colwell, R.R., *Global climate and infectious disease: the cholera paradigm.* Science, 1996. **274**(5295): p. 2025-31.
2. Hasan, N.A., et al., *Genomic diversity of 2010 Haitian cholera outbreak strains.* Proc Natl Acad Sci U S A, 2012. **109**(29): p. E2010-7.
3. Chun, J., et al., *Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic Vibrio cholerae.* Proc Natl Acad Sci U S A, 2009. **106**(36): p. 15442-7.
4. Huq, A. and R.R. Colwell, *Environmental factors associated with emergence of disease with special reference to cholera.* Eastern Mediterraneal Health Journal, 1996. **2**: p. 37-45.
5. Lobitz, B., et al., *Climate and infectious disease: use of remote sensing for detection of Vibrio cholerae by indirect measurement.* Proc Natl Acad Sci U S A, 2000. **97**(4): p. 1438-43.
6. Jutla, A.S., et al., *Warming oceans, phytoplankton, and river discharge: implications for cholera outbreaks.* Am J Trop Med Hyg, 2011. **85**(2): p. 303-8.
7. Colwell, R.R., J. Kaper, and S.W. Joseph, *Vibrio cholerae, Vibrio parahaemolyticus, and other vibrios: occurrence and distribution in Chesapeake Bay.* Science, 1977. **198**(4315): p. 394-6.
8. Huq, A., et al., *Ecological relationships between Vibrio cholerae and planktonic crustacean copepods.* Appl Environ Microbiol, 1983. **45**(1): p. 275-83.
9. Islam, M.S., B.S. Drasar, and R.B. Sack, *The aquatic flora and fauna as reservoirs of Vibrio cholerae: a review.* J Diarrhoeal Dis Res, 1994. **12**(2): p. 87-96.
10. Kaper, J.B., J.G. Morris, Jr., and M.M. Levine, *Cholera.* Clin Microbiol Rev, 1995. **8**(1): p. 48-86.
11. Nalin, D.R., et al., *Adsorption and growth of Vibrio cholerae on chitin.* Infect Immun, 1979. **25**(2): p. 768-70.
12. Rawlings, T.K., G.M. Ruiz, and R.R. Colwell, *Association of Vibrio cholerae O1 El Tor and O139 Bengal with the Copepods Acartia tonsa and Eurytemora affinis.* Appl Environ Microbiol, 2007. **73**(24): p. 7926-33.
13. Cash, R.A., et al., *Response of man to infection with Vibrio cholerae. I. Clinical, serologic, and bacteriologic responses to a known inoculum.* J Infect Dis, 1974. **129**(1): p. 45-52.
14. Huq, A., et al., *A simple filtration method to remove plankton-associated Vibrio cholerae in raw water supplies in developing countries.* Appl Environ Microbiol, 1996. **62**(7): p. 2508-12.
15. Colwell, R.R., et al., *Reduction of cholera in Bangladeshi villages by simple filtration.* Proc Natl Acad Sci U S A, 2003. **100**(3): p. 1051-5.
16. Lipp, E.K., A. Huq, and R.R. Colwell, *Effects of global climate on infectious disease: the cholera model.* Clin Microbiol Rev, 2002. **15**(4): p. 757-70.
17. Constantin de Magny, G., et al., *Environmental signatures associated with cholera epidemics.* Proc Natl Acad Sci U S A, 2008. **105**(46): p. 17676-81.
18. Alam, M., et al., *Seasonal cholera caused by Vibrio cholerae serogroups O1 and O139 in the coastal aquatic environment of Bangladesh.* Appl Environ Microbiol, 2006. **72**(6): p. 4096-104.
19. Alam, M., et al., *Toxigenic Vibrio cholerae in the aquatic environment of Mathbaria, Bangladesh.* Appl Environ Microbiol, 2006. **72**(4): p. 2849-55.
20. Huq, A., et al., *Critical factors influencing the occurrence of Vibrio cholerae in the environment of Bangladesh.* Appl Environ Microbiol, 2005. **71**(8): p. 4645-54.
21. Sack, D.A., et al., *Cholera.* Lancet, 2004. **363**(9404): p. 223-33.
22. Huq, A., et al., *Coexistence of Vibrio cholerae O1 and O139 Bengal in plankton in Bangladesh.* Lancet, 1995. **345**(8959): p. 1249.
23. Islam, M.S., et al., *Detection of non-culturable Vibrio cholerae O1 associated with a cyanobacterium from an aquatic environment in Bangladesh.* Trans R Soc Trop Med Hyg, 1994. **88**(3): p. 298-9.
24. Islam, M.S., et al., *Isolation of Vibrio cholerae O139 synonym Bengal from the aquatic environment in Bangladesh: implications for disease transmission.* Appl Environ Microbiol, 1994. **60**(5): p. 1684-6.

Dhaka University Institutional Repository

25. Nandi, B., et al., *Rapid method for species-specific identification of Vibrio cholerae using primers targeted to the gene of outer membrane protein OmpW.* J Clin Microbiol, 2000. **38**(11): p. 4145-51.
26. Rivera, I.N., et al., *Method of DNA extraction and application of multiplex polymerase chain reaction to detect toxigenic Vibrio cholerae O1 and O139 from aquatic ecosystems.* Environ Microbiol, 2003. **5**(7): p. 599-606.
27. Hoshino, K., et al., *Development and evaluation of a multiplex PCR assay for rapid detection of toxigenic Vibrio cholerae O1 and O139.* FEMS Immunol Med Microbiol, 1998. **20**(3): p. 201-7.
28. Hoshino, K., et al., *Development and evaluation of a multiplex PCR assay for rapid detection of toxigenic Vibrio cholerae O1 and O139.* Fems Immunology and Medical Microbiology, 1998. **20**(3): p. 201-207.
29. Rivera, I.N.G., et al., *Method of DNA extraction and application of multiplex polymerase chain reaction to detect toxigenic Vibrio cholerae O1 and O139 from aquatic ecosystems.* Environmental Microbiology, 2003. **5**(7): p. 599-606.
30. Louis, V.R., et al., *Predictability of Vibrio cholerae in Chesapeake Bay.* Appl Environ Microbiol, 2003. **69**(5): p. 2773-85.
31. Stine, O.C., et al., *Seasonal cholera from multiple small outbreaks, rural Bangladesh.* Emerg Infect Dis, 2008. **14**(5): p. 831-3.

Dhaka University Institutional Repository

CHAPTER 5:

VIBRIO CHOLERAE SEASONALITY AS A FUNCTION OF COMPLEX CHANGES WITHIN MICROBIAL COMMUNITIES OF THE AQUATIC ENVIRONMENT

Dhaka University Institutional Repository

Abstract:

The recent advent of metagenomics has made it possible to investigate the dynamics and changes in microbial communities in order to understand their spatial and temporal relationship with seasonal emergence and reemergence of infectious disease. In this connection, we have carried out a year-long systemic biweekly surveillance of a microbial community in a cholera endemic region in the vicinity of Bay of Bengal, the home land of Asiatic cholera, using 16S metagenomic analyses. The aim of this study was to understand the dynamics of microbial community over time, and to correlate any changes within this community with clinical surveillance data gathered during the same study period. The primary focus was not only to understand the microbial community influence on the emergence of a cholera epidemic but also to identify a community structure that might possibly serve as harbinger for forthcoming cholera epidemics, or, at the very least, to understand what happens in the community before, during, or after an outbreak. Here we report successful implementation of metagenomics to characterize the environmental community and demonstrate that aquatic environment of Mathbaria harbor a genetically diverse and complex microbial community including primary members of both freshwater and coastal ecosystems. Data generated from this study may serve as a preliminary baseline for monitoring bacterial complexity in the environment. Enormous diversity was uncovered at every taxonomic level and a large number of bacterial genera pathogenic to human and animals, and indication for possible existence of novel bacterial taxa in this ecosystem were also detected, which may have a significant impact on livestock, agriculture and human health. The abundance and diversity of the microbial community including *Vibrio* populations, exhibited a complex seasonal distribution, and multiple analyses suggest that a specific consortium within the natural community, perhaps aroused as a result of positive selection, might be crucial in triggering *Vibrio cholerae* to overcome natural competition barriers and emerge in epidemic proportion. Moreover, we observed a good correlation of cholera cases when the ecosystem experience high community diversity (alpha and gamma) with least dissimilarity (beta) among the habitats.

Dhaka University Institutional Repository

Introduction:

Natural environments harbor a large diversity of microorganisms, most of which have not yet been identified/or characterized. Although numerous studies have documented a rich diversity of bacterial communities in different natural habitats, it is an acceptable norm that microbial diversity is enormous and not fully recognized, *i.e.*, microorganisms represent the largest reservoir of uncharacterized biodiversity [1]. The primary focus of the present study is on changes in the microbial community and environmental factors associated with emergence of disease, *i.e.*, what happens in a community when there is an outbreak of disease. Pathogenic microorganisms may depend partially or entirely on other physical, chemical, or biological factors for their existence. Many microbes are strictly vector-dependent, while others are not. For both types, environmental factors can affect directly, or indirectly, survival, persistence and ability to produce disease. Temperature is a key factor [2]. Sunlight can affect the persistence and spread of a pathogen if it is associated with phytoplankton and/or algae [2]. For example, when algae and phytoplankton increase in biomass, zooplankton blooms rapidly follow. Bacteria associated with zooplankton also increase. Also, humidity, resulting from evaporation due to elevation of temperature, may play an important role in the occurrence of many diseases. With appropriate humidity and moisture, most bacteria survive longer than they would in less humid or dry areas. There are many diseases common to tropical climates that are linked to water transmission. If transmission between hosts does not involve vectors, then water, or at least humid conditions, can be involved in transmission [3]. For example, in warm and humid regions, where water is available as a transmitting medium, *Vibrio cholerae* may proliferate rapidly to the level of an infective dose. In general, it is impossible to separate environmental factors from biological factors, as can be seen from interrelationships in nature that play a significant role in the emergence of infectious diseases.

Microbiologists have been constrained in their efforts to describe the compositions of natural microbial communities using traditional methods. Few microorganisms have sufficiently distinctive morphology to be recognized by

Dhaka University Institutional Repository

microscopy. Culture dependent methods are biased, as a microorganism can be cultivated only after its physiological niche is perceived and duplicated experimentally. It is therefore widely believed that fewer than 20% of the extant microroganisms have been discovered [4] and that culture methods are inadequate for studying microbial community composition. In view of the physiological and phylogenetic diversity among microorganisms [5] and that 80% or more of microbes remain undiscovered, one questions how well we know the Earth's microbiota and its biochemical potential. Metagenomic studies, a culture independent approach to investigate the composition of mixed microbial communities from environmental niches, allows for the identification of not only culturable but also nonculturable undescribed bacteria [6-8]. Direct access to genomic DNA of co-existing microbial species can give a better understanding of evolution, lifestyle and diversity of the microorganisms and expose more of the hidden world of microbes [9]. We have performed a culture-independent analysis of the composition of water samples from a cholera endemic region to study microbial community, using a common but distinctive cellular component, 16S ribosomal RNA. This has lead to the hypothesis that *Vibrio cholerae* seasonality might be a function of more complex changes within microbial communities in the aquatic environment in cholera endemic areas, including changes in microbial populations that are not culturable or may not even be identified with traditional microbiologic techniques. As the majority of bacteria are yet to be discovered or brought into culture, it is clear that the majority of their genetic diversity is unknown. The prevailing bacterial diversity in a habitat can be affected by physical and chemical environmental variables that changes over time. Monitoring microbial changes in environmental sources across time thus will be very useful not only to see if changes in the microbial ecology influences when *Vibrio cholerae* appears but also to get a feeling for the overall diversity/stability of the microbial community over time, in particular when cholera is present or absent in the environment.

Dhaka University Institutional Repository

Methods:

Sample Collection and DNA Extraction. Water and plankton samples were collected biweekly from six ponds in Mathbaria during August 2005 to August 2006. Five hundred milliliter of water was filtered successively through 64 μm and 20 μm mesh nylon nets to remove zooplankton and phytoplankton, respectively, and concentrated by filtering through a 0.22 μm bacteriological membrane filter (Millipore Corp., Bedford, MA). Each membrane was washed off using 5 mL of sterile phosphate buffered saline (PBS), and stored at −80°C prior to DNA extraction. DNA was extracted from each of the collected water samples using the QIAamp DNA Mini Kit (QIAGEN) according to the manufacturer's protocol. The purified DNA was stored at −80°C until PCR amplification.

Amplification of 16S rRNA Genes and Pyrosequencing. The extracted DNA was amplified using primers targeting the V1 to V3 hypervariable regions of the bacterial 16S rRNA gene [10]. The primers used for bacteria were V1-9 F: 5′-X-AC-GAGTTTGATCMTGGCTCAG-3′ and V3-541R: 5′-X-AC-WTTACCGCGGCTGCTGG -3′ (where X barcode is uniquely designed for each soil sample, followed by a common linker AC). Polymerase chain reactions were carried out under the following conditions: initial denaturation at 94°C for 5 min, followed by 10 cycles of denaturation at 94°C for 30 s, annealing at 60°C to 55°C with a touchdown program for 45 s, and elongation at 72°C for 90 s. This was followed by an additional 20 cycles of denaturation at 94°C for 30 s, annealing at 55°C for 45 s, and elongation at 72°C for 90 s. The amplified products were purified using the QIAquick PCR purification kit (Qiagen, CA, USA). Amplicon pyrosequencing was performed by using 454/Roche GS-FLX Titanium Instrument (Roche, NJ, USA).

Processing of Pyrosequencing Data and Taxonomic Analysis. DNA sequences were processed and analyzed following the procedures described previously [10]. Total sequencing reads were divided and assigned to each sample by recognition of unique barcodes, followed by trimming sequences by removing barcode, linker, and primer

Dhaka University Institutional Repository

sequences from both sides. Resultant sequences were subjected to a filtering process where only reads containing 0-1 ambiguous base calls (Ns) and 300 or more base pairs were selected for further analysis. Nonspecific PCR amplicons that showed no match with the 16S rRNA gene database upon BLASTn search (expectation value of $>10^{-5}$) were also removed from subsequent analyses. Putative chimeric sequences were detected and screened using a similarity-based approach, which splits each query sequence into two even-length fragments and then assigns each fragment to a taxon using BLAST search against EzTaxon-extended database (http://eztaxon-e.ezbiocloud.net/), followed by removal of the sequences when two fragments differ at the order level or percent identities at greater than 95% for both fragments despite assigned to different taxonomies. All sequences were classified using EzTaxon-extended database.

Results and Discussion:

The sequences collected for this study provided an overview of the fresh water microbiota in this region. Across all 144 water samples collected over a period of one year through a biweekly surveillance (24 visits) of six natural water reservoirs in Mathbaria, Bangladesh, a total of 50 bacterial and five eukaryotic phyla were identified representing 92% *Bacteria* and 8% *Eukarya* in the population. Figure 5.1 demonstrates the global population diversity observed in this study. The bacterial phylotypes detected are shown in Figure 5.2, most sequences (96%) were related to seven major phyla: Proteobacteria (33%), Cyanobacteria (29%), Actinobacteria (13%), Firmicutes (10%), Bacteroidetes (8%), Gemmatimonadetes (2%), and Chloroflexi (1%). One hundred and twenty five bacterial classes belonging to these 50 bacteria phyla were documented, among them only eight major classes represented 92% of the total bacterial sequences, namely, Chroobacteria (28%), Gammaproteobacteria (14%), Actinobacteria (13%), Betaproteobacteria (11%), Bacilli (7%), Alphaproteobacteria (6%), Flavobacteria (4%), Spingobacteria (4%) and Clostridia (3%). Altogether, a total of 1600 bacterial genera and 3311 species were identified through out the study. On an average, each habitat constitute about 1988 bacteria species and out of these, only 28% (n=984) of them represented well-characterized species and genera in bacteria, while the vast majority of remaining genera

Dhaka University Institutional Repository

and species represents putative genera and species that were assigned following the similarity based taxonomic assignment criteria used in EzTaxon-extended database (http://eztaxon-e.ezbiocloud.net/), and includes representatives of uncultured phylotypes [11]. Detection of such a large number of bacterial taxa indicates the existence of vast genetic diversity and complex community structure in the coastal ecosystem of Bay of Bengal. Table 5.1 depicts the most abundant bacterial genera and species identified in this study and demonstrate that, along with some commensal bacterial populations, this coastal ecosystem serves as a reservoir for many bacterial genera pathogenic to human and animals.

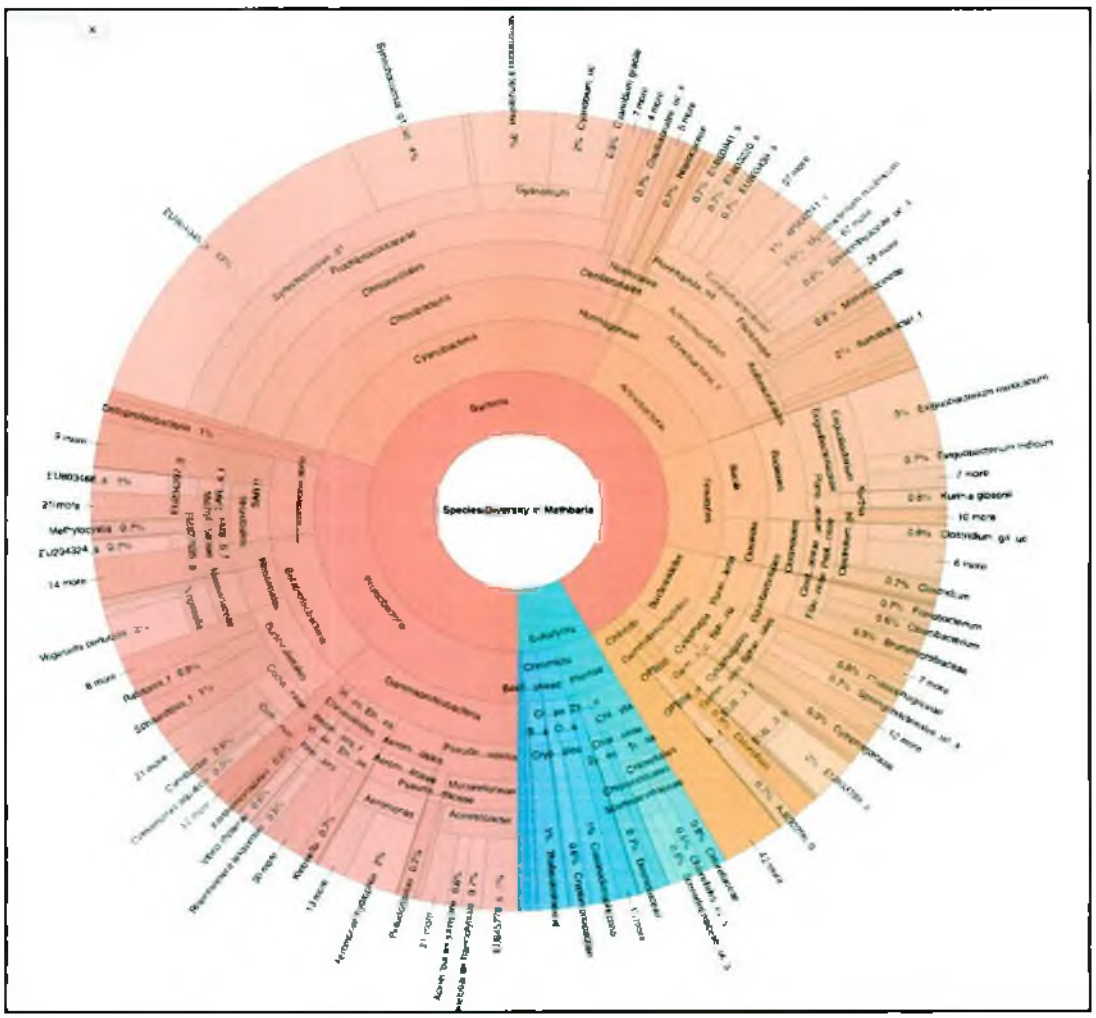
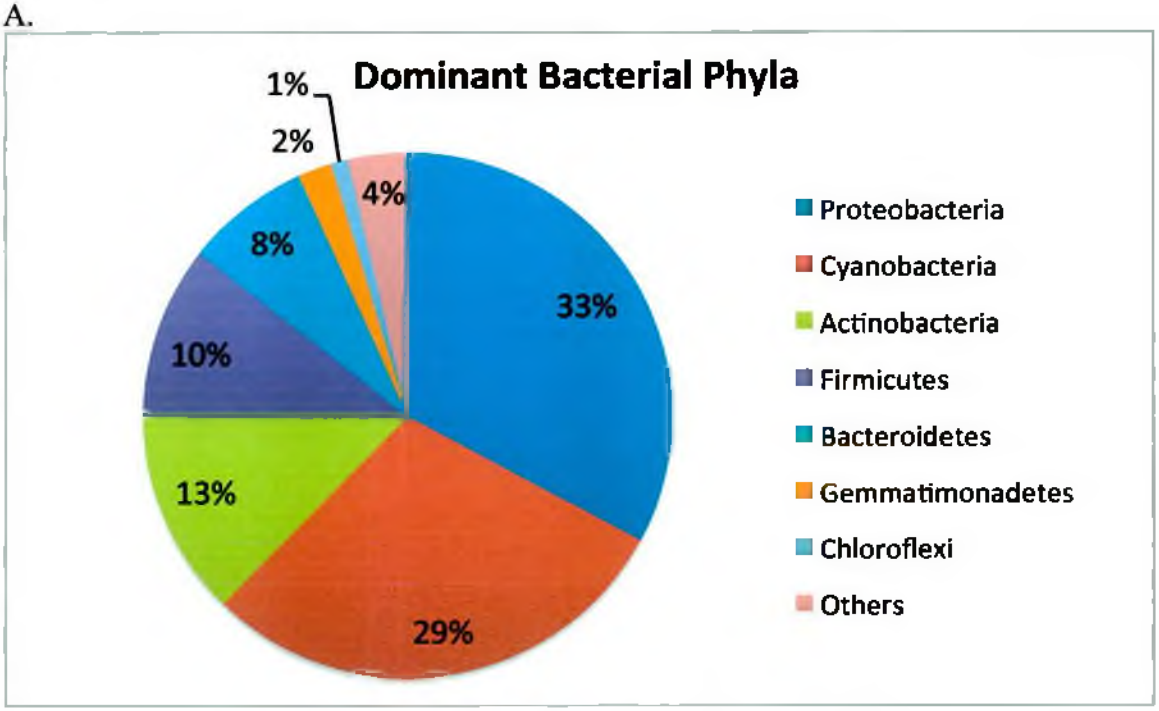


Figure 5.1: Global population diversity and overall community composition of coastal ecosystem of Mathbaria. The figure generated by Krona visualization tool demonstrates the hierarchical taxonomic distribution of each identity (kingdom to species level) and represented by from inner circle to outer circle, respectively.

Dhaka University Institutional Repository

A.



B.

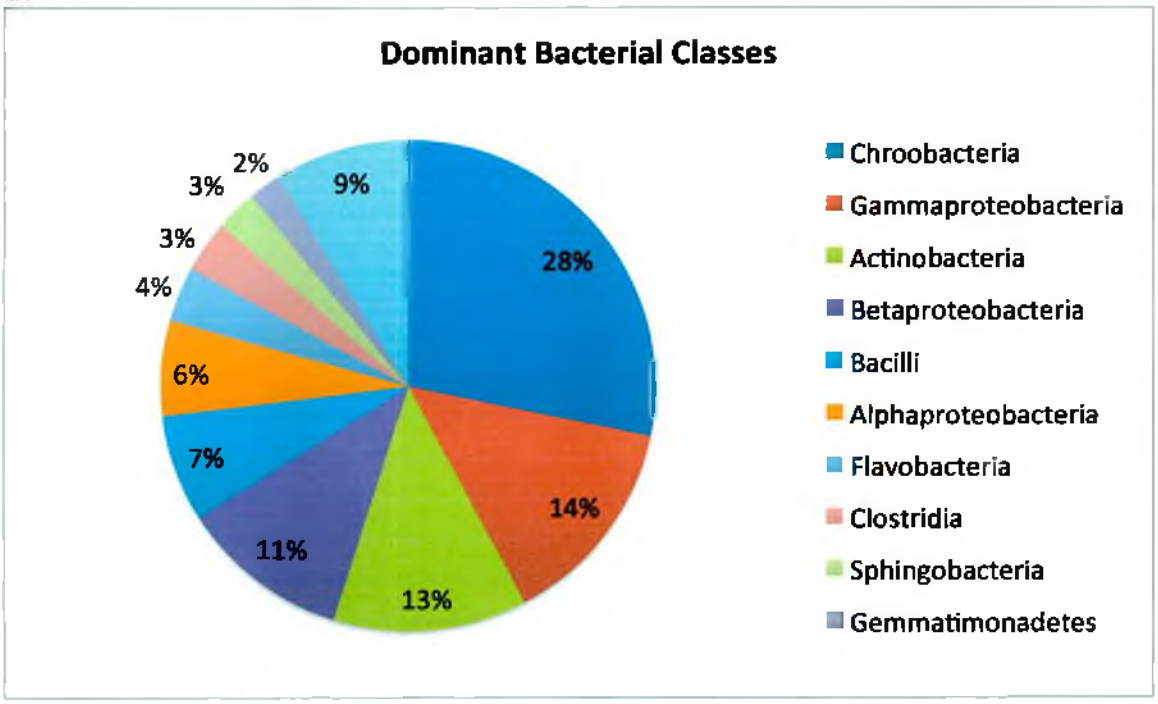


Figure 5.2: Relative abundance of major microbial phyla (A) and classes (B) in coastal ecosystem of Mathbaria, Bangladesh. A total of 50 different phyla and 125 classes were identified in this study, of them, 7 major phyla shown (A) and 8 major classes shown in (B) represent, 96% and 928% of the entire community, respectively.

Dhaka University Institutional Repository

Genus	Relative abundace	Species	Relative abundace
Cyanobium	12.84	*Exiguobacterium mexicanum*	10.55
Exiguobacterium	10.72	*Merismopedia tenuissima*	10.11
Acinetobacter	7.93	*Aeromonas hydrophila*	6.42
Aeromonas	6.74	*Vogesella perlucida*	5.71
Corynebacterineae	5.16	*Kurthia gibsonii*	2.40
Vogesella	4.77	*Acinetobacter haemolyticus*	2.33
Frankineae	3.51	*Exiguobacterium indicum*	2.32
Micrococcineae	1.98	*Acinetobacter johnsonii*	1.77
Kurthia	1.68	*Vibrio cholerae*	1.76
Flavobacterium	1.61	*Cyanobium gracile*	1.69
Clostridium	1.60	*Mycobacterium insubricum*	1.64
Klebsiella	1.53	*Comamonas aquatica*	1.62
Rheinheimera	1.52	*Rheinheimera texasensis*	1.62
Methylocystis	1.47	*Clostridium glycolicum*	1.54
Pseudomonas	1.47	*Klebsiella pneumoniae*	1.50
Cloacibacterium	1.38	*Cronobacter muytjensii*	1.44
Comamonas	1.33	*Aeromonas fluvialis*	1.34
Vibrio	1.30	*Clostridium bifermentans*	1.34
Curvibacter	1.27	*Acidovorax temperans*	1.31
Cronobacter	1.03	*Aeromonas jandaei*	1.25
Acidovorax	0.97	*Exiguobacterium profundum*	1.09
Polynucleobacter	0.96	*Flectobacillus roseus*	1.08
Ralstonia	0.96	*Clostridium lituseburense*	1.06
Beijerinckia	0.90	*Exiguobacterium acetylicum*	1.03
Aquincola	0.80	*Cloacibacterium rupense*	1.00
Deinococcus	0.78	*Synechococcus rubescens*	0.86
Flectobacillus	0.78	*Piscinibacter aquaticus*	0.85
Sediminibacterium	0.73	*Acinetobacter gyllenbergii*	0.78
Others	22.28	*Ralstonia pickettii*	0.78
		Cloacibacterium normanense	0.76
		Burkholderia phenazinium	0.76
		Flavobacterium indicum	0.72
		Others	29.57

Table 5.1: Relative abundance of some well-characterized bacterial genera and species frequently identified in this study by 16S metagenomics. Altogether, a total of 1600 bacterial genera and 3598 species were identified through out the study.

Dhaka University Institutional Repository

Monitoring the community composition across different habitats and over time, we observed that composition varied considerably less between habitats (Fig 5.3) than within habitats over time (Fig 5.4). Each habitat harbored a characteristic microbiota with a dynamic set of abundant taxa, which changes over time (Fig. 5.3). Within habitats, variation in natural community and their relative abundance over time was remarkably noticeable than the variation observed in between different habitats on a given day. Finally, after accounting for habitat and seasonality, variation was significantly less in each sampling round (sampling visit) than over seasonality, which might be the result of community response to the changes in the environmental parameters.

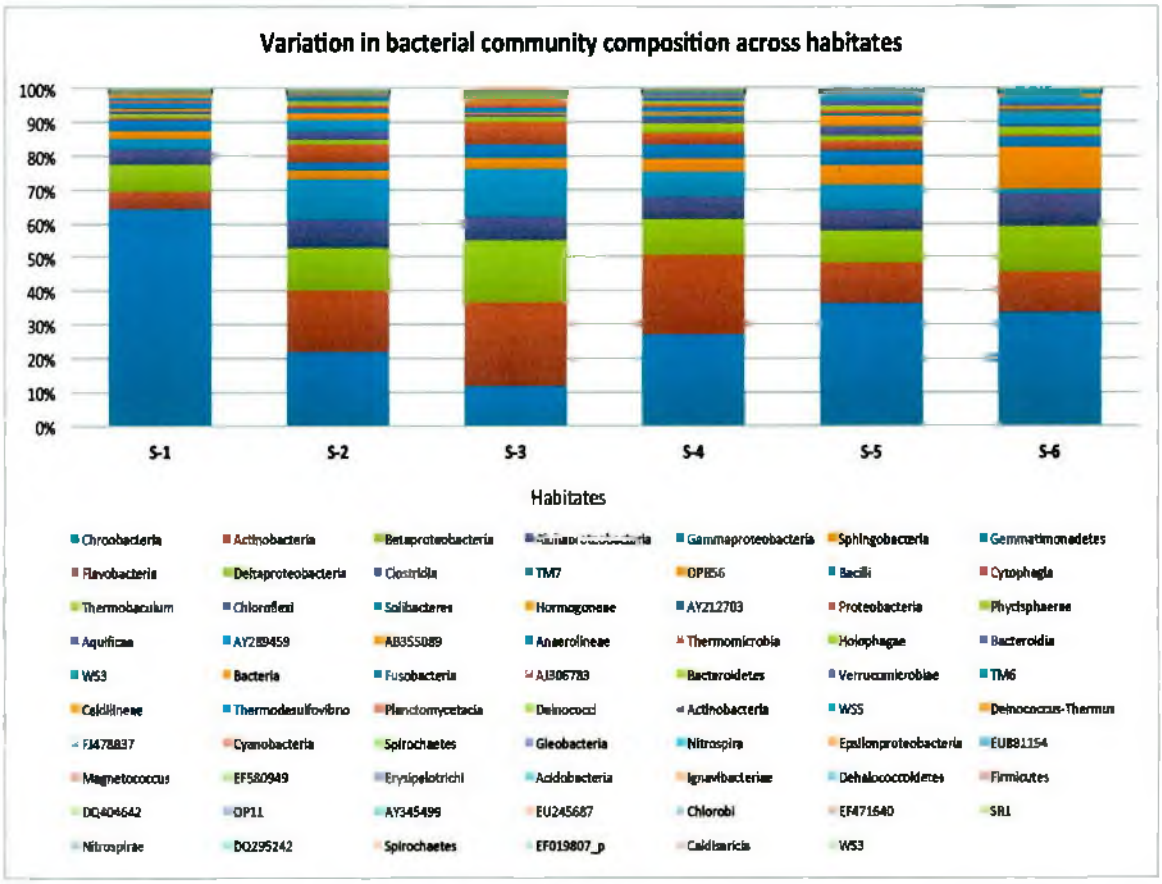
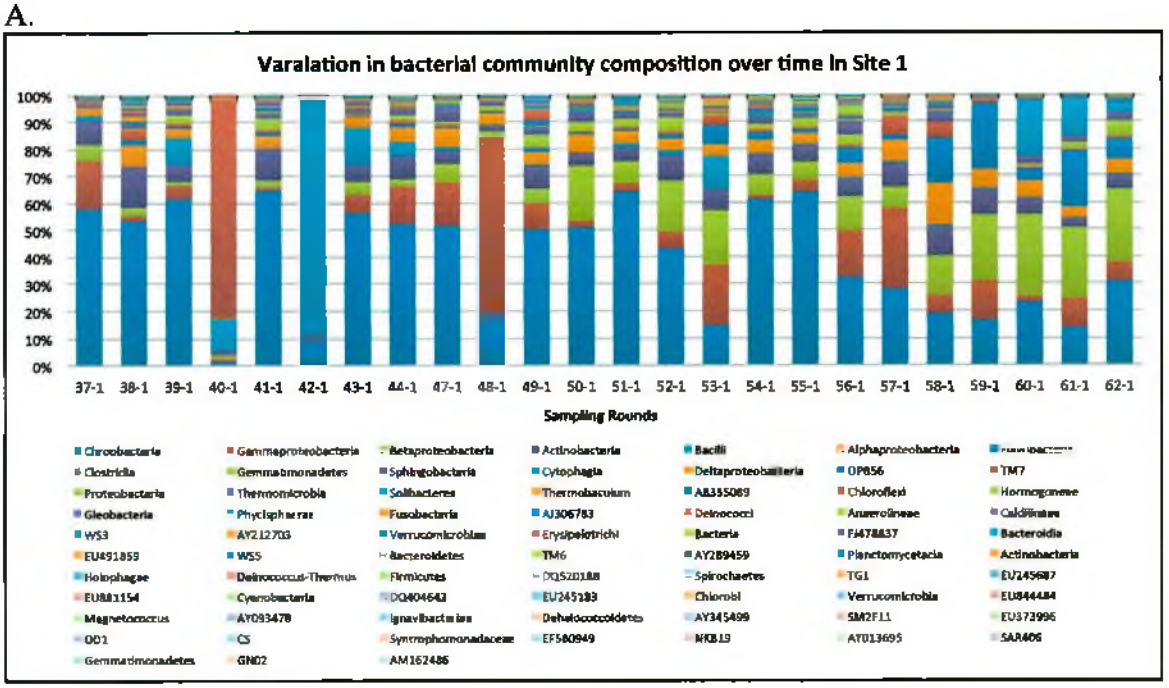


Figure 5.3: Variation in bacterial community composition across different habitats (site 1 to 6) of Mathbaria, Bangladesh at a certain time (March 2006). The figure also represents the relative abundance of different bacterial classes identified in each site.

Dhaka University Institutional Repository

A.



B.

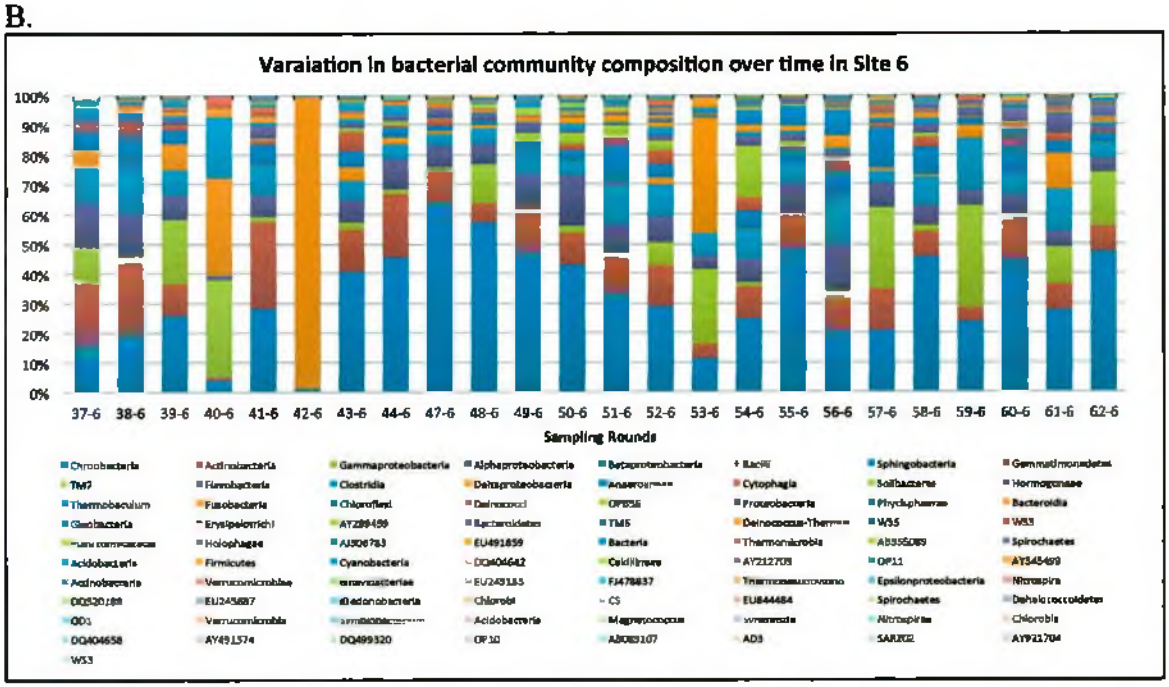


Figure 5.4: Characteristics microbiota of two distinct habitats (site 1 and 6) over time. The figures represent the relative abundance of different bacterial classes identified in site 1 (A) and 6 (B) over 24 sampling visits (sampling rounds) over a period of one year (Aug 2005 to Aug 2006).

Dhaka University Institutional Repository

We used a rank abundance curve, which orders each taxonomic group by their relative abundance and provide important insights into understanding the species richness and species evenness across habitats. Figure 5.5 demonstrated the rank abundance curves across different taxonomic level (class to species) for all six habitats sampled in this study, and indicated that, at each taxonomic level, habitat to habitat rank abundance were comparable. However, when progressing from higher to lower taxonomic level (i.e., form class to species), we observed a decreasing trend for number of organisms over total community concentrating at 95% of the abundance, as indicated by the fewer number of black lines crossing over the red line. This means at species level, curves are more dispersed or less tight, indicating that there were significant variations in their relative abundance over different sampling rounds. This is in strong agreement with our previously noted observation that within habitats there were significant temporal variations.

Organism's richness accumulation curves shown in Figure 5.6 demonstrated the species richness for combinations of habitats sampled at each time point (rounds). These curves portray the average pooled species richness when rounds were combined successively together at each habitat for the five taxonomic levels (class, order, family, genus and species). The species accumulation curves calculated the average species richness (and the standard deviation) for larger combinations of sampling rounds, such as combinations of three or combinations of 24 rounds, until all possible sizes of combinations were investigated. This demonstrated that most samples showed no sign of reaching an asymptote in species richness, as the curves don't flatten off even after accumulating all of the sampling rounds together. This suggests that we have not sampled enough to cover the entire biodiversity, and therefore, did not capture nearly all the species in the study area. It also indicated that, if we were to continue to collect samples or increase the number of sampling rounds, the diversity would increase. However, this trend appeared to be similar between habitats, but when we progressed with the taxonomic levels from class to species, the trends were becoming more distinguishable. For example, at genus level, maximum species richness for habitat (or site) 1, 5 and 6

Dhaka University Institutional Repository

was around 900, while for habitat 2 to 4, it was over 1000 (Fig. 5.6). Since species richness depends on sample size, it is expected that we would not have recorded all the species that occurred in the survey area.

We used the Shannon diversity index as a quantitative measure to reflect the evenness and diversity of different taxa over time. It also helps quantifying the uncertainty associated with the prediction. If the index value is low, the uncertainty in identifying an individual that is taken at random is reduced, in other words the diversity and the evenness is lower. Interestingly, there were no cholera cases observed when the Shannon Index is high (i.e., over 250) meaning that during a cholera epidemic period, microbial communities in the environment become relatively less diverse and less even (Fig. 5.7) in distribution. While plotting the mean Shannon index along with environmental isolation of *V. cholerae* O1, and Non-O1/O139, we observed that *V. cholerae* Non-O1/O139 were readily isolated when Shannon index was relatively low. However, in most cases, toxigenic *V. cholerae* O1 isolation demonstrated a similar relationship with Shannon index as has been seen for cholera cases, except for some isolations in the Fall (sampling round 37-42). However, it has already been documented that cholera seasonality in Mathbaria was quite distinct (Alam, Hasan, 2006; Alam, Sultana, 2006). In our previous analysis (see Chapter 2), we demonstrated that despite toxigenic *V. cholerae* being isolated twice a year (Fall and Spring) in this region, only the Spring peak of *V. cholerae* occurrence, when the highest number (66%) of isolations was observed, led to localized cholera outbreaks. Our observation of Shannon index relationship with *V. cholerae* isolation and cholera cases help explain these dynamics. In the Spring, although accumulated frequencies of all species were high, the overall species diversity and their evenness in distribution appeared to be much lower, as evidenced by lower Shannon indexes, but isolation of both toxigenic and non-toxigenic *V. cholerae* was much higher than other times of the year. Furthermore, we observed disruption in the *Vibrio* value at round 49 in three of habitats (i.e., site 5, 4 and 3) before cholera cases rose (Figure 5.9) followed by a steady increase in the value in other sites when first cases were observed in round 51 (Figure 5.9). A large increase was observed next in site 3 (round 52) then in site 2 (round 53). Levels went back down at round 54 and the number of

Dhaka University Institutional Repository

cholera cases followed these trends. This prompted us to hypothesize that selection is occurring within the community, enabling enrichment of certain species like *V. cholerae* to a level adequate to start the first index case, which after gaining additional fitness and amplification cycles through human passage appear in epidemic proportions. In Fall, toxigenic *V. cholerae* was also isolated, though in lower frequency (27%), accumulated frequencies of all species were higher than average but lower than that of Spring, and no cholera cases were observed. The Shannon index was high, indicating an increase in species diversity and their evenness in the community. This indicated that seasonal drivers might trigger the growth of the entire community with no selection, as both species diversity and their evenness in distribution was also increased. This level of enrichment thus might not be adequate to initiate the index case, thereby obviating an epidemic. These observations, however, must be replicated in other areas of endemic zones like Bakergonj, Dhaka, and Matlab, Bangladesh where two seasonal cholera outbreaks occur.

Dhaka University Institutional Repository

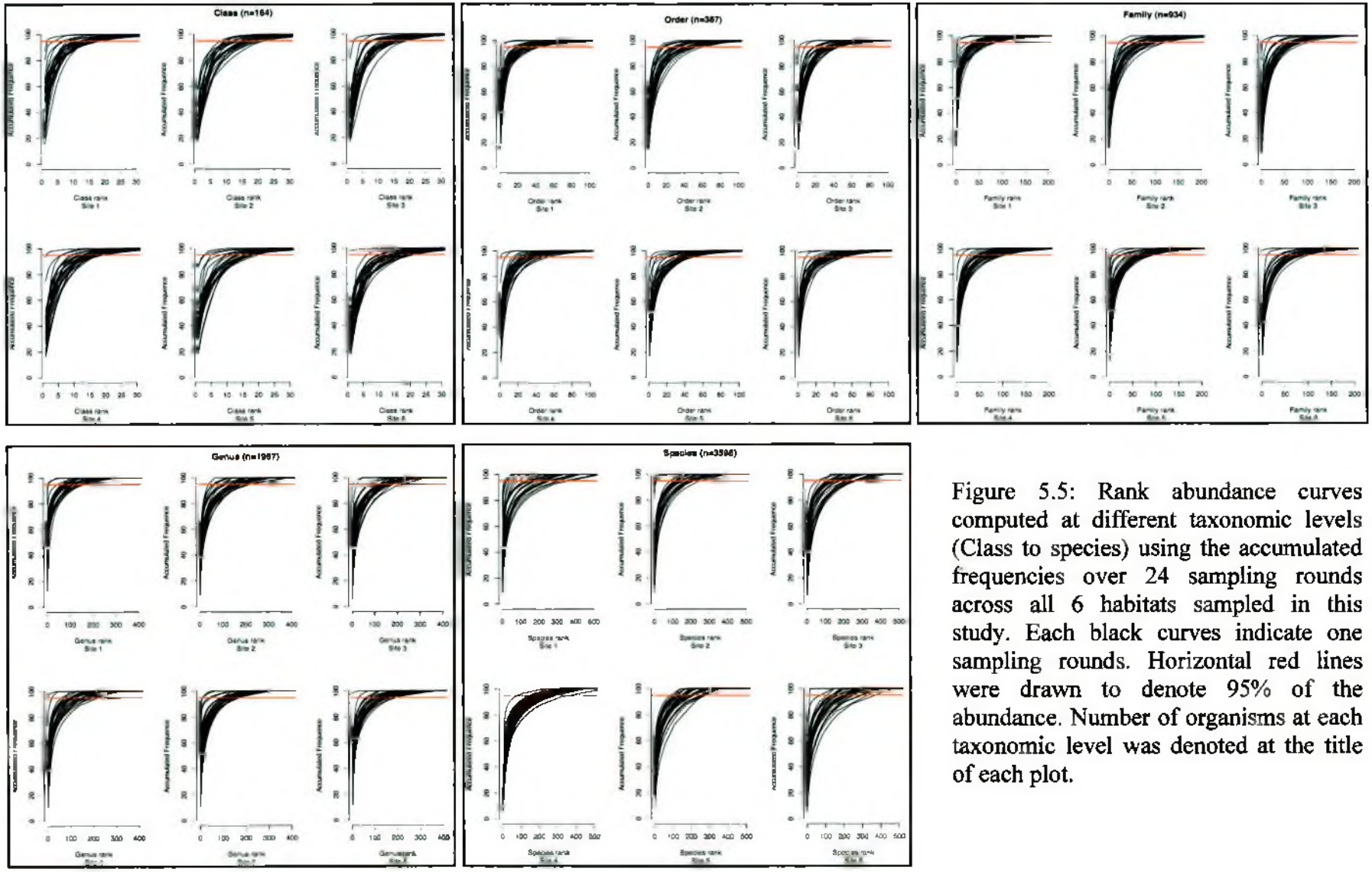


Figure 5.5: Rank abundance curves computed at different taxonomic levels (Class to species) using the accumulated frequencies over 24 sampling rounds across all 6 habitats sampled in this study. Each black curves indicate one sampling rounds. Horizontal red lines were drawn to denote 95% of the abundance. Number of organisms at each taxonomic level was denoted at the title of each plot.

Dhaka University Institutional Repository

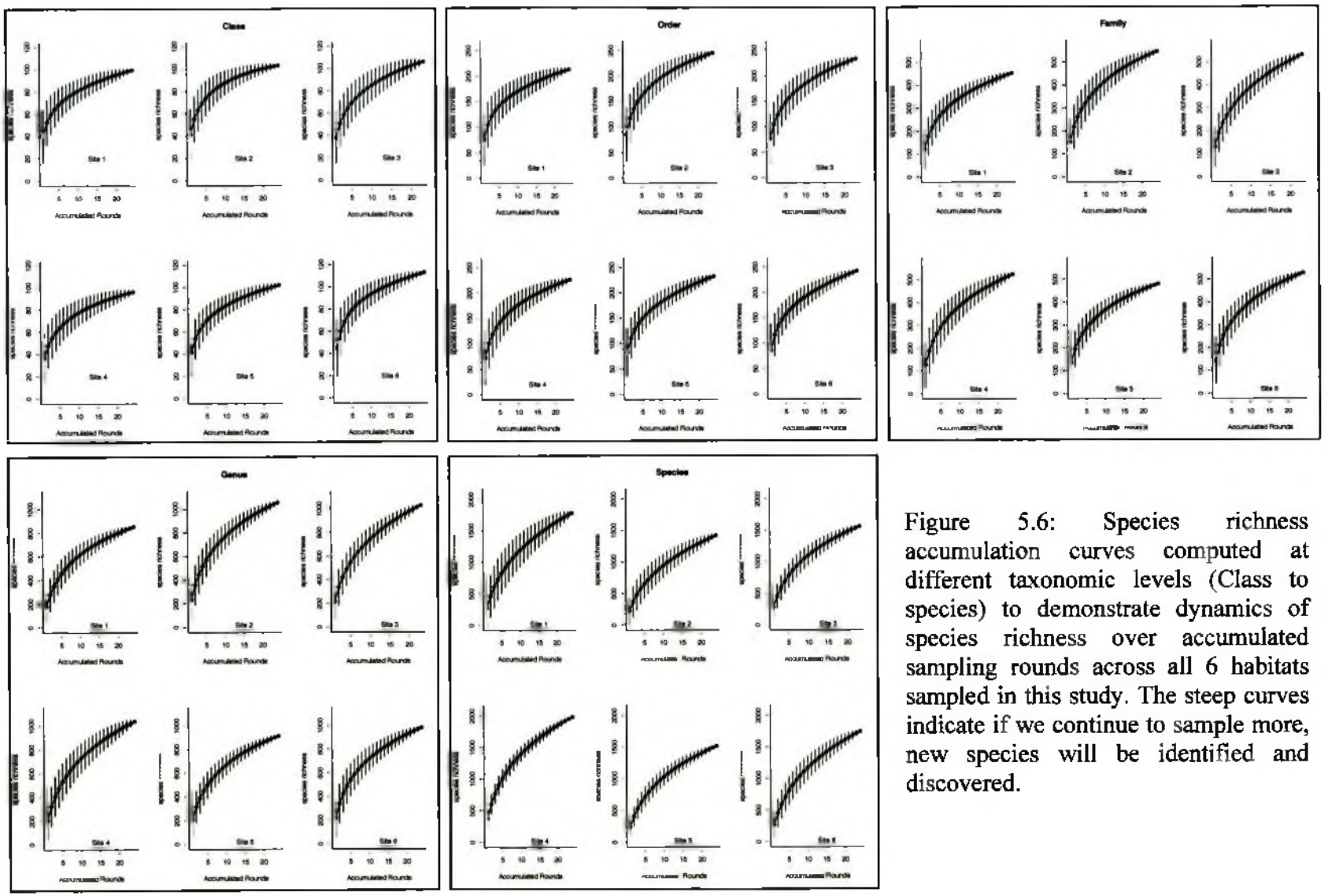


Figure 5.6: Species richness accumulation curves computed at different taxonomic levels (Class to species) to demonstrate dynamics of species richness over accumulated sampling rounds across all 6 habitats sampled in this study. The steep curves indicate if we continue to sample more, new species will be identified and discovered.

Dhaka University Institutional Repository

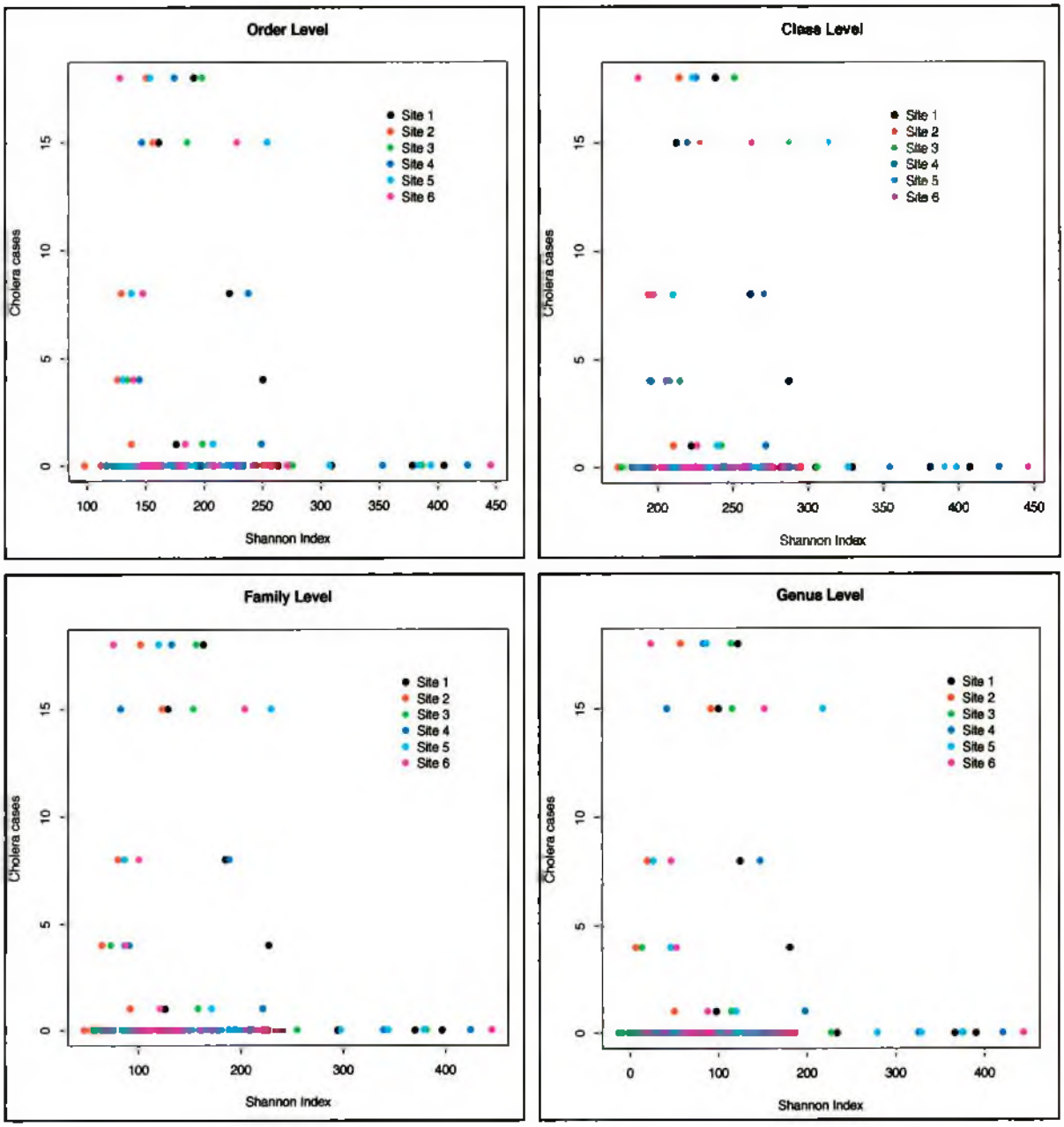


Figure 5.7: Shannon index plotted with recorded cholera cases in this study. Shannon diversity was calculated separately for each habitat (site) and was colored accordingly. Typically higher Shannon index indicate increased diversity and evenness.

Dhaka University Institutional Repository

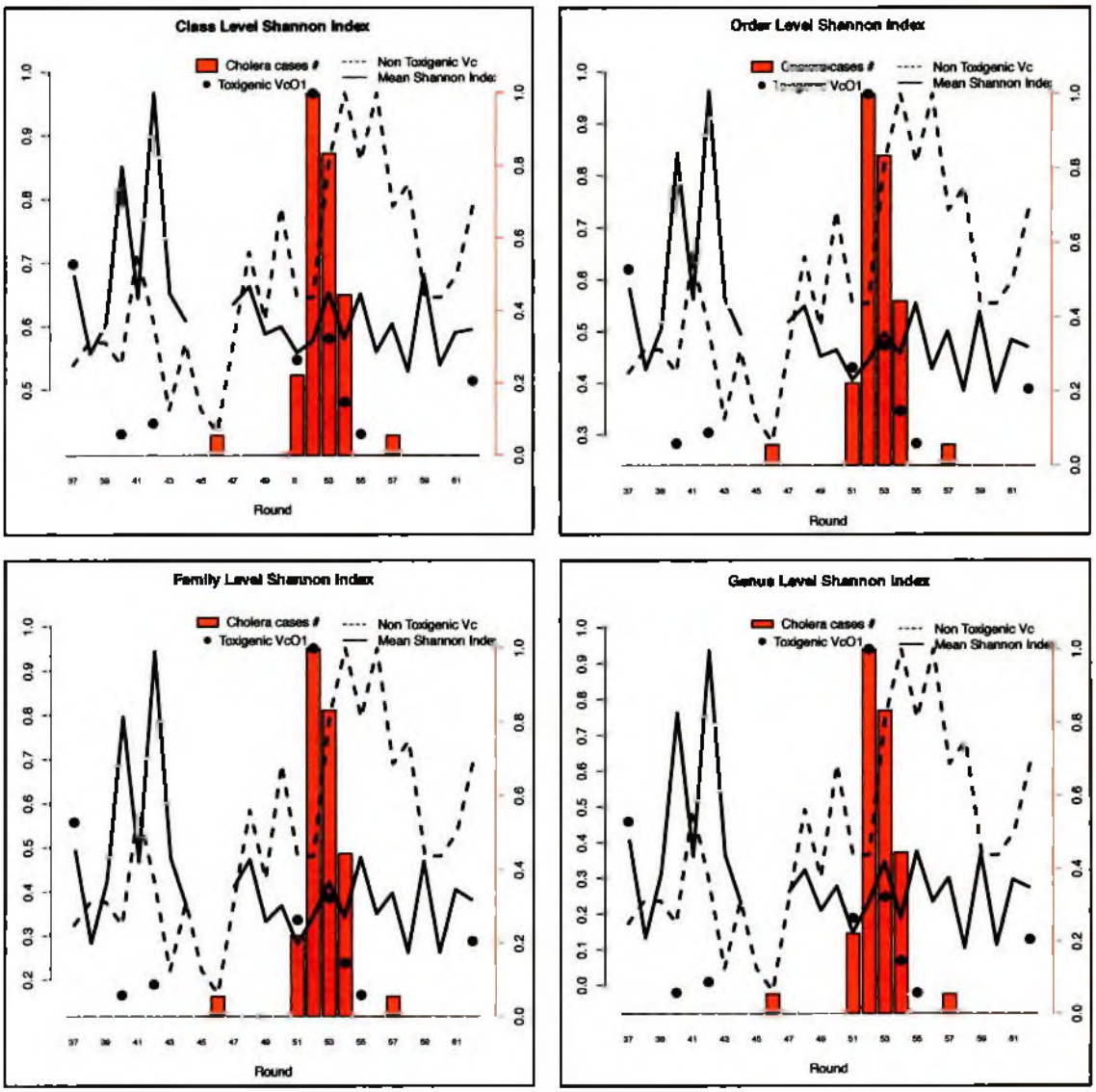


Figure 5.8: Temporal diversity of Shannon index plotted along with environmental *V. cholerae* detection and recorded cholera cases in this study. Shannon diversity was calculated separately for each habitat, and mean value of Shannon index was used here. The gap in the Shannon index curve was due to lack of samples for Round #46.

Dhaka University Institutional Repository

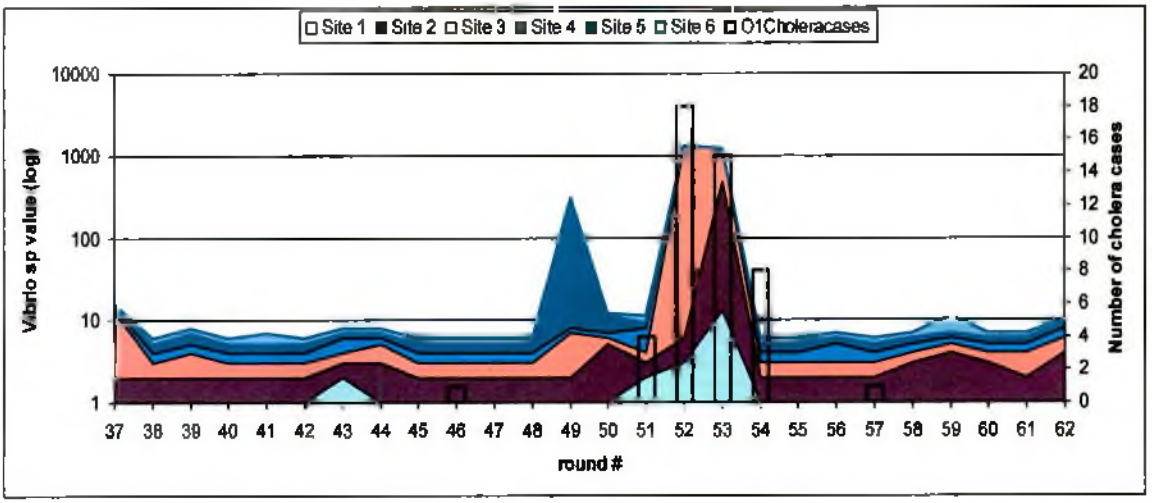


Figure 5.9: Plot of *Vibrio* sp. values against the number of cholera cases observed in Mathbaria, Bangladesh.

To further understand the dynamics of microbial communities among these habitats, we measured the biodiversity over spatial scales by computing alpha, beta, and gamma diversity as described by Whittaker (1972). Within the context of the present study, alpha diversity refers to the mean number of species over the six habitats, beta diversity, on the other hand, was measured comparing the total number of species unique to each habitat. Gamma diversity was measured for the entire geographic-scale species diversity by calculating total number of different species over six different habitats. The beta diversity index offers a quite interesting trend from Family to genus level (Figure 5.9), to begin with, it confirmed the Shannon Index analyses by showing a large variability and high values indicating lot of heterogeneity between sites for round 40 and 42. Then for Family to genus levels, the beta diversity was the lowest during the time of cholera in the Mathbaria population. It suggests that all the six habitats during this time were very similar in terms of diversity composition. This also appeared to be true for the detection of the toxigenic *V. cholarae* as they were observed simultaneously with cholera cases. The pairwise comparisons between the three indexes at different taxonomic levels are shown in Figure 5. 11. From Class to genus level, pairwise comparison followed similar trends, which was expected as the calculation of the index shared some properties. However, at the exception of the three red points (during cholera time), there was no

Dhaka University Institutional Repository

specific trend before or after cholera time. The cholera cases for these specific points correlated with highest value of Gamma and Alpha diversity and lowest value for Beta diversity (Figure 5.12).

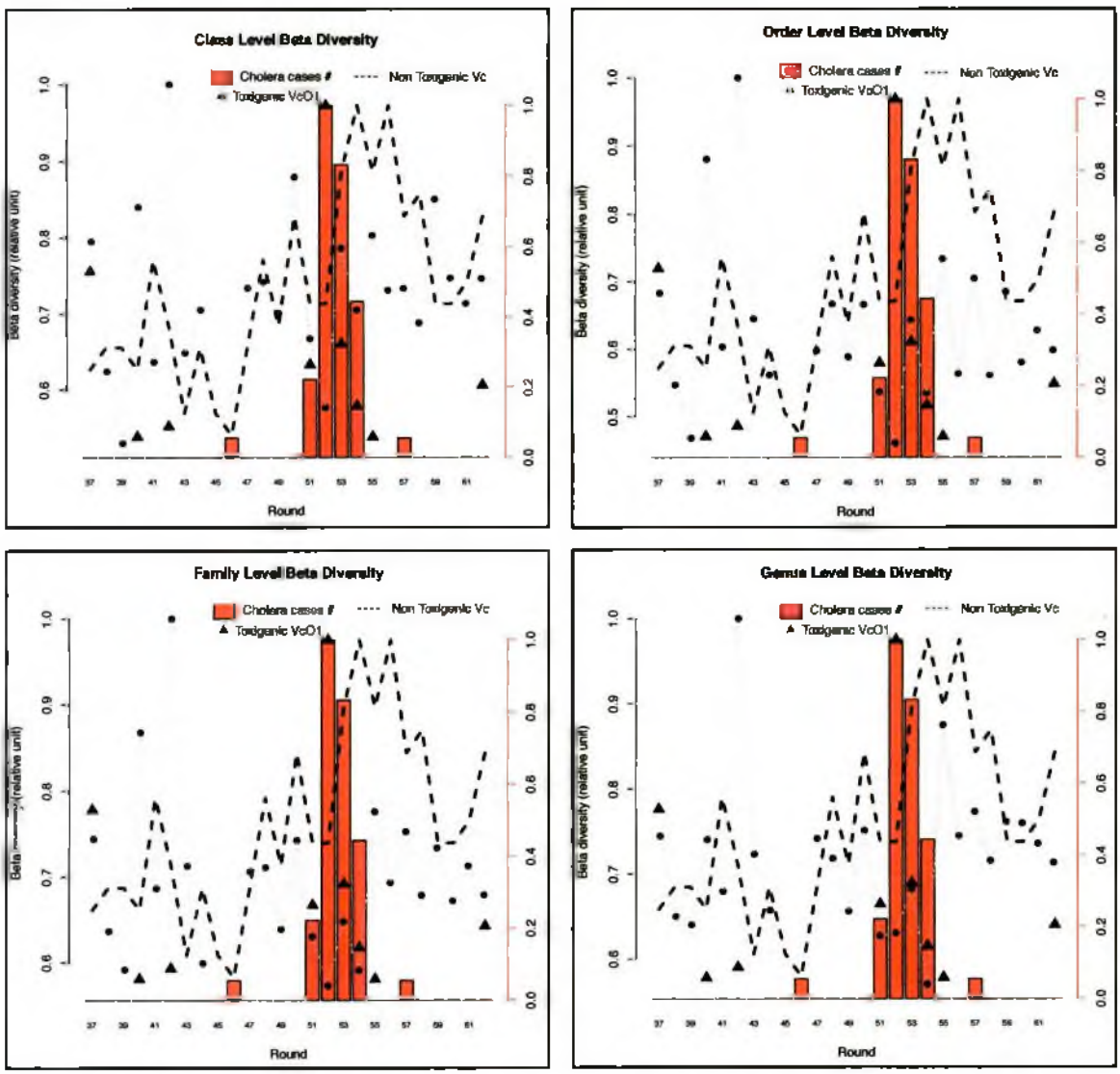


Figure 5.10: Temporal variation of beta diversity with the occurrence of cholera cases and environmental isolation of toxigenic and non-toxigenic *V. cholerae* in Mathbaria, Bangladesh. Dotted line with solid black circles represents beta diversity.

Dhaka University Institutional Repository

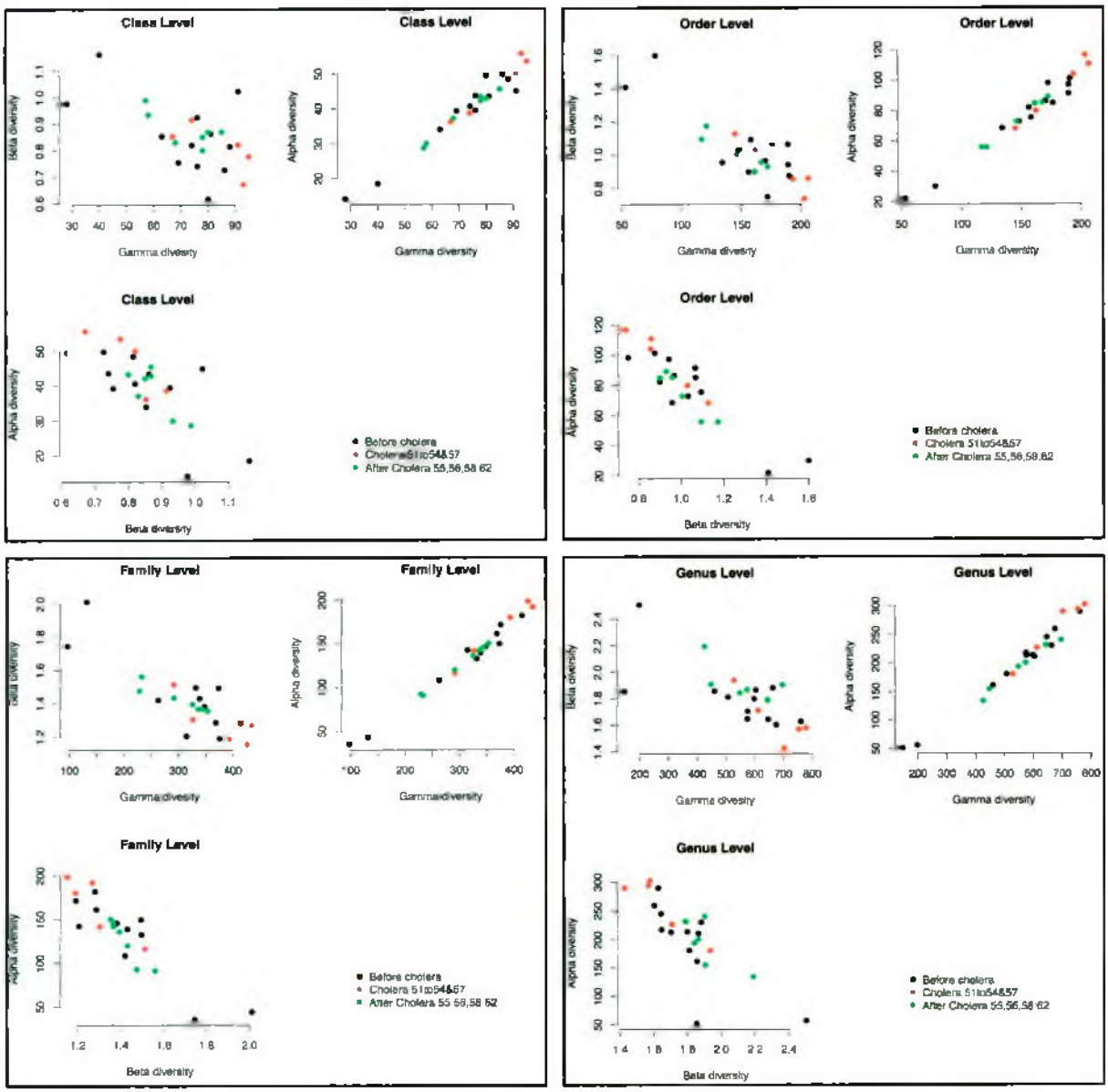


Figure 5.11: Pairwise comparison between alpha, beta and gamma indexes at different taxonomic levels. Data from all six sites were pooled together to represent the indexes for each sampling rounds. Rounds were color coded to distinguish rounds before cholera, during cholera and after cholera.

Dhaka University Institutional Repository

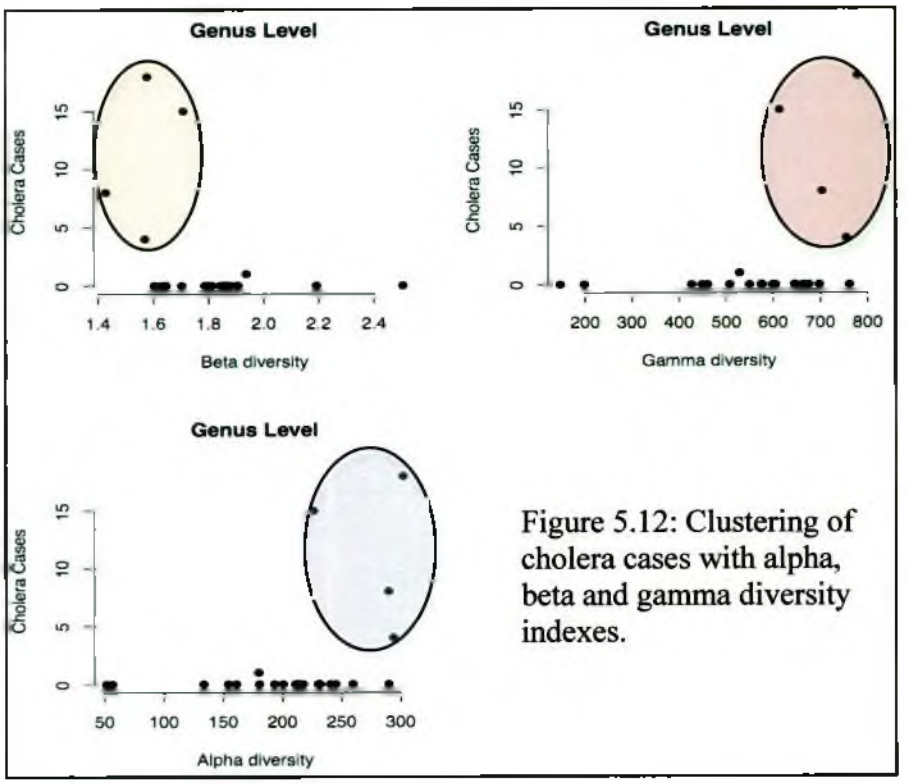


Figure 5.12: Clustering of cholera cases with alpha, beta and gamma diversity indexes.

Finally, we assessed differences in the overall community composition at species level during and in between seasonal cholera epidemics in Mathbaria. Figure 5.13 demonstrated a shift in microbial community with increased abundance of certain species including *V. cholerae* during cholera epidemics, and provided additional support for our previous observation that a selection is going on in the community when there was an outbreak. Figure 5.14 demonstrated that during cholera epidemic certain species became over represented in the natural community while others became underrepresented or absent and there were less heterogeneity in the community during the time of an outbreak compared to the other time when no cholera case was recorded. Therefore, monitoring the endemic regions for such distortion in the community may provide an important clue to predict a forthcoming outbreak.

Dhaka University Institutional Repository

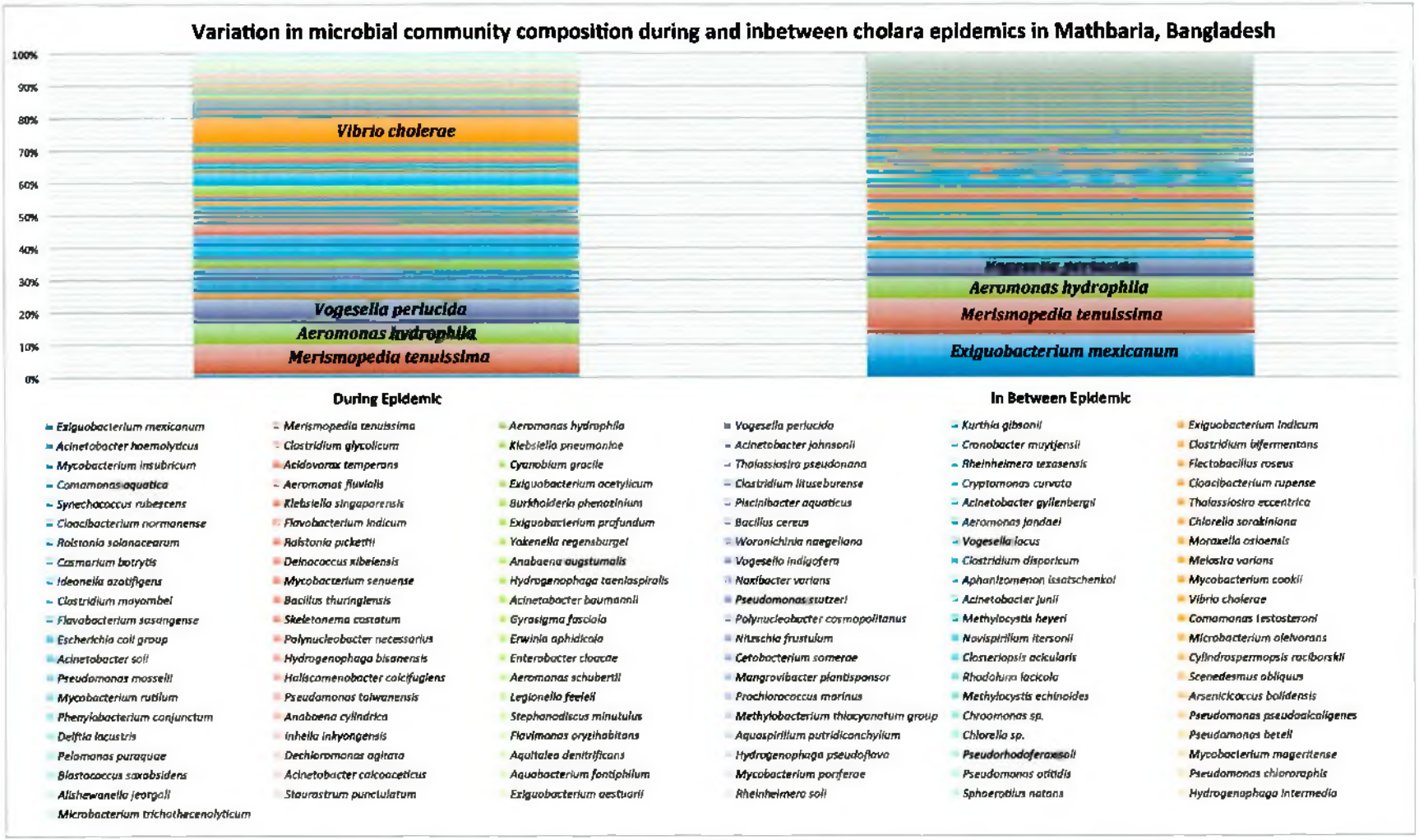


Figure 5.13: Variation in microbial community composition during and in between seasonal cholerae epidemics in Mathbaria, Bangladesh

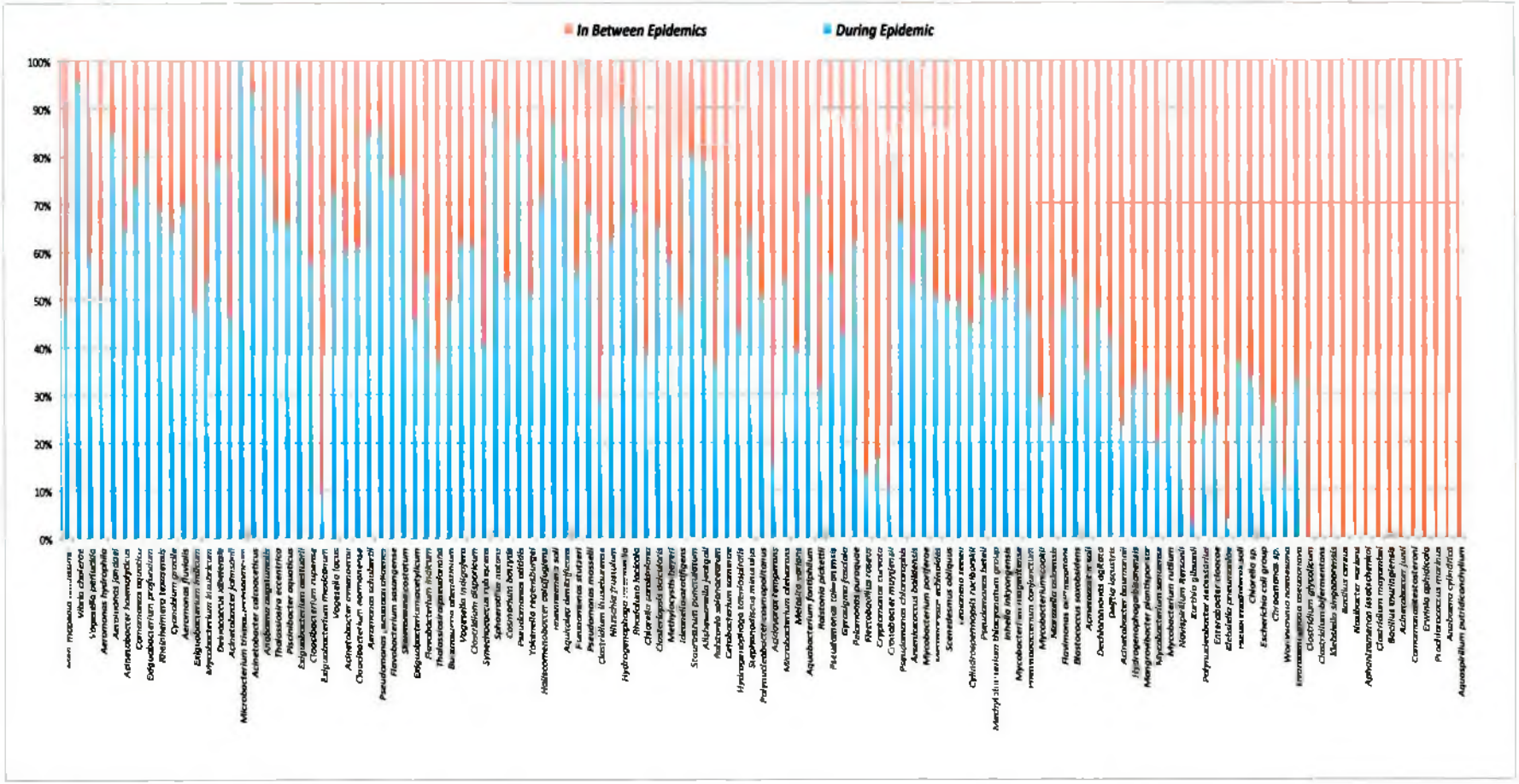


Figure 5.14: Relative abundance of major species detective during and in between seasonal cholera epidemics in Mathbaria, Bangladesh.

Dhaka University Institutional Repository

Dhaka University Institutional Repository

Conclusion:

Metagenomics, a powerful tool that recently has been introduced to analyze natural microbial communities, is a sequence-based analysis of the collective microbial genomes contained in an environmental sample. In this study we successfully implemented metagenomics to characterize the environmental community composition and demonstrated that the aquatic environment of Mathbaria harbor genetically diverse and complex microbial community including primary members of both freshwater and coastal ecosystems, which appears to be the result of frequent coastal water intrusion into the fresh water system in the region. Tremendous diversity was uncovered at every taxonomic level, and many bacterial genera pathogenic to human and animals were detected with possible implications for agriculture and public health. Careful analysis of data for spatial and temporal variation in microbial communities demonstrated that temporal variation in a natural community and their relative abundance was more pronounced than the variation between habitats on a given day, and *Vibrio cholerae* abundance and distribution followed a similar trend. This suggests that the abundance and diversity of the microbial community including the *Vibrio* populations, exhibit a complex seasonal distribution driven by temporal fluctuations of environmental parameters, which helps explain the seasonal occurrence of *V. cholerae* in the environment. Our observation from diversity dissimilarity analysis, as well as species diversity and evenness in distribution, reaffirmed each other and suggests that selection is going on in the circulating community, and a comparison of the community diversity along with observed cholera cases in this region demonstrate that high diversity, alone cannot explain cholera epidemics. Rather, we found a good correlation of cholera cases when the ecosystem experiences high community diversity (alpha and gamma) with least dissimilarity (beta) among the habitats. Taken together, the results of this study suggests that specific consortium among the indigenous population, arising as a result of positive selection, might be crucial in responding on changes in environmental parameters that allow *V. cholerae* to overcome the natural competition and emerge in epidemic proportions. However, these observations need to be replicated in other endemics zone like Bakergonj, Dhaka, and Matlab, Bangladesh where two seasonal cholera outbreaks

Dhaka University Institutional Repository

occur. Moreover, a large number of bacterial sequences, not homologous to the well-characterized bacterial sequences in the databases, indicate existence of novel bacterial taxa in this ecosystem. Nevertheless our data do not unequivocally support identification of such novel bacterial taxa as whole genome extrapolation of rRNA gene polymorphism may always not be accurate. This is in part due to the fact that 16S rRNA based analysis can be influenced by artifacts, including chimeric sequences caused by PCR amplification and sequencing errors [12]. According to a study by Ashelford K.E et al, at least 1 in 20 16S rRNA sequences currently in public repositories contains substantial anomalies [13] and it was shown in one study[12] that some metagenomics projects may overestimate species diversity because of the presence of sequencing errors and chimeric sequences. With this caveat in mind however, our study demonstrates that data generated from metagenomic studies may serve as a preliminary baseline for monitoring bacterial complexity in the environment. Similar research could be extended to other endemic areas with similar seasonality as well as geographic proximity to human residence to establish a deep inventory of circulating bacterial species using whole genome shotgun sequencing to enable a genome centric analysis of species identity and to avoid amplification bias of 16 rRNA genes. This would also provide functional characteristics of the community. Such additional studies would establish essential nucleic acid sequence baselines for monitoring of disease-causing bacterial species that may have a significant impact on livestock, agriculture and human health as well as providing data with possible implications in the forecasting of emergence and re-emergence of an infectious disease like cholera.

Dhaka University Institutional Repository

References:

1. Tekere, M., et al., *Metagenomic analysis of bacterial diversity of Siloam hot water spring, Limpopo, South Africa.* 2011.
2. Huq, A. and R.R. Colwell, *Environmental factors associated with emergence of disease with special reference to cholera.* Eastern Mediterraneal Health Journal, 1996. **2**: p. 37-45.
3. Brinkman, U., *Economic development and tropical disease. In: Wilson ME, Levins R, Spielman A, eds. 1994. Disease in Evolution: Global Changes and Emergence of Infectious Diseases. Proceedings of a conference. Woods Hole, Massachusetts, November 7-10, 1993.* Ann N Y Acad Sci, 1994. **740**: p. 1-503.
4. Wayne, L.G., *International Committee on Systematic Bacteriology: announcement of the report of the ad hoc Committee on Reconciliation of Approaches to Bacterial Systematics.* Zentralbl Bakteriol Mikrobiol Hyg A, 1988. **268**(4): p. 433-4.
5. Woese, C.R., *Bacterial evolution.* Microbiol Rev, 1987. **51**(2): p. 221-71.
6. Petrosino, J.F., et al., *Metagenomic pyrosequencing and microbial identification.* Clin Chem, 2009. **55**(5): p. 856-66.
7. Krause, L., et al., *Finding novel genes in bacterial communities isolated from the environment.* Bioinformatics, 2006. **22**(14): p. e281-9.
8. Chen, K. and L. Pachter, *Bioinformatics for whole-genome shotgun sequencing of microbial communities.* PLoS computational biology, 2005. **1**(2): p. e24.
9. Krause, L., et al., *Phylogenetic classification of short environmental DNA fragments.* Nucleic Acids Res, 2008. **36**(7): p. 2230-9.
10. Unno, T., et al., *Use of barcoded pyrosequencing and shared OTUs to determine sources of fecal bacteria in watersheds.* Environmental science & technology, 2010. **44**(20): p. 7777-7782.
11. Kim, O.S., et al., *Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species.* Int J Syst Evol Microbiol, 2012. **62**(Pt 3): p. 716-21.
12. Quince, C., et al., *Accurate determination of microbial diversity from 454 pyrosequencing data.* Nat Methods, 2009. **6**(9): p. 639-41.
13. Ashelford, K.E., et al., *At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies.* Appl Environ Microbiol, 2005. **71**(12): p. 7724-36.

Dhaka University Institutional Repository

CHAPTER 6:

COMPARATIVE GENOMICS OF CLINICAL AND ENVIRONMENTAL *VIBRIO MIMICUS*

Dhaka University Institutional Repository

Abstract

Vibrio mimicus has been the subject of taxonomic controversy, specifically whether it is a variant of *Vibrio cholerae* or a separate species. A genomic analysis was undertaken to resolve the issue. The genomes of *V. mimicus* MB451, a clinical isolate, and VM223, an environmental isolate, comprise ca. 4,347,971 and 4,313,453 base pairs (bp) and encode 3,802 and 3,290 open reading frames, respectively. As in other vibrios, chromosome I (C-I) predominantly contains genes necessary for growth and viability, while chromosome II (C-II) bears genes for adaptation to environmental change. C-I harbors many virulence genes, including some not previously reported in *V. mimicus*, such as *Vibrio* Pathogenicity Island 2 (VPI-2), mannose-sensitive hemagglutinin (MSHA), and enterotoxigenic hemolysin (*hlyA*); while C-II encodes a *Vibrio* Seventh Pandemic Island II (VSP-II) cluster of genes. Extensive genomic rearrangement in C-II indicates it is a hot spot for evolution and genesis of speciation for the genus *Vibrio*. The number of virulence regions discovered in this study (VSP-II, MSHA, HlyA, PilE, and Intl4) with no notable difference in potential virulence genes between clinical and environmental strains, suggests these genes may also play a role in the environment and that pathogenic strains may arise in the environment. Significant genome synteny with prototypic pre-7th pandemic strains of *V. cholerae* was observed and results of phylogenetic analysis support the hypothesis that, in the course of evolution, *V. mimicus* and *V. cholerae* diverged from a common ancestor with a prototypic 6th pandemic genomic backbone.

Introduction:

Vibrio mimicus is a Gram-negative, gamma-Proteobacterium, closely related to *V. cholerae*. It was first described as a biochemically atypical *Vibrio cholerae* (1). However, it is phenotypically and genotypically distinct from *V. cholerae* and can be differentiated from *V. cholerae* by 12 specific biochemical reactions including sucrose fermentation, Voges-Proskauer reaction (acetoin production from glucose), lipase production, sodium tartrate fermentation, and polymyxin sensitivity. Comparative sequence analysis of the malate dehydrogenase (*mdh*) housekeeping gene of *V. mimicus* showed mean pairwise divergence from *V. cholerae* to be

Dhaka University Institutional Repository

approximately 10 %, equivalent to the divergence of *Salmonella enterica* LT2 from *E. coli* K-12 (2, 3).

The natural habitat of *V. mimicus* is similar to that of *V. cholerae*; i.e., the aquatic ecosystem, including seawater, freshwater, and brackish water, where it has been found both as a free-living bacterium and in association with zooplankton, crustaceans, filter-feeding mollusks, turtle eggs, and fish. Infections in humans occur from consumption or exposure to these sources (4-9). *V. mimicus* human gastroenteritis is characterized by diarrhoea, nausea, vomiting, abdominal cramps, and fever. However, unlike *V. cholerae*, *V. mimicus* has not been associated with epidemics of cholera-like diarrhoea, most likely due to the fact that most isolates of *V. mimicus* do not produce cholera toxin. In fact, Chowdhury *et al.* (5) reported that fewer than 10% of clinical isolates and less than 1% of environmental strains produce CT enterotoxin. Although a number of hypothesized virulence factors have been identified in *V. mimicus,* including cholera-like enterotoxin (10), heat stable enterotoxin (11), heat labile enterotoxin (12), hemolysin (13), protease (14), phospholipase (15), arylesterase (16), siderophore (aerobactin) (17), and hemagglutinin (18), the mechanism of its pathogenesis remains unclear.

The objective of this study was to determine the genetic basis of *V. mimicus* physiology, pathogenicity, and evolution and to clarify its relationship with *V. cholerae*. To this end, we sequenced two strains of *V. mimicus*, VM223, an environmental isolate collected from a bivalve in Saõ Paulo, Brazil, and MB451, a clinical isolate from a patient with diarrhoea in Matlab, Bangladesh. Genome sequence comparison of these strains with each other and *V. cholerae* demonstrates clear species delineation for *V. mimicus* but also provides evidence of probable evolution of *V mimicus* and *V. cholerae* from a common ancestor.

Materials and Methods:

Genome sequencing. The genome of *Vibrio mimicus* was sequenced, in part, by the Joint Genome Institute (JGI) and all general aspects of the sequencing performed at the JGI are available at http://www.jgi.doe.gov/. Draft sequences were obtained from a blend of Sanger and 454 sequences and involved paired end Sanger

Dhaka University Institutional Repository

sequencing on 5- 8kb plasmid libraries to 5X coverage and 20X coverage of 454 data. The Phred/Phrap/Consed software package (www.phrap.com) was used for both sequence assembly and quality assessment (52-54). After the shotgun stage, reads were assembled with parallel phrap (High Performance Software, LLC). Repeat resolution was performed using Dupfinisher (55). Gaps between contigs were closed by editing in Consed and by several targeted finishing reactions that included transposon bombs, primer walks on clones, primer walks on PCR products, and adapter PCR reactions. Gene-finding and annotation were achieved using the RAST server (56).

Comparative genomics. To compare more precisely, we reannotated six completed *Vibrio* genomes available in the database, *Vibrio cholerae* O1 El Tor str. N16961, *Vibrio cholerae* O1 classical O395, *Vibrio harveyi* ATCC BAA-1116, *Vibrio parahaemolyticus* RIMD 2210633, *Vibrio vulnificus* YJ016, and *Vibrio fischeri* MJ11, using the same annotation (RAST) as used for the two *V mimicus* genomes. Genome to genome comparison was performed using three approaches as described by Chun *et al.* (26). To determine average nucleotide identity (ANI) and genetic distance between strains and to assign strains to species groups, a reciprocal best match BLASTN analysis was performed for each genome. The average similarity between genomes was measured as the average nucleotide identity (ANI) of all conserved genes, as described by Konstantinidis and Tiedje (57).

Identification of genomic islands and repeat sequences. We employed the Island Viewer application for GI predictions, which uses three methods for genomic island prediction, namely, IslandPick, IslandPath-DIMOB and SIGI-HMM (28). As this application requires the genome to be completed, the clinical *V mimicus* MB451 sequence was used for GI prediction and the tandem repeat finder program (58) to identify the repeat sequences.

Phylogenetic analyses based on genome sequences. A set of orthologues for each ORF of *V. cholerae* N16961 was obtained through comparison with different sets of strains and individually aligned using the CLUSTALW2 program (59). The resultant multiple alignments were concatenated to generate genome scale alignments that were subsequently used to reconstruct the neighbor-joining phylogenetic tree

Dhaka University Institutional Repository

(60). The evolutionary model of Kimura (61) was used to generate the distance matrix and the MEGA program (62) was used for phylogenetic analysis.

Results and Discussion:

Genome features. The combination of traditional Sanger sequencing and pyrosequencing yielded high quality assemblies of the genomes of *V. mimicus* MB451 (*Accession# ADAF00000000*) and VM223 (*Accession# ADAJ01000000*). The genome of *V. mimicus* MB451 consists of two circular chromosomes of 2,971,217 and 1,304,309 bp with an average G+C content of 46% and 45%, respectively (Figure 6.1, Table 6.1) and a plasmid of 37,927 bp. The C-II of *V. mimicus* MB451 was closed, whereas C-I was obtained in a single contig but not closed due to a gap of *ca.* 3.5-4.0 Kb, most likely a part of an rDNA repeat. RAST subsystem based annotation (15) identified 3,802 predicted coding sequences (CDSs) and 119 RNAs in the genome of *V. mimicus* MB451. The genome of *V. mimicus* VM223 comprises eight contigs encoded on the two chromosomes, C-I (contig 53-51-47-50-46-49, in order), and C-II (contig52 and 48). The predicted size of the genome, ca. 4,347,971 bp, is similar to that of *V. mimicus* MB451 with 3,290 CDSs and 111 RNAs (Table 6.1). *Ca.* 19% and 17% of the CDSs in *V. mimicus* MB451 and VM223 were annotated as hypothetical proteins, including proteins that are conserved in other bacteria. Like *V. cholerae*, most of the genes required for growth and viability are located on C-I. C-II contains relatively more hypothetical proteins. However, several genes believed to be important for normal cell function, for example, genes encoding ribosomal proteins L20 and L35, alkylphosphonate utilization operon protein PhnA, Uxu operon transcriptional regulator encoded by *uxuR*, are also encoded on C-II. The overall subsystem category distributions of both *V. mimicus* genomes were similar to each other and to *V. cholerae* (Fig. 6.2).

***V. mimicus* plasmid**. *V. mimicus* MB451 contains a plasmid with high G+C content (50%), and encodes 56 CDSs, with an average length of 534bp (Fig 6.1, Table 6.1). Ca. 79% of the CDSs were annotated as hypothetical proteins. Most of the coding sequences with functional assignment are phage related proteins including phage integrase, phage portal protein, bacteriophage tail assembly protein etc. No noticeable similarity with sequences in the database was observed, except for a few

Dhaka University Institutional Repository

ORFs at the beginning of the sequence with significant similarity with genes of Prophage 1 of *Salmonella typhimurium* str. LT2. Moreover, similarity of the plasmid G+C content along with the prophage G+C (51%), presence of multiple phage proteins, and similarity of the few ORFs to those of a prophage in *Salmonella* indicate that this plasmid may, in fact, be an extra chromosomal temperate phage.

***V. mimicus* Super Integron**. Super integrons (SI) are prevalent among *Vibrio* species and can be highly variable even within the same species (19). Both *V. mimicus* strains possess SI on C-II that are 20-50 Kb larger than the SI of *V. cholerae* O1 El Tor N16961 (20). Significant variation in the SI of *V. mimicus* was observed when compared to each other and against *V. cholerae* SI (Table 6.1). However, integron integrase *IntI*4, of *V. cholerae* N16961 shows 82.5 and 83.7% nucleotide similarity to those of *V. mimicus* MB451 and VM223, respectively. In *V. mimicus* MB451, the SI encompasses VII_000636 -774 (145 KB) encoding 139 ORFs, whereas in VM223, the SI encompasses VMA_000683-779 (175 KB) on contig 52 and encodes 96 ORFs. Most of the genes (55% in MB451 and 57% in VM223) are hypothetical in nature and the coding sequences of the functional ORFs are related to homologous sequences of a wide range of microorganisms. This finding is in accordance with the postulate that the SI of the genus *Vibrio* serves as a capture system for acquiring DNA from the surrounding environment (21, 22). The SIs of both strains of *V. mimicus* encode several gene products involved in drug resistance, DNA metabolism, transcriptional regulation, and plasmid stabilization.

Dhaka University Institutional Repository

Table 6.1: General features of the *Vibrio mimicus* genome compared with other *Vibrio* spp.

Feature	*V. mimicus* MB-451	*V. mimicus* VM223	*V. cholerae* N16961	*V. parahaemolyticus* RIMD 2210633	*V. harveyii* ATCC BAA-1116	*V. vulnificus* CMCP6
Chromosome size						
Chromosome I (Mb)	2.97	4.347	2.96	3.29	3.77	3.28
Chromosome II (Mb)	1.3		1.07	1.88	2.2	1.84
Plasmid (Kb)	37.9					
G+C content (mol%)						
Chromosome I	46	46	47	45	45	46
Chromosome II	45		46	45	45	47
Plasmid	50					
Protein coding sequences						
Chromosome I	2,321	3290	2,742	3,080	3,546	2,926
Chromosome II	1,125		1,093	1,752	2,374	1,562
Plasmid	56					
Average length (bp)						
Chromosome I	984	1032	948	952	923	887
Chromosome II	955		832	946	890	977
Plasmid	534					
Percent of coding region						
Chromosome I	87	78	87	86	85	83
Chromosome II	82		84	86	86	86
Plasmid	79					
Ribosomal RNA operons						
Chromosome I	7	7	8	10	10	8
Chromosome II	0		0	1	1	1
Number of tRNAs						
Chromosome I	97	90	94	112	105	98
Chromosome II	0		4	14	16	13
Number of Integrons						
Chromosome I	0	SI*	0	1	NK	SI
Chromosome II	SI		SI	0	NK	0

SI= Super Integron; *: found at small chromosome (C-II); NK= Not known

Dhaka University Institutional Repository

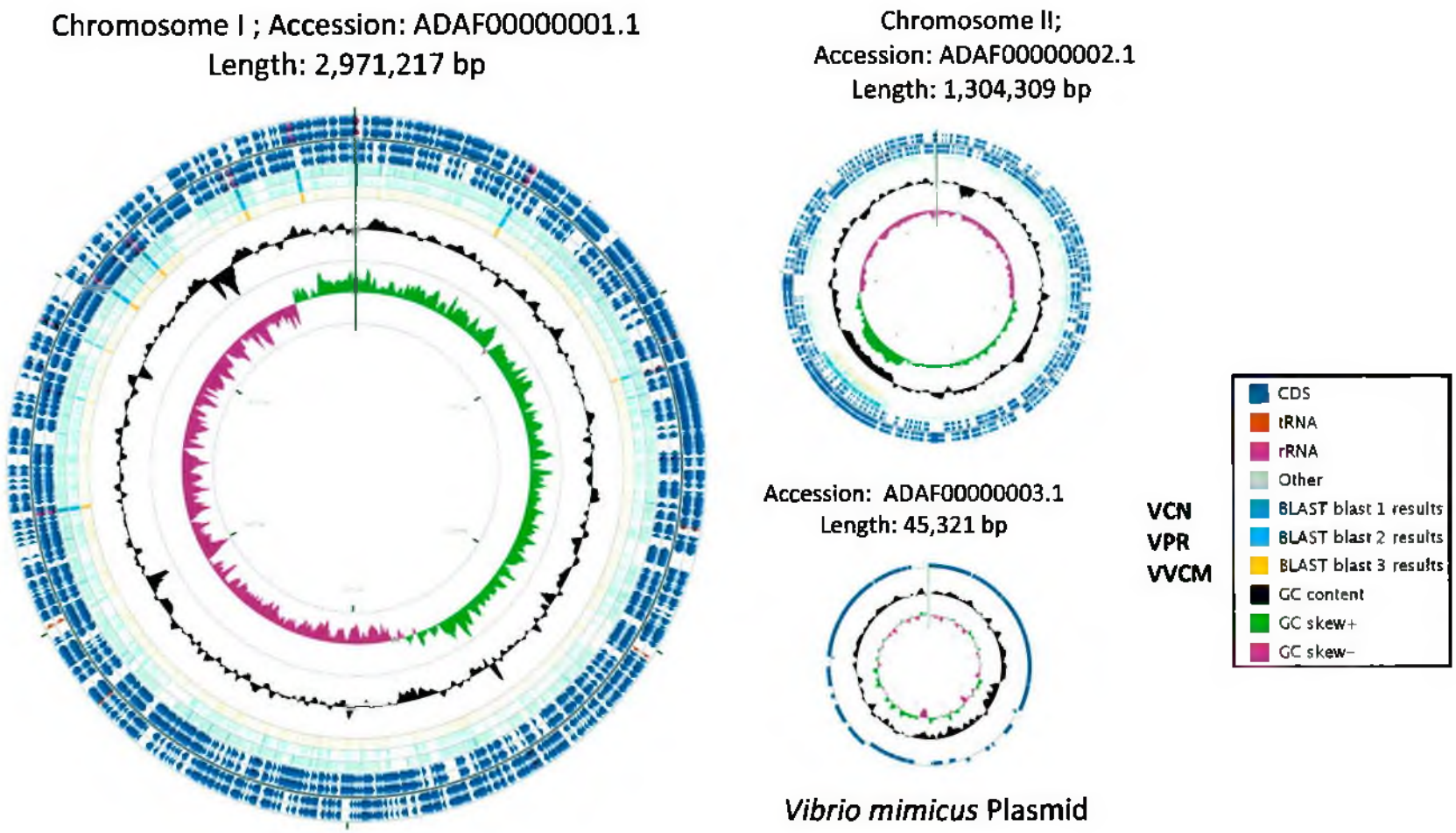


Figure 6.1: Circular maps of the chromosomes and plasmids of *V. mimicus* MB451Genome. From the outside in, the first-second and third-fourth circles show CDSs transcribed clockwise and anticlockwise, respectively. Circle 5, 6 & 7 displayed the Blast results against *V. cholerae* N16961 (VCN), *V. parahaemolyticus* RIMD 22110633 (VPRIMD) and *V. vulnificus* CMCP6 (VVCM) , the eighth circle shows the GC content, and the ninth circle shows the GC skew.

Dhaka University Institutional Repository

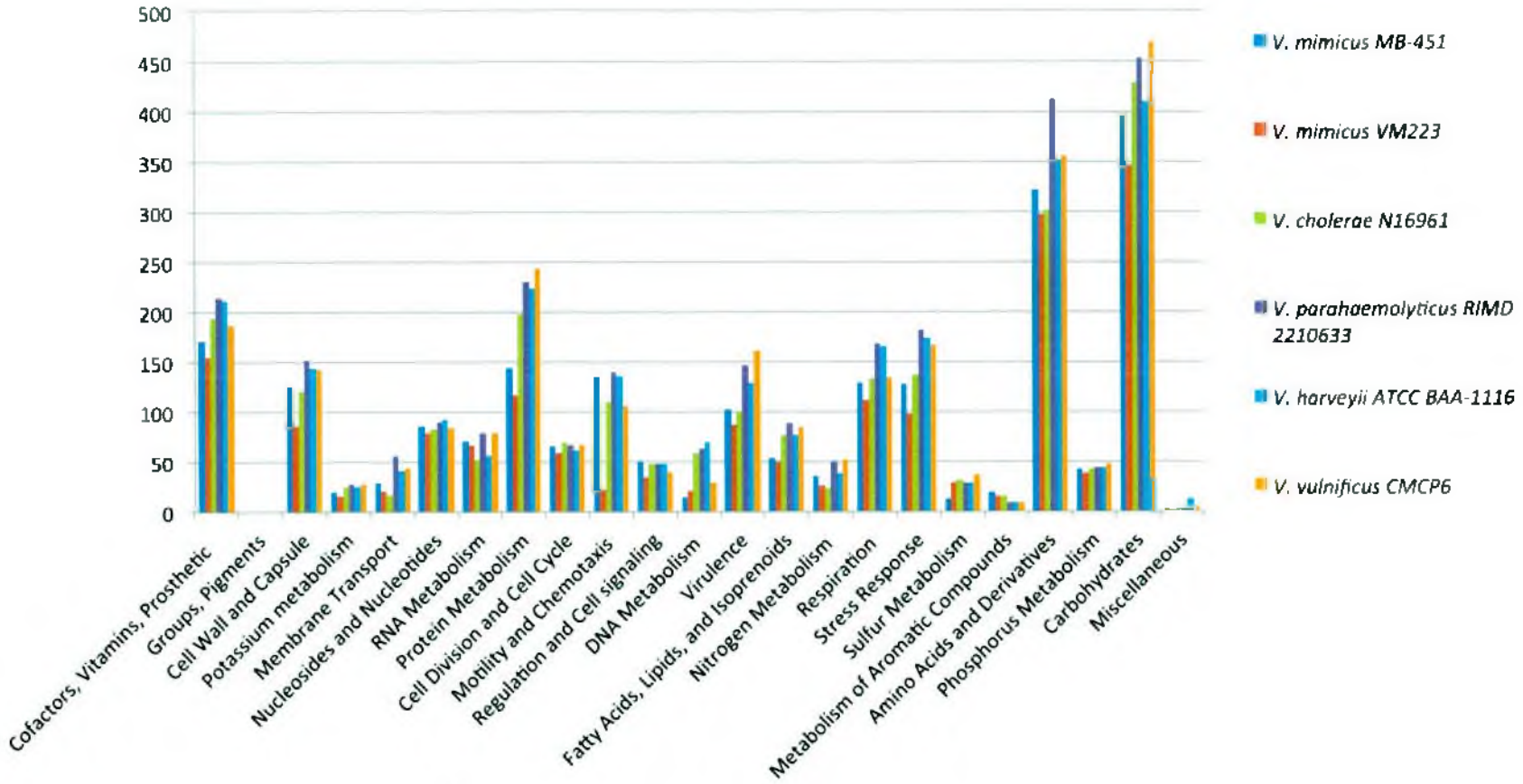


Figure 6.2: Distribution of different subsystem features in Vibrios by RAST subsystem category distribution

Dhaka University Institutional Repository

Comparative genomics. The genomes of *V. mimicus* MB451 and *V. mimicus* VM223 were compared to each other and to six other completed *Vibrio* genomes by pairwise reciprocal BLASTP analysis (Fig. 6.3), as well as by MUMmer (Fig. 6.4). Overall, they are most similar to the genomes of *V. cholerae*, but ca. 7-8% larger than *V. cholerae* N16961 (4.03Mb), due mainly to the significant size difference in C-II (~23Kb), and ca.15% smaller than both *V. parahaemolyticus* RIMD 2210633 and *V. vulnificus* CMCP6 (Table 1). A total of 1,727 non-duplicated ORFs (45.4% of MB451 and 52% of VM223 protein-coding genome) are shared by all strains, representing a "core" genome for the genus *Vibrio* and providing clear evidence of significant conservation among the genomes of *Vibrio* species. There were at least six and nine large regions on C-II and C-I, respectively, where significant mismatch was detected (Fig. 6.3), explained by insertion of different genomic islands and/or acquisition of other mobile genetic elements, resulting in generation of strain specific CDSs. *V. mimicus* MB451 and VM223 encoded 215 (5.6% of MB451 genome) and 108 (3.4% of VM223 genome) strain specific ORFs, respectively, most (83-85%) of which were annotated as hypothetical or proteins of unknown function. Ca. 1.1-1.3% of the genomes were found to be conserved, having no reciprocal match with any of the reference genomes, and therefore, are considered to be *V. mimicus* specific ORFs. Among these, 57% were annotated as hypothetical proteins, the remaining ORFs having functional assignment falls under diverse functional category, including metabolism, signal transduction, antibiotic resistance, virulence and pathogenicity, etc. Interestingly, 36% of the species-specific ORFs were encoded on C-II of *V. mimicus*.

Comparative ortholog cluster analysis revealed, *V mimicus* MB451 and VM223 share 2,428 non-duplicated ORFs (64% of MB451 genome) and 2,285 non-duplicated ORFs (69% of VM223 genome), respectively, with *V. cholerae*, of which, 2,254 ORFs are conserved among all four strains examined. Chun et al. (26) determined that the *V. cholerae* core genome contains 2,613 ORFs. Taken together, these data support a close genomic relatedness of *V. mimicus* with *V. cholerae*. *V. mimicus* MB451 and VM223 shared 66 ORFs with *V. cholerae* O1 El Tor N16961, and331 ORFs with *V. cholerae* O1 classical O395, indicating that *V. mimicus* has a closer relatedness with the classical

Dhaka University Institutional Repository

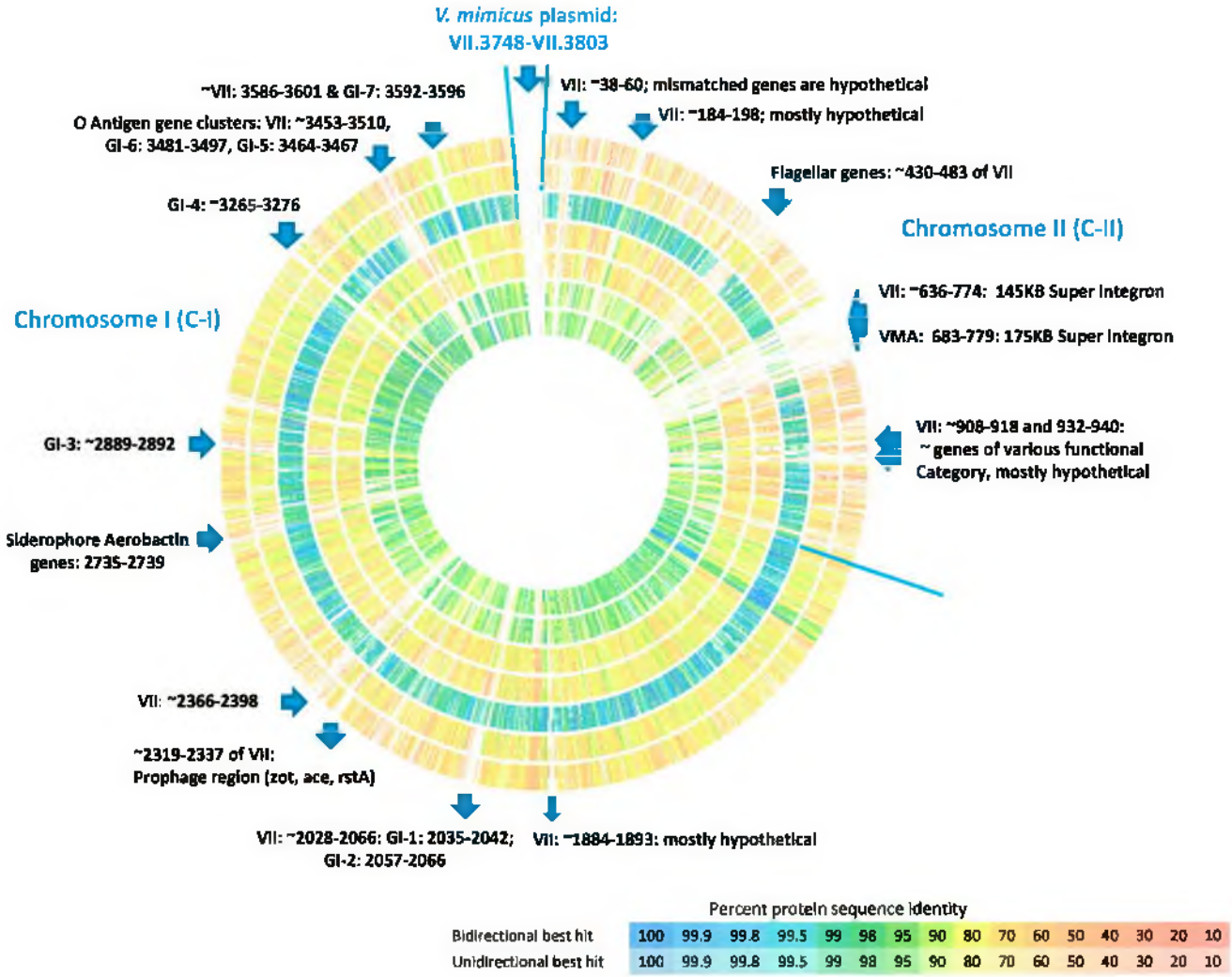


Figure 6.3: Genome comparison of *V. mimicus* MB451 with reference genomes. From outer ring to inner ring; *Vibrio fischeri* MJ11 (388396.10), Comparison Organism 2: *Vibrio harveyi* ATCC BAA-1116 (338187.6), Comparison Organism 3: *Vibrio mimicus* VM223 / VMA (666666.1800), Comparison Organism 4: *Vibrio parahaemolyticus* RIMD 2210633 (223926.3), Comparison Organism 5: *Vibrio vulnificus* CMCP6 (216895.14), Comparison Organism 6: *Vibrio cholerae* O1 biovar eltor str. N16961 (243277.1), Comparison Organism 7: *Vibrio cholerae* O395 (345073.6). Horizontal lines are drawn to separate chromosomes and plasmid. Arrow indicates area of rearrangement or insertion.

Dhaka University Institutional Repository

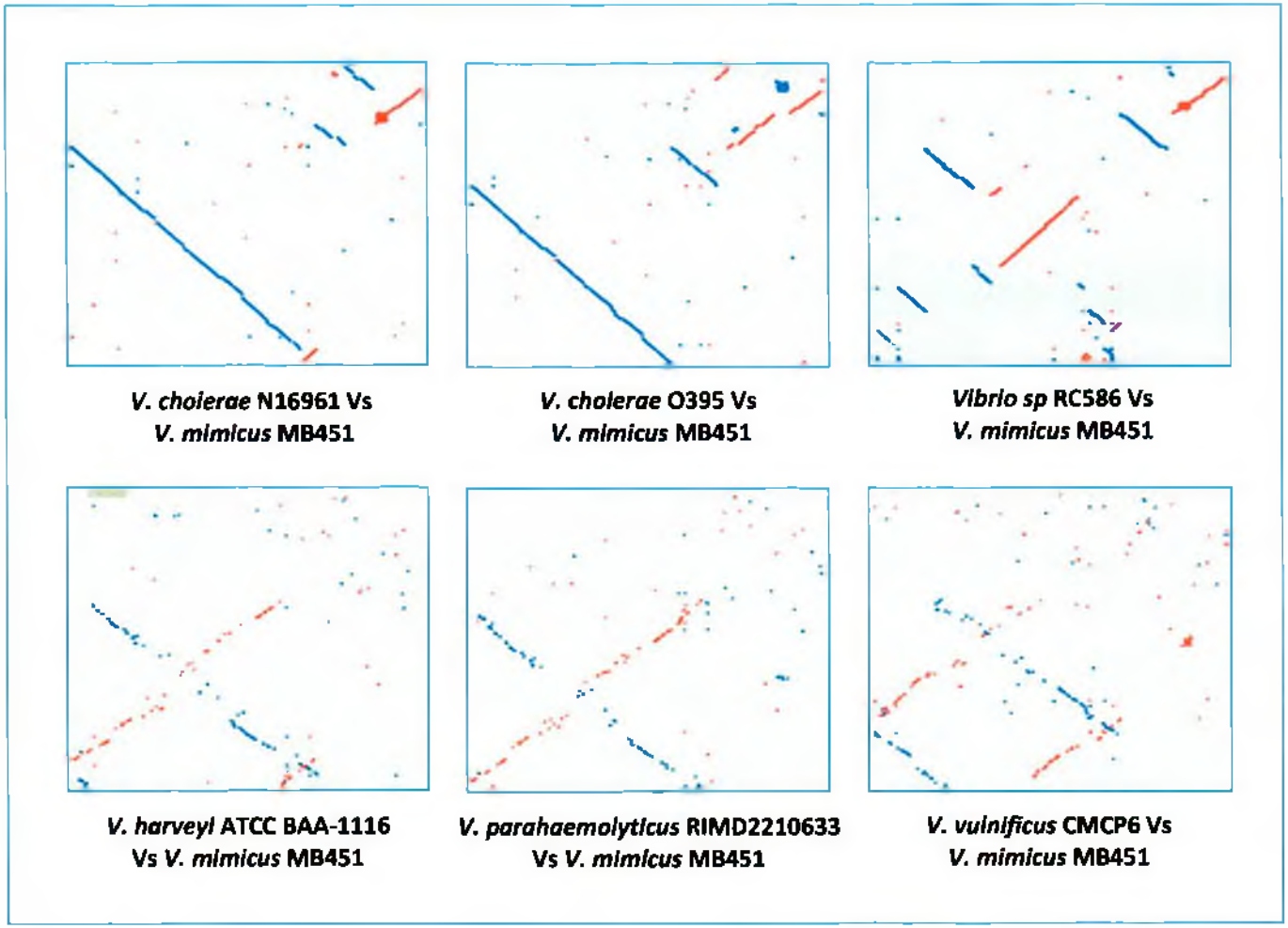


Figure 6.4: Whole genome comparison of *Vibrio mimicus* by using MUMmer program demonstrates evidence of genomic inversion and rearrangement. Blue lines indicate inversion.

Dhaka University Institutional Repository

biotype of *V. cholerae* than with El Tor, although the Average Nucleotide Identity (ANI) of conserved genes (>70% similarity) revealed no significant difference between *V. cholerae* classical (85.22%) and El Tor biotypes (85.25%) with *V. mimicus* but demonstrates clear species delineation for *V. mimicus*.

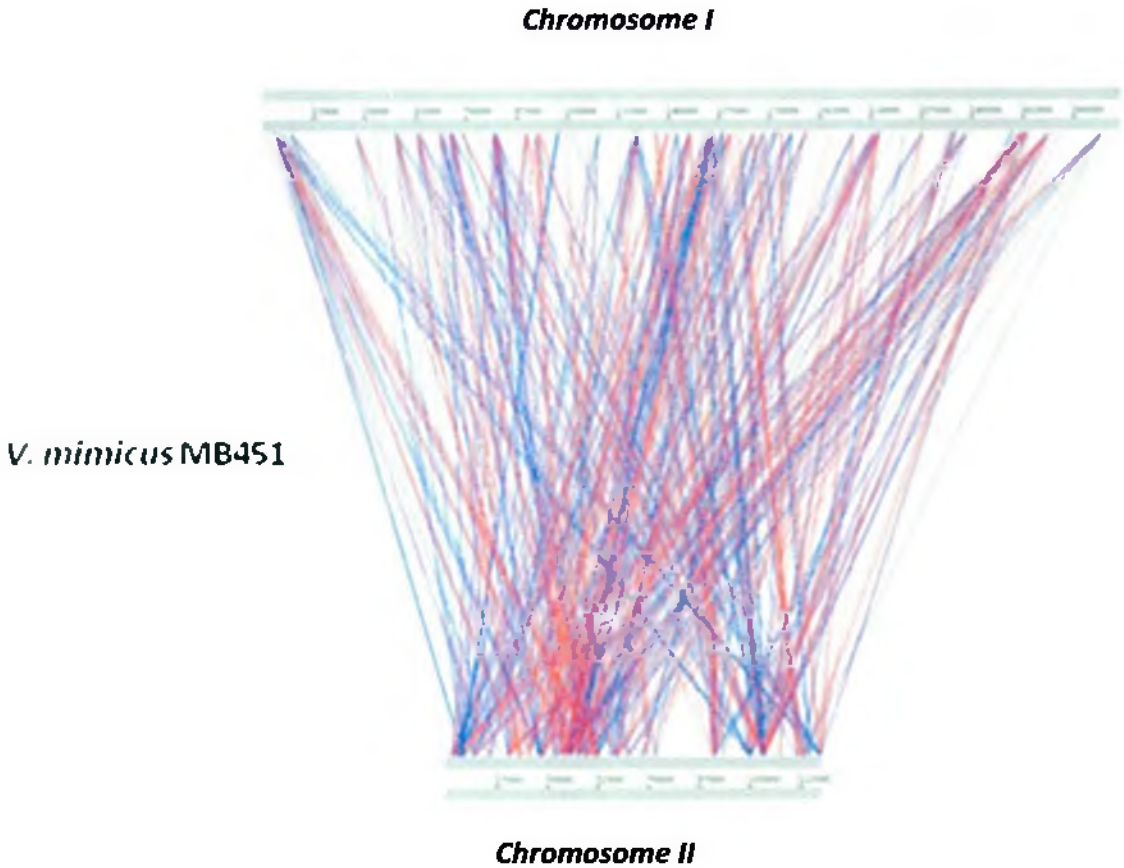
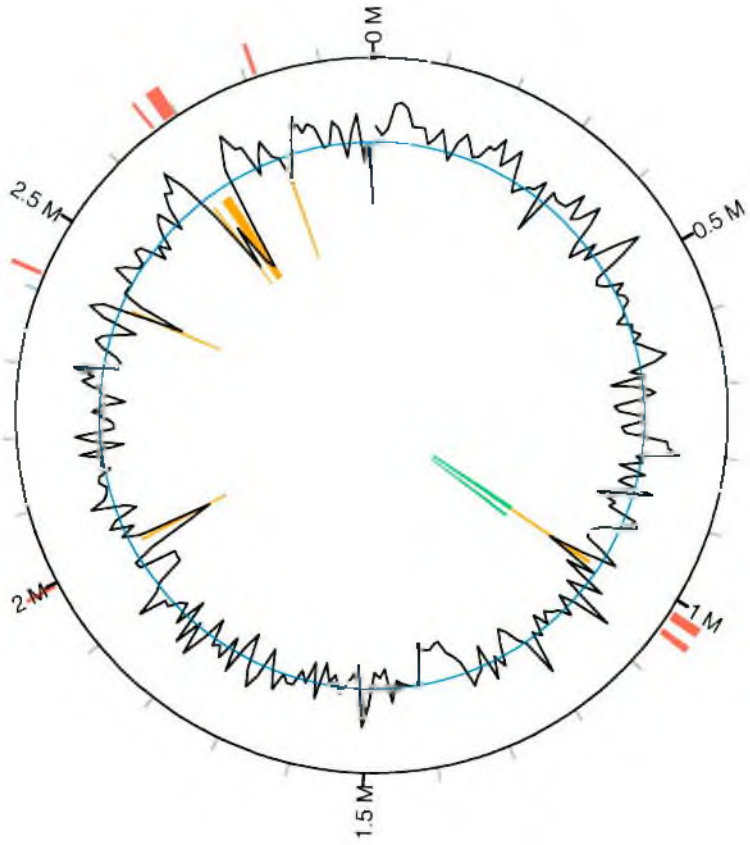


Figure 6.5: Linear pairwise comparison between chromosomes of *Vibrio mimicus:* evidence of intra chromosomal rearrangement.

The genome of clinical (MB451) and environmental (VM223) *V. mimicus* were compared and a set of 2,990 "core" ORFs was identified in the *V. mimicus* genome, possibly an underestimation because the genome of *V. mimicus* VM223 is incomplete. Nonetheless, this level of core gene content for *V mimicus* is higher than reported for either *V. cholerae* or *Escherichia coli,* where 2,613 and 2,200 core genes were found,

Dhaka University Institutional Repository



GI Prediction Program
Predicted by at least one method
IslandPick
SIGI-HMM

GI	Start*	End*	Size (bp)	GC (%)	ORFs
GI-I	1,014,173	1,026,434	12,261	41.6	8 (VII_002035-2042)
GI-II	1,039,924	1,048,870	8,946	42.9	10 (VII_002057-2066)
GI-III	1,994,173	1,999,068	4,895	38.6	4 (VII_002889-2892)
GI-IV	2,418,934	2,424,455	5,521	39.7	12 (VII_003265-3276)
GI-V	2,660,780	2,664,956	4,176	30	4 (VII_003464-3467)
GI-VI	2,680,442	2,698,767	18,325	37.3	17 (VII_003481-3497)
GI-VII	2,811,031	2,816,324	5,293	43.4	5 (VII_003592-3596)

**nucleotide position at C-I*

Figure 6.6: Distribution of genomic islands (GI's) in the large chromosome (C-I) of *V. mimicus* MB451. Color corresponds to the GI prediction method. The table lists Nucleotide position and size of predicted Genomic Islands.

Dhaka University Institutional Repository

respectively (26, 27). The core gene ratios of the large (C-I) and the small chromosome (C-II) are 84.7 and 68.4%, respectively. The greater core gene content of C-I reiterates functional importance for those genes compared with those of C-II. A total of 238 non-duplicated ORFs were identified exclusively in the genome of environmental *V. mimicus* VM223 and 68 % (155) of these ORFs are either hypothetical or of unknown function. Most of the ORFs with functional assignments are encoded on mobile genetic elements or are proteins of non-essential function and therefore most likely acquired by lateral gene transfer, and having a role in environmental fitness, such as the Rhs elements (VMA_000863-861). The clinical *V. mimicus* strain MB451 has 607 non-duplicated ORFs and 53 of these are plasmid-borne. The majority (56%) of these ORFs are hypothetical or proteins of unknown function and the ORFs with functional assignment were notably pre-CTX-Φ encoded genes, VPI-2 encoded genes, and multiple transcriptional regulators (TetR, Cro/Cl, ArsR, MarR family and others).

Genome Plasticity in *V. mimicus*

Genomic Islands. Genomic islands (GI's) are non-self-mobilizing integrative and excisive elements that encode diverse functional characteristics and their acquisition is central to bacterial evolution, serving as a mechanism of diversification and adaptation. The Island Viewer application (28) identified seven GI's in C-I of *V. mimicus* MB451, predicted by at least one method (Fig. 6.6); C-II and the plasmid have no predicted genomic island. Despite majority of GI encoded ORF's are hypothetical protein or protein of unknown functions, the ORF's with functional assignment revealed; GI's of *V. mimicus* carry important functions for metabolism, and adaptative traits that might be beneficial for the bacteria under certain growth or environmental conditions.

Repeat sequences. DNA repeats are both causes and consequences of genome plasticity, inducing deletions, amplifications, rearrangements and gene conversion. *V. mimicus* MB451 genomes contain many perfect and approximate tandem repeats, 29 and 15 in C-I and C-II, with period lengths ranging from 6 to 387 and 6 to 429, respectively (*Table 6.2S3*). As tandem repeats are generated by duplication, which

Dhaka University Institutional Repository

Indices	Period Size	Copy Number	Consensus Size	Percent Matches	Percent Indels	Score	A	C	G	T	Entropy (0-2)
***A. V. mimicus* MB451 large chromosome (C-I)**											
358024--358644	300	2.1	300	81	3	739	27	23	28	20	1.99
358610--358649	18	2.2	18	86	0	53	22	42	25	10	1.84
442362--442392	6	5.2	6	100	0	62	19	0	48	32	1.49
591746--592193	242	1.9	242	91	1	747	26	25	23	25	2
620948--620977	15	2	15	93	0	51	56	6	23	13	1.6
982227--982272	18	2.6	18	78	21	53	50	17	28	4	1.65
1196387--1202480	303	20	303	99	0	11981	24	27	35	12	1.92
1276047--1276074	9	3.1	9	100	0	56	21	10	25	42	1.85
1359866--1359895	9	3.3	9	95	0	51	26	46	20	6	1.75
1511913--1511937	13	1.9	13	100	0	50	32	0	28	40	1.57
1687083--1687127	24	2	22	82	8	54	24	15	35	24	1.94
1897165--1897953	126	6.3	126	94	0	1247	33	29	26	10	1.89
1912215--1913467	129	9.7	129	88	0	1581	22	30	20	27	1.98
1912213--1913467	387	3.2	386	88	0	1776	22	30	20	27	1.98
2007226--2007833	170	3.6	170	98	0	1173	25	26	24	22	2
2065846--2065887	22	1.9	22	85	0	57	47	11	11	28	1.76
2073236--2073286	15	3.3	15	86	2	66	54	9	17	17	1.69
2104113--2104137	12	2.1	12	100	0	50	36	40	16	8	1.77
2109348--2109379	16	2	16	100	0	64	31	31	0	37	1.58
2243396--2243464	24	2.9	24	75	0	66	5	30	15	47	1.69
2243421--2243530	48	2.3	48	82	6	125	6	34	20	39	1.78
2345165--2345812	281	2.3	281	98	0	1262	23	26	25	25	2
2438749--2439117	199	1.9	196	99	0	560	27	22	24	24	2
2450213--2450237	7	3.6	7	100	0	50	12	16	44	28	1.83
2566253--2566572	119	2.7	119	97	0	586	19	27	29	23	1.98
2566855--2567153	110	2.7	110	99	0	589	25	27	24	23	2
2599448--2599512	8	8.1	8	100	0	130	24	12	38	24	1.9
2956694--2956729	14	2.6	13	91	8	54	2	0	25	72	0.98
2958757--2958811	6	9.2	6	100	0	110	67	16	16	0	1.24
***B. V. mimicus* MB451 small chromosome (C-II)**											
7525--7572	16	3	16	90	0	78	41	12	27	18	1.86
104619--104742	6	20.7	6	100	0	248	50	33	16	0	1.46
260644--260676	15	2.1	16	88	5	50	48	18	6	27	1.71
368505--368566	15	4	15	81	12	54	40	24	8	27	1.83
368517--368560	15	2.9	15	89	0	70	40	25	6	27	1.8
382064--382098	18	1.9	18	88	0	52	14	11	34	40	1.82
473882--473909	14	2	14	100	0	56	28	7	28	35	1.84
652473--655070	432	6	432	93	1	4061	25	25	25	23	2
704461--704496	14	2.6	14	90	0	54	30	0	19	50	1.48
811213--811266	6	9	6	100	0	108	16	16	16	50	1.79
1052690--1052719	9	3.3	9	100	0	60	66	20	0	13	1.24
1138126--1138155	15	2	15	93	0	51	46	23	13	16	1.82
1179750--1179789	16	2.5	16	91	0	62	42	17	17	22	1.89
1242212--1242249	20	1.9	20	88	0	58	31	28	7	31	1.86
1267165--1267206	18	2.3	18	87	0	57	26	45	19	9	1.8

Table 6.2: Tandem repeats in *V. mimicus* MB451 identified using the Tandem repeats finder program.

Dhaka University Institutional Repository

genome structures, this may be a key process in the evolution of *V. mimicus*. Interestingly, many (59%) of the tandem repeats are in protein encoding genes, which may exhibit higher mutation rates, allowing targeted sequence variation and thereby, enabling a rapid response to challenging and hostile conditions of both the external environment and the human intestine.

Genomic rearrangement. Our analysis of the *V. mimicus* MB451 genome, using MUMmer (29) and the Artemis Comparison Tool (ACT) (30), inferred that evolution of the *V. mimicus* MB451 genome structure is marked by several intra (Fig. 6.5) and inter-chromosomal (Fig. 6.3, 6.4 and 6.7) rearrangements. It is apparent from the overall genomic comparison (*Supp. Fig. S2*) that *V. mimicus* MB451 is more closely related to *V. cholerae* and *Vibrio* sp. RC586, than to any of the other *Vibrio spp.* despite a number of repeated inversions occurred across the genome. These rearrangements appear to be involved in generation of strain-specific CDSs and suggest that extensive genome plasticity is common among species of the genus *Vibrio*. Chromosome-wise comparison of *V. mimicus* with El Tor and Classical biotypes of *V. cholerae* demonstrate that both chromosomes have greater synteny with the classical biotype of *V. cholerae* than with El Tor and the overall gene content and position in C-I is much better conserved than in C-II (Fig. 6.7). Extensive genomic rearrangements (inversions, insertions and deletions) obviously have occurred in C-II. These findings lead us to an important conclusion, namely that C-II has a critical function in the evolution and genesis of speciation in the genus *Vibrio*. Moreover, the high degree of genome synteny, as well as the larger number of conserved ORFs shared with pre-7th pandemic strains of *V. cholerae*, lead to the conclusion that, in the course of evolution, *V. mimicus* and *V. cholerae* most likely diverged from a common ancestor having prototypic 6th pandemic clones of *V. cholerae* as their "core". Once separated, the *V. mimicus* genome underwent extensive rearrangement, especially in C-II.

Pathogenicity. Although virulence determinants of *V. mimicus* have not yet been well characterized, the ability of *V. mimicus* to produce multiple toxins has been reported (7-10, 31). Further study also demonstrated the presence of CTX prophage and *Vibrio*

Dhaka University Institutional Repository

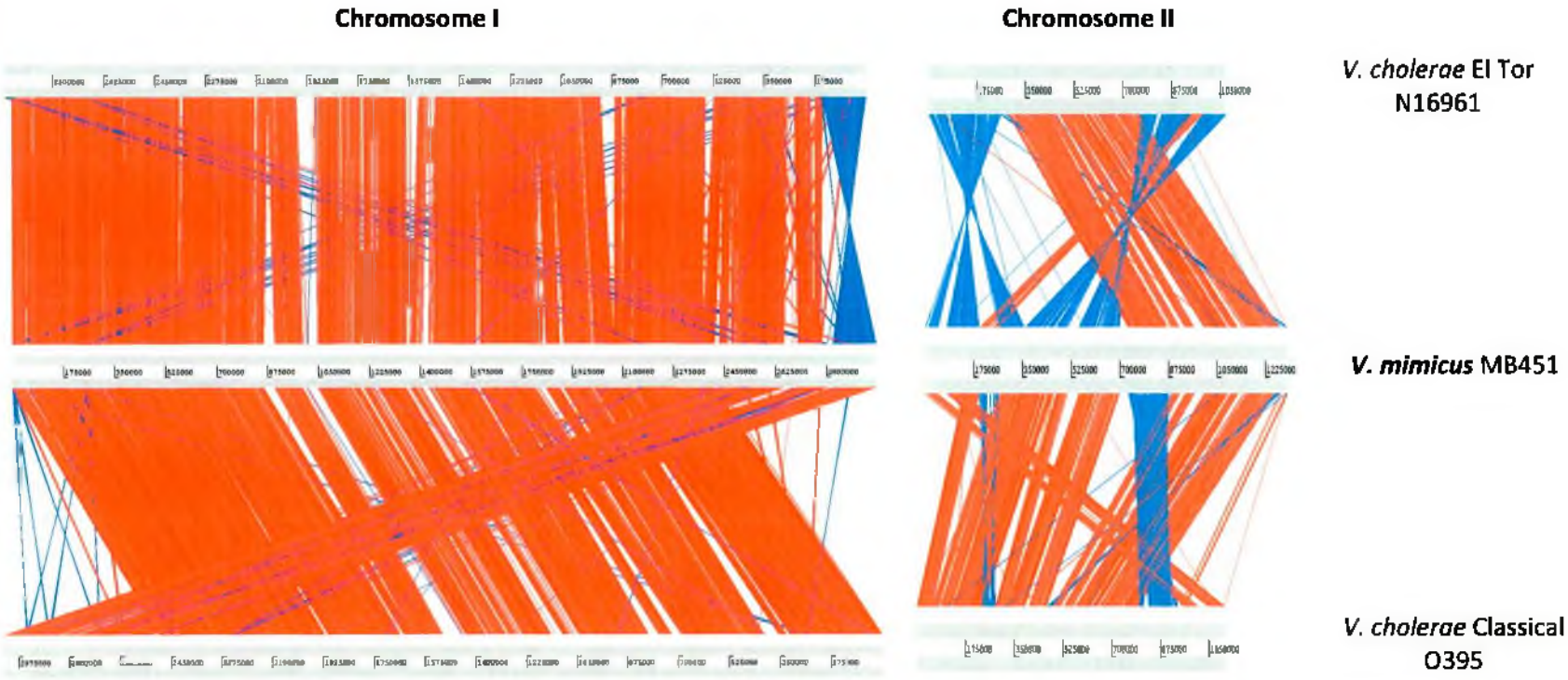


Figure 6.7: Linear pairwise comparison of the *Vibrio mimicus* genome. The analysis was performed using Artemis Comparison Toll (ACT). Regions with similarity are highlighted by connecting red or blue lines between the genomes; red lines indicate homologous blocks of sequence and blue lines indicate inversions, whereas gaps are indicative of unique DNA. The grey bars represent forward and reverse strands.

Dhaka University Institutional Repository

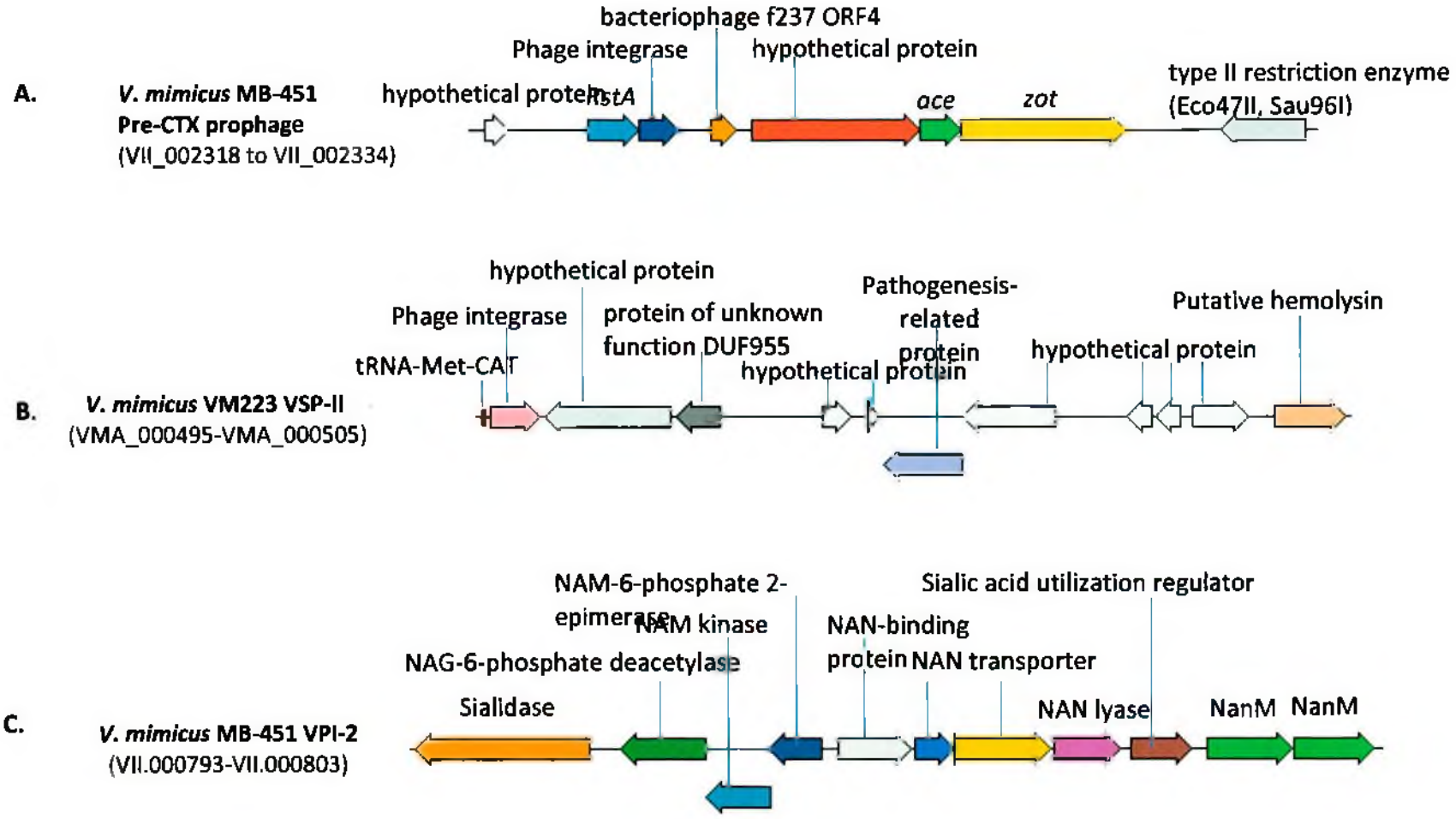


Figure 6.8: Pathogenicity Islands found in the genome of *Vibrio mimicus*

Dhaka University Institutional Repository

Pathogenecity Island (VPI) in *V. mimicus* (2), virulence factors found in toxigenic *V. cholerae*. The genomes of MB451 and VM223 were found to possess virulence determinants previously known to occur in *V. mimicus* as well as a number of elements, such as the *Vibrio* Seventh Pandemic Island II (VSP-II), mannose-sensitive hemagglutinin (MSHA), enterotoxigenic hemolysin (HlyA), type IV pilin (PilE), and Integron Integrase (*Int*I4), described here for the first time in *V. mimicus*.

CTXΦ Insertion Site. Inserted at the CTXΦ insertion locus on C-I of *V. mimicus* VM223 is a ca. 7.5 kb element encoding four ORFs (VMA_002135-38), three of which are annotated as hypothetical proteins and one as a phage protein. Inserted at this same locus in *V. mimicus* MB451 is a ca. 14 kb element encoding 16 ORFs (VII_002318 to VII_002334), eight of which are annotated as hypothetical proteins. Interestingly, this element constitutes a pre-CTX prophage (Fig.6.8A) evidenced by the presence of ORFs homologous to the zona occludens toxin (Zot), accessory cholera enterotoxin (Ace), RstA phage-related replication protein and RstB phage related integrase, found in strains of *V. cholerae* harboring pre-CTX phages. This phage-like element also encodes an ORF homologous to ORF9 of the filamentous bacteriophage f237, suggesting that the insertion locus for CTXΦ is conserved in *V. mimicus* and serves as an integration locus for various phages. The sequence of the putative prophage was compared and found to be similar to that of environmental non-toxigenic *V. cholerae*, i.e., 65-70% of the *V. mimicus* MB451 pre-CTX prophage had 80-83% sequence similarity to pre-CTX prophages of environmental *V. cholerae* non-O1/O139 strains AM-19226, 1587, MZO-3, VL426, 12129(1), and MZO-2. Moreover, 65% of the entire prophage has 82% identity with *V. cholerae* KSF-1phi, whereas similarity with other *V. cholerae* phages (such as VGJphi, fs1, VEJphi, VSKK, VSK etc.) was not significant.

***Vibrio* Pathogenicity Island 2 (VPI-2).** VPI-2 is a 57.3 kb gene cluster that encodes genes for neuraminidase and sialic acid metabolism, and has characteristic features of a pathogenicity island. The genome of *V. mimicus* MB451 encodes a ca. 12.8 kb portion of VPI-2 (Fig. 6.8C) that contains genes responsible for sialic acid metabolism (homologous to VC1773 to VC1784 of *V. cholerae* N16961), surprisingly located on the

Dhaka University Institutional Repository

C-II (VII_000793-VII_000803). It has been suggested that this sialic acid metabolism region allows the bacterium to metabolize sialic acid residues on the gangliosides of human intestinal epithelial cells, thereby providing a nutrient source for the cell (32). Interestingly, in this segment of VPI-2, regions containing homologues of genes VC1758 (phage integrase) to VC1782, encoding the type I restriction modification system, and VC1785 to VC1810, encoding the mu-phage, are deleted.

***Vibrio* Seventh Pandemic Island II.** The genome of *V. mimicus* VM223 harbors a variant of *Vibrio* Seventh Pandemic Island II (VSP-II), located on C-I (Fig. 6.8B). The 21.5 kb island encompasses 11 ORFs (VMA_000495-VMA_000505), which include a phage integrase, protein of unknown function DUF955, pathogenesis-related protein, putative hemolysin and several (63%) hypothetical proteins, and has a G+C content of 40%, different from the average chromosomal G+C content of 46%. Further scrutiny of this site revealed eight additional ORFs, including a DNA repair protein, RadC, a transcriptional regulator, Ribonuclease HI, and five hypothetical proteins, comprising a total of 19 ORFs encoded by *V. mimicus* VM223 VSP-II. In 7th pandemic isolates of *V. cholerae* O1 biotype El Tor and O139, VSP-II is a 26.9 kb region (*V. cholerae* O1 El Tor N16961 genes VC0490–VC0516) and 41 % of *V. mimicus* VM223 VSP-II is identical to *V. cholerae* VSP-II. However, there are noticeable variations in the organization and content of ORFs in *V. mimicus* VSP-II compared with *V. cholerae* VSP-II.

Hemolysins. Both *V. mimicus* genomes contained one copy of heat labile (VII_000023 & VMA_001263), heat stable (VII_000310 & VMA_001040), and enterotoxigenic (HlyA, VII_000877 & VMA_000630) hemolysin, a major virulence factor of *V. mimicus* (5). The heat labile hemolysins (VMH) share 97% nucleotide similarity between each other and 77% with *V. cholerae* El Tor hemolysin (HlyA). The *V. mimicus* heat stable hemolysins are 98% identical to each other and 76% with the thermostable hemolysin of *V. cholerae,* 70% to the *Listonella anguillarum vah*4 gene for hemolysin, and 67% to the thermostable hemolysin of *V. vulnificus* YJ016. Interestingly, no significant similarity was observed with *V. parahaemolyticus tdh.* The enterotoxigenic hemolysin (HlyA) of *V. mimicus* is 82% similar to *V. cholerae* HlyA

Dhaka University Institutional Repository

(VCA0219). Both *V. mimicus* genomes contained 12 other hemolysins, five of which are located on C-II. This finding is in agreement with the presence of multiple hemolysins in the genomes of pathogenic and non-pathogenic *Vibrio* spp., including *V. cholerae* (26). Smith and Oliver (33) suggested these hemolysins play a role in the cold-shock response of *V. vulnificus* and this certainly can be extended to *V. mimicus*.

Proteases. *Vibrio mimicus* protease (VMP or Vm-HA/protease) has recently been shown to modulate the activity of *V. mimicus* hemolysin (VMH) by limited proteolysis (52). Since enhanced hemagglutinability surrogates strong bacterial cellular affinity for binding to host mucosa, VMP can be arrived as a significant component of *V. mimicus* virulence. VMP (VII_000553 and VMA_000854) of the *V. mimicus* genome showed 97% nucleotide similarity with each other and with the VMP of *V. mimicus* ES-39. VMP of *V. mimicus* VM223 and MB451 also showed significant similarity at both the nucleotide (80%) and amino acid level (88%) to HA/protease of *V. cholerae*. Both genomes of *V. mimicus* also contained homologs (VII_000021; VMA_001265) of the prtV protease of *V. cholerae* on C-II, which has recently been shown to be a virulence factor in a *Caenorhabditis elegans* model (34). In addition, both genomes possessed a number of other homologs of proteases, including a membrane-bound zinc metalloprotease (VII_001591; VMA_002784), Protease IV (VII_001833; VMA_002551; and htpX protease (VII_002671; VMA_001839). Interestingly, the genome of *V. mimicus* MB451 encodes a second putative zinc metalloprotease (VII_001021), which is 97% identical (amino acid) to the zinc metalloproteases of *V. mimicus* VM573 and VM603 and 77% identical to that of *V. ulnificus* CMCP6 and YJ016. All five homologs share amino acid homology to bacillolysin neutral protease of *Bacillus* spp.

Pili and adherence. Attachment is a critical step in pathogenesis and both pili and flagella are important virulence factors in many pathogenic bacteria (35-38). Mannose-sensitive hemagglutinin (MSHA) pilus (39), a type IV pilus (TFP) has been shown to play a role in adherence to plankton and in biofilm formation (40-42); but its role in intestinal colonization is not clear. Both *V. mimicus* genomes possess putative MSHA gene clusters (VII_003323-3339; VMA_000581-591), which are of nearly equal

Dhaka University Institutional Repository

length and highly similar. Both are slightly smaller in size, compared to that of *V. cholerae*, with the majority of the deleted sequence contained in the coding region of MshN. BlastN analysis of *V. mimicus* MSHA demonstrated that these clusters are 77% identical to that of *V. cholerae* (VCO398-411) and their G+C content differs slightly from *V. cholerae*. Additionally, there are no apparent integrase or transposases defining this region as a pathogenicity island or suggesting another origin. Taken together, these results suggest that the MSHA gene cluster was acquired before *V. cholerae* and *V. mimicus* diverged from their most recent common ancestor. Since hemagglutination plays an important role in intestinal adherence of *V. mimicus* (43), the MSHA of *Vibrio mimicus* may serve as a dual function appendage; i.e., intestinal adherence and attachment to plankton or other surfaces to form a biofilm. In addition, MSHA has been shown to serve as receptor for KSF-1Φ, VGJΦ, fs-1, fs-2, and 493 filamentous phages in *V. cholerae* (44-48). Therefore, the same family of phages may be expected to infect *V. mimicus*.

The *V. mimicus* genomes also harbored additional TFP genes. Both *V. mimicus* genomes contained a homolog of the type IV pilin gene, *pil*A (VII_001442; VMA_002922), which encodes subunits of fimbriae and often expressed during human infection, and two copies of the *tad* locus, with one copy on each chromosome (C-I: VMA_002363-70 and VII_002046-56; C-II: VMA_001394-1402 and VII_000975-85). A homolog of the TFP biogenesis protein PilE, a marker often employed to assess presence of a TFP associated with virulence in *V. cholerae*, was also identified in the intergenic sequence between the type IV fimbrial biogenesis protein FimT (VII_002948) and Chaperone protein DnaJ (VII_002949) in *V. mimicus* VM223 with 85% sequence similarity to *V. cholerae pilE* (VC0857).

Dhaka University Institutional Repository

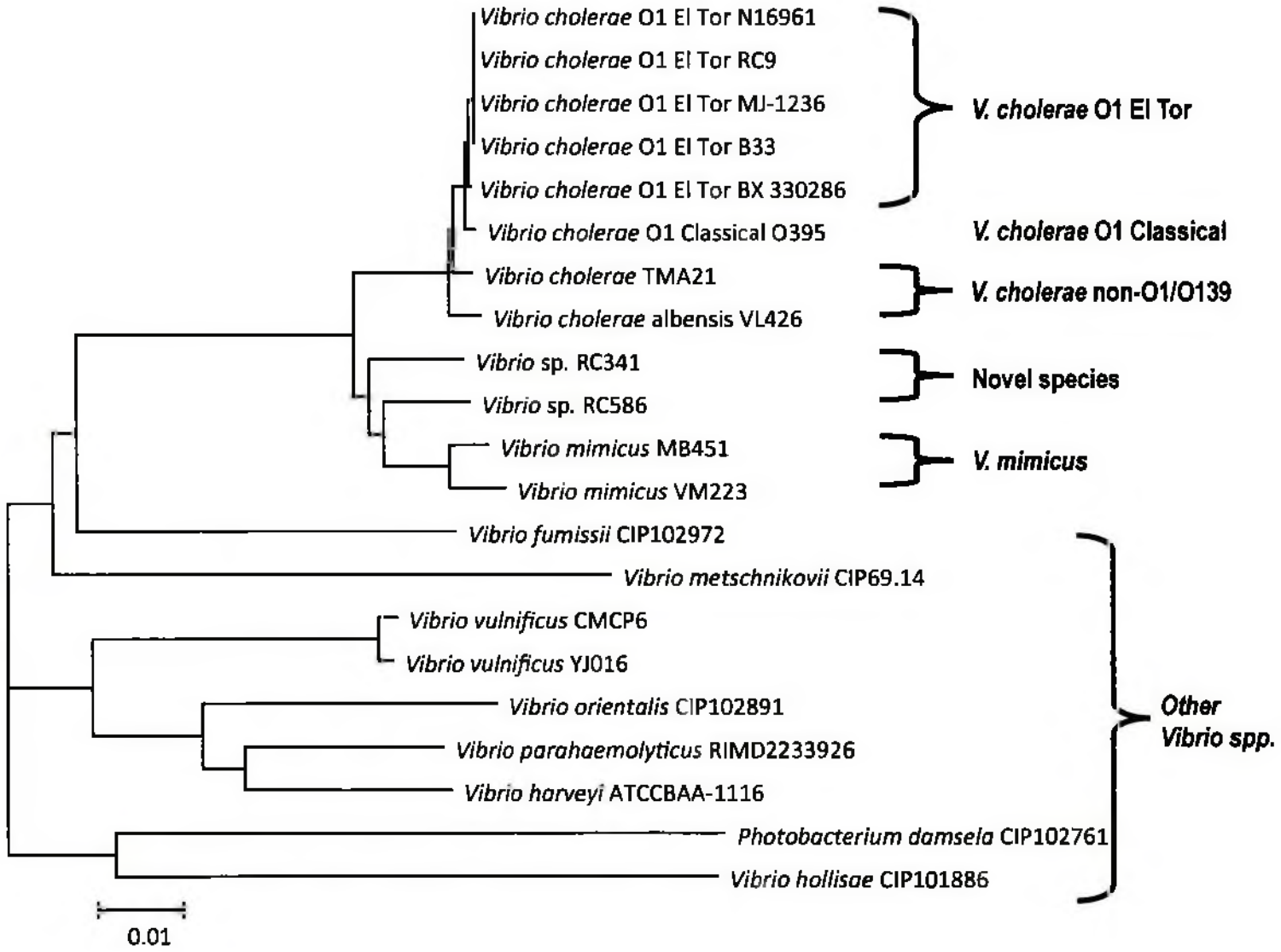


Figure 6.9. Neighbor-joining tree based on alignment of homologous sequences of 75 conserved ORFs. Bar represents 0.01 substitutions per site.

Dhaka University Institutional Repository

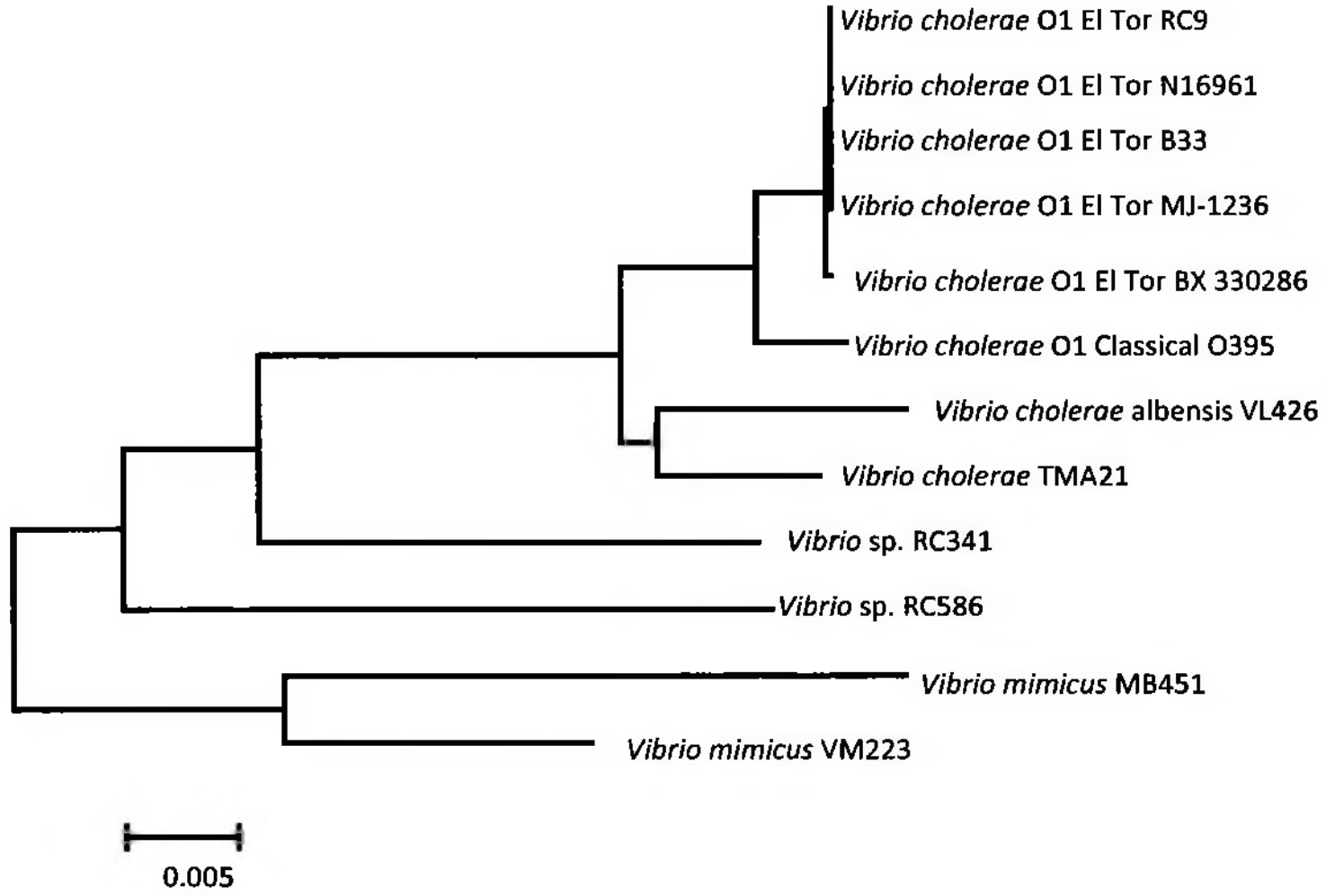


Figure 6.10. Neighbor-joining tree based on alignment of homologous sequences of 925 conserved ORFs. Bar represents 0.005 substitutions per site.

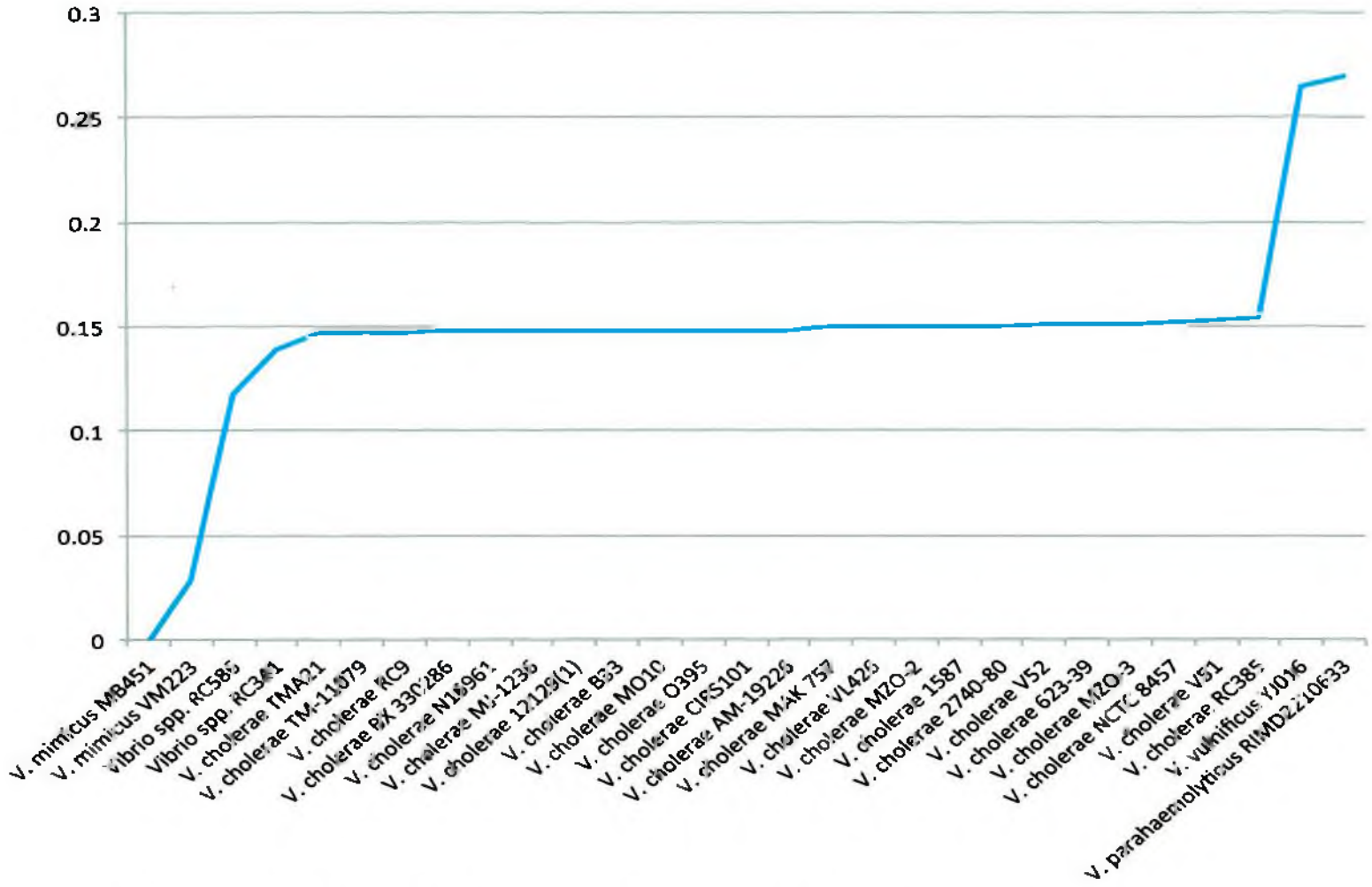


Figure 6.11. Evolutionary Distance Tree based on the divergence of the conserved ORFs among *Vibrio* spp.

Dhaka University Institutional Repository

Dhaka University Institutional Repository

Phylogenomics of *V. mimicus*. The phylogeny of *V. mimicus* is inferred from a neighbor-joining tree, using homologous alignment of 75 conserved ORFs of different *Vibrio* species (Fig. 6.9). *V. cholerae* strains tightly clustered together and the two *V. mimicus* strains branched separately from *V. cholerae*, clustering with the two novel *Vibrio* species (51), *Vibrio* spp. RC586 and RC341. Therefore, *V. cholerae* and *V. mimicus* may have evolved recently from a common ancestor and more likely are distinguished from each other by presence or absence of mobile genetic elements, based on the high rate of insertion and deletion in the *V. cholerae* genome (26). A second genome-based neighbor-joining tree (Fig. 6.10), constructed using 925 ORFs, highly conserved among *V. cholerae* and *V. mimicus* strains, demonstrates that *V. mimicus* are deeply rooted as independent branches at ancestral nodes, indicating evolutionary divergence from a progenitor of an ancestral *V. cholerae*. The evolutionary relationships inferred by this tree suggest that *V. mimicus* is more closely related to the *Vibrio* spp. RC586 and RC341, than to *V. cholerae*, consistent with results of the previous analysis (Fig. 6.10). A closer phylogenetic relationship of *V. mimicus* with *Vibrio* spp. RC586 and RC341 is also supported by the higher ANI value between them. An evolutionary distance tree (Fig. 6.11), based on divergence of nucleotide sequences in different *Vibrio* genomes, yielded results consistent with those shown in Fig. 6.9 and Fig. 6.10, and allows two important conclusions. First, the *V. cholerae-V. mimicus* clade forms a monophyletic group constituting a sister group of *V. mimicus* and two new *Vibrio* spp. RC586 and RC341, with *V. cholerae* as an outgroup. Second, the *V. cholerae-V. mimicus* clade diverged from a common ancestor before the two lineages diverged from each other. The first conclusion is strongly supported by ANI of the conserved ORFs, 84.5 to 85.5% with *V. cholerae* and 86 to 88% with RC341 and RC586 and the second by the evolutionary distance tree.

Dhaka University Institutional Repository

Conclusion:

Results of the genome sequence analysis reported here for *V. mimicus* provide insight into the biology of this organism, notably, in the refinement and expansion of our knowledge on metabolism, virulence, evolution, and phylogeny of this organism. These new *Vibrio mimicus* genome sequence can now serve as reference for studies of communities where *Vibrio* spp. are present. Even though this organism was selected for genome sequencing prominently because of the controversy concerning its phylogenetic novelty, our genome sequence-based phylogeny reveals *V. mimicus* have most likely originated from a common ancestor of *V. cholerae*, with subsequent divergence of both species through extensive rearrangement, especially in C-II. The presence of *V. cholerae* virulence regions previously reported (VPI and CTXΦ) in *V. mimicus,* as well as those discovered in this study (VSP-II, MSHA, *hlyA*, *pilE*, and *Intl4*), and their greater similarity of these newly discovered virulence regions to that of *V. cholerae,* indicate recent interspecies lateral transfer between *V. cholerae* and *V. mimicus* and suggests transfer of virulence factors among isolates is an ongoing process. The higher genomic relatedness of *V. mimicus* to the 6th pandemic biotype of *V. cholerae,* as well as acquisition of the novel features, VSP-II, MSHA, HlyA, and Intl4 that are characteristic of the 7th pandemic biotype of *V. cholerae*, provide insight into how genomic rearrangement can enhance virulence and environmental fitness. Finally, the genome of *V. mimicus* provides a starting point for understanding how a free-living, environmental bacterium can emerge as a human pathogen. It is anticipated that many of the genes identified as virulence genes will prove to be important for the fitness of the bacterium in its native aquatic environments as evidenced by the presence of these genes in both the clinical and environmental as well as conservation of homologs in environmental *V. cholerae* genomes.

Dhaka University Institutional Repository

References:

1. Davis BR, *et al.* (1981) Characterization of biochemically atypical *Vibrio cholerae* strains and designation of a new pathogenic species, *Vibrio mimicus*. *J Clin Microbiol* 14(6):631-639.
2. Boyd EF, Moyer KE, Shi L, & Waldor MK (2000) Infectious CTXPhi 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.
3. Byun R, Elbourne LD, Lan R, & Reeves PR (1999) Evolutionary relationships of pathogenic clones of *Vibrio cholerae* by sequence analysis of four housekeeping genes. *Infect. Immun.* 67(3):1116-1124.
4. Chowdhury MA, Yamanaka H, Miyoshi S, Aziz KM, & Shinoda S (1989) Ecology of *Vibrio mimicus* in aquatic environments. *Appl Environ Microbiol* 55(8):2073-2078.
5. Chowdhury MA, Aziz KM, Kay BA, & Rahim Z (1987) Toxin production by *Vibrio mimicus* strains isolated from human and environmental sources in Bangladesh. *J. Clin. Microbiol.* 25(11):2200-2203.
6. Colwell RR & Huq A (1994) Environmental reservoir of *Vibrio cholerae*. The causative agent of cholera. *Disease in Evolution: Global Changes and Emergence of Infectious Diseases*, eds Wilson ME, Levins R, & Spielman A (Ann. New York Acad. Sci., New York), Vol 740, pp 44-54.
7. Campos E, *et al.* (1996) *Vibrio mimicus* diarrhea following ingestion of raw turtle eggs. *Appl. Environ. Microbiol.* 62(4):1141-1144.
8. Chowdhury MA, Hill RT, & Colwell RR (1994) A gene for the enterotoxin zonula occludens toxin is present in *Vibrio mimicus* and *Vibrio cholerae* O139. *FEMS Microbiol Lett* 119(3):377-380.
9. Shi L, *et al.* (1998) Detection of genes encoding cholera toxin (CT), zonula occiudens toxin (ZOT), accessory cholera enterotoxin (ACE) and heat-stable enterotoxin (ST) in *Vibrio mimicus* clinical strains. *Microbiol Immunol* 42(12):823-828.
10. Spira WM & Fedorka-Cray PJ (1984) Purification of enterotoxins from *Vibrio mimicus* that appear to be identical to cholera toxin. *Infect Immun* 45:679-684.
11. Nishibuchi M & Seidler RJ (1983) Medium-dependent production of extracellular enterotoxins by non-O-1 Vibrio cholerae, Vibrio mimicus, and Vibrio fluvialis. *Appl Environ Microbiol* 45(1):228-231.
12. Gyobu Y, Kodama H, & Uetake H (1988) Production and partial purification of a fluid-accumulating factor of non-O1 Vibrio cholerae. *Microbiol Immunol* 32(6):565-577.
13. Nishibuchi M, Khaeomanee-iam V, Honda T, Kaper JB, & Miwatani T (1990) Comparative analysis of the hemolysin genes of Vibrio cholerae non-01, V. mimicus, and V. hoilisae that are similar to the tdh gene of V. parahaemolyticus. *FEMS Microbiol Lett* 55(3):251-256.

14. Chowdhury MA, Miyoshi S, & Shinoda S (1990) Purification and characterization of a protease produced by Vibrio mimicus. *Infect Immun* 58(12):4159-4162.

15. Kang JH, Lee JH, Park JH, Huh SH, & Kong IS (1998) Cloning and identification of a phospholipase gene from vibrio mimicus. *Biochim Biophys Acta* 1394(1):85-89.

16. Shaw JF, *et al.* (1994) Nucleotide sequence of a novel arylesterase gene from Vibro mimicus and characterization of the enzyme expressed in Escherichia coli. *Biochem J* 298 Pt 3:675-680.

17. Okujo N & Yamamoto S (1994) Identification of the siderophores from Vibrio hollisae and Vibrio mimicus as aerobactin. *FEMS Microbiol Lett* 118(1-2):187-192.

18. Alam M, Miyoshi S, Maruo I, Ogawa C, & Shinoda S (1994) Existence of a novel hemagglutinin having no protease activity in Vibrio mimicus. *Microbiol immunol* 38(6):467-470.

19. Mazel D, Dychinco B, Webb VA, & Davies J (1998) A distinctive class of intergron in the *Vibrio cholerae* genome. *Science* 280:605-608.

20. Heidelberg JF, *et al.* (2000) DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* 406:477-483.

21. Rowe-Magnus DA, Guerout AM, & Mazel D (1999) Super-integrons. *Res Microbiol* 150(9-10):641-651.

22. Rowe-Magnus DA, *et al.* (2001) The evolutionary history of chromosomal super-integrons provides an ancestry for multiresistant integrons. *Proc Natl Acad Sci U S A* 98(2):652-657.

23. Dyda F, Klein DC, & Hickman AB (2000) GCN5-R ELATED NA CETYLTRANSFERASES: A Structural Overview. *Annu. Rev. Biophys. Biomol. Struct.* 29(1):81-103.

24. Chain PSG, *et al.* (2006) Complete genome sequence of *Yersinia pestis* strains Antiqua and Nepal516: evidence of gene reduction in an emerging pathogen. *J. Bacteriol.* 188(12):4453.

25. Johnson DC, Dean DR, Smith AD, & Johnson MK (2005) Structure, function, and formation of biological iron-sulfur clusters. *Annu Rev Biochem* 74:247-271.

26. Chun J, *et al.* (2009) Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. *Proc. Natl. Acad. Sci. U.S.A.* 106(36):15442.

27. Rasko DA, *et al.* (2008) The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *J. Bacteriol.* 190(20):6881.

28. Langille MGI & Brinkman FSL (2009) IslandViewer: an Integrated interface for computational identification and visualization of genomic islands. *Bioinformatics* 25(5):664.

29. Kurtz S, *et al.* (2004) Versatile and open software for comparing large genomes. *Genome Biol* 5(2):R12.

30. Carver TJ, *et al.* (2005) ACT: the Artemis comparison tool. *Bioinformatics* 21(16):3422.

31. Ramamurthy T, *et al.* (1994) *Vibrio mimicus* with multiple toxin types isolated from human and environmental sources. *J Med Microbiol* 40(3):194-196.

32. Jermyn WS & Boyd EF (2002) Characterization of a novel Vibrio pathogenicity island (VPI-2) encoding neuraminidase (nanH) among toxigenic *Vibrio cholerae* isolates. *Microbiol* 148(Pt 11):3681-3693.

33. Smith B & Oliver JD (2006) In situ and in vitro gene expression by *Vibrio vulnificus* during entry into, persistence within, and resuscitation from the viable but nonculturable state. *Appl Environ Microbiol* 72(2):1445-1451.

34. Vaitkeviclus K, *et al.* (2006) A *Vibrio cholerae* protease needed for killing of *Caenorhabditis elegans* has a role in protection from natural predator grazing. *Proc Natl Acad Sci U S A* 103(24):9280-9285.

35. Sauer FG, Mulvey MA, Schilling JD, Martinez JJ, & Hultgren SJ (2000) Bacterial pili: molecular mechanisms of pathogenesis. *Curr Opin Microbiol* 3(1):65-72.

36. O'Toole GA & Kolter R (1998) Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. *Mol. Microbiol.* 30(2):295-304.

37. Morand PC, Tattevin P, Eugene E, Beretti JL, & Nassif X (2001) The adhesive property of the type iV pilus-associated component PilC 1 of pathogenic *Neisseria* is supported by the conformational structure of the N-terminal part of the molecule. *Mol. Microbiol.* 40(4):846-856.

38. Zolfaghar I, Evans DJ, & Fleiszig SMJ (2003) Twitching motility contributes to the role of pili in corneal infection caused by *Pseudomonas aeruginosa*. *Infect. Immun.* 71(9):5389.

39. Marsh JW & Taylor RK (1999) Genetic and transcriptional analyses of the *Vibrio cholerae* mannose-sensitive hemagglutinin type 4 pilus gene locus. *J. Bacteriol.* 181(4):1110.

40. Chiavelli D, Marsh J, & Taylor R (2001) The mannose-sensitive hemagglutin of *Vibrio cholerae* promotes adherence to zooplankton. *Appl Environ Microbiol* 67(7):3220-3225.

41. Watnick Pi, Fullner KJ, & Kolter R (1999) A role for the mannose-sensitive hemagglutinin in biofilm formation by *Vibrio cholerae* El Tor. *J Bacteriol* 181(11):3606-3609.

42. Watnick PI & Kolter R (1999) Steps in the development of a *Vibrio cholerae* El Tor biofilm. *Mol Microbiol* 34(3):586-595.

43. Alam M, Miyoshi S, Tomochika K, & Shinoda S (1997) *Vibrio mimicus* attaches to the intestinal mucosa by outer membrane hemagglutinins specific to polypeptide moieties of glycoproteins. *Infect Immun* 65(9):3662-3665.

44. Faruque SM, *et al.* (2005) Genomic sequence and receptor for the Vibrio cholerae phage KSF-1phi: evolutionary divergence among filamentous vibriophages mediating lateral gene transfer. *J Bacteriol* 187(12):4095-4103.

45. Campos J, *et al.* (2003) VGJ phi, a novel filamentous phage of *Vibrio cholerae*, integrates into the same chromosomal site as CTX phi. *J Bacteriol* 185(19):5685-5696.

Dhaka University Institutional Repository

46. Honma Y, Ikema M, Toma C, Ehara M, & Iwanaga M (1997) Molecular analysis of a filamentous phage (fsl) of *Vibrio cholerae* O139. *Biochim Biophys Acta* 1362(2-3):109-115.

47. Ikema M & Honma Y (1998) A novel filamentous phage, fs-2, of *Vibrio cholerae* O139. *Microbiology* 144 (Pt 7):1901-1906.

48. Jouravleva EA, *et al.* (1998) The *Vibrio cholerae* mannose-sensitive hemagglutinin is the receptor for a filamentous bacteriophage from *V. cholerae* O139. *Infect. Immun.* (66):2535-2539.

49. Jianying G, *et al.* (2009) Genomic and systems evolution in Vibrionaceae species. *BMC Genomics* 10.

50. Boyd JM, *et al.* (2008) Contribution of Type IV Pili to the Virulence of *Aeromonas salmonicida* subsp. *salmonicida* in Atlantic Salmon (Salmo salar L.). *Infect. Immun.* 76(4):1445.

51. Haley BJ, *et al.* (Comparative genomic analysis reveals evidence of two novel Vibrio species closely related to V. cholerae. *BMC microbiology* 10(1):154.

52. Ewing B & Greene P (1998) Base-Calling of Automated Sequencer Traces Using Phred. II. Error Probabilities. *Genome Research* 8:186-194.

53. Ewing B, Hillier L, Wendi M, & Greene P (1998) Base-Calling of Automated Sequencer Traces Using Phred. I. Accuracy Assessment. *Genome Research* 18:175-185.

54. Gordon D, Abajian C, & Green P (1998) Consed: A Graphical Tool for Sequence Finishing. *Genome Research* 1998(8).

55. Han CS & Chain P (2006) Finishing repeat regions automatically with Dupfinisher. *International Conference on Bioinformatics and Computational Biology*, eds Arabnia HR & Valafar H (CSREA Press), pp 141-146.

56. Aziz R, *et al.* (2008) The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9(1):75.

57. Konstantinidis KT & Tiedje JM (2005) Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci U S A* 102(7):2567-2572.

58. Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* 27(2):573-580.

59. Larkin MA, *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947.

60. Saitou N & Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol biol evol* 4(4):406.

61. Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 16(2):111-120.

62. Kumar S, Nei M, Dudley J, & Tamura K (2008) MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in bioinformatics*.

Dhaka University Institutional Repository

CHAPTER 7:

DEEP SEA BACTERIA RELATED TO HUMAN PATHOGENIC *VIBRIO* SPECIES

Dhaka University Institutional Repository

Abstract:

Vibrio species are both ubiquitous and abundant in marine coastal waters, estuaries, ocean sediment, and aquaculture settings worldwide. We report here the isolation, characterization, and genome sequence of a novel *Vibrio* species, *Vibrio antiquarius*, isolated from a mesophilic bacterial community associated with hydrothermal vents located along the East Pacific Rise, near the southwest coast of Mexico. Phenotypic and 16S rDNA similarity to *V. alginolyticus* and *V. parahaemolyticus* was found, but divergence was sufficient to warrant separate species status. The average nucleotide identity (ANI) of *V. antiquarius* with *V. alginolyticus* 12G01 and *V. parahaemolyticus* RIMD 2210633 was 92 and 85%, respectively, below the 95% threshold that has been suggested as a species demarcation criterion. Comparative genome analysis revealed a variety of genomic events including chromosomal rearrangements, gene deletions and acquisitions, chromosomal integration of phages, all of which may be important driving forces in the presence of *V. antiquarius* in the hydrothermal vent environment and evolution of coastal and estuarine species of *Vibrio*. Several genes and operons potentially related to its deep-sea vent habitat, as well as factors known to be involved in virulence in freshwater, coastal, and brackish water vibrios, were identified in *Vibrio antiquarius*. The presence of virulence factors associated with human illness in this deep sea *Vibrio* species suggests these genes play a role outside of the human body.

Introduction:

Deep-sea hydrothermal vents comprise a suite of diverse microenvironments, with gradients of nutrient abundance and extreme physicochemical conditions that vary from high to low with respect to temperature, oxygenation states, fluid velocities and, therefore, support diverse heterotrophic and chemoautotrophic communities across three domains of life (1). These extensive findings have also demonstrated the presence of mesophilic microbial communities as discovered in explorations the Mediterranean Sea and the Pacific and Atlantic Oceans (2-5). Although not directly vent-related, an

Dhaka University Institutional Repository

interesting set of 34 bacterial strains was isolated from mud-water samples collected at a depth of 4,940 m, 150 miles east of Cape Canaveral, Florida, containing *Vibrio, Aeromonas,* and *Pseudomonas* spp. (7). Clearly, vibrios are ubiquitous and highly abundant in the aquatic environment on a global scale, including deep sea sediments, (8-13) and it has been repeatedly shown that these bacteria are present at high densities in and/or on marine organisms, e.g., corals (14), fish (15-17), mollusks (18), seagrass, sponges, shrimp (16, 19) and zooplankton (16, 20, 21) (22, 23).

During the deep-sea submersibles Alvin and Nautile dives in 1999 along the East Pacific Rise, southwest of the Mexican coast, samples of water surrounding sulfide chimneys of a hydrothermal vent community were collected and four mesophilic bacterial strains were isolated and subsequently tested for phenotypic traits, including growth on TCBS agar (*Oxoid, Cambridge*, UK). Single gene phylogenies, 16S rRNA, 23S rRNA, and *toxR*, a well conserved gene in *Vibrionaceae* frequently used to reflect the phylogenetic relationship of this family, were used to identify the isolates, comparing the results with all known ribosomal RNA sequences and all *Vibrio toxR* sequences. The four isolates were identified as *Shewanella algae*, *V. harveyi* and two novel *Vibrio* species which were labeled as *Vibrio* sp. Ex25 and *Vibrio* sp. Ex97. *Vibrio* sp. Ex25 showed phenotypic and genetic similarity to the human pathogens *V. parahaemolyticus*, *V. alginolyticus*, and *V. cholerae*. Therefore, this genome was fully sequenced using a combination of short and long read technology. Results of the genomic analysis confirm a close relationship to both *V. alginolyticus* and *V. parahaemolyticus*, but warrant separate species status. The possession of a repertoire of virulence factors typically found in *V. parahaemolyticus* and *V. cholerae*, both of which are human pathogens, suggest functions of these genes in the natural environment may also be critical to life in the deep sea environment.

Dhaka University Institutional Repository

Methods:

Sample collection and culture isolation. Samples were collected from two sites on the East Pacific Rise, 9°50'N, 104°17' W by DSV *Alvin* and 13° N (12°49'N, 103°56'W) at a depth of 2,500 meters by DSV *Nautile*. Samples were inoculated into three different heterotrophic, anaerobic seawater-based media immediately after sample retrieval aboard ship. Samples collected from 9° N were inoculated into MSH medium (24)with 4 mM FeS, 4 mM H_2S, 0.5 g/L yeast extract, 0.5 g/L trypticase peptone, with and without 5 mM acetate. Samples from 13° N were inoculated into MSH medium amended with 0.05 g/L of yeast extract, 0.05 g/L trypticase peptone, 20 mM acetate, 0.3 g/L coenzyme M, and 15 mM of iron pyrophosphate. All samples were incubated at 25° C (9° N) or 30° C (13° N). Four pure cultures were obtained by serial dilution and from colonies picked from the roll tubes, designated A6 and BB4 (13°N), and EX25 and EX97 (9°N). The cultures were enriched aerobically in alkaline peptone water (APW) and Luria-Bertani broth (LB) and spread on TCBS (thiosulfate-citrate-bile-salts-sucrose agar). Yellow colonies (sucrose positive) on TCBS were reisolated on the same agar medium (APW or LB) used for enrichment. The strains were subjected to biochemical tests using API20E strips (Biomerieux Vitek, Inc. MO, USA). Salt tolerance was assayed in nutrient broth containing 3%, 6%, and 8% NaCl.

Identification by single gene phylogeny. DNA was extracted and 16S rRNA sequences were obtained, as described previously (25). Sequences were manually aligned in the Genetic Data Environment (GDE, available from ftp.bio.indiana.edu/soft/molbio/unix/GDE). Approximately 1,500 nucleotides were obtained from sequencing and approximately 1,250 nucleotides in evolutionarily conserved regions were used for maximum likelihood analysis. Maximum likelihood trees were constructed using fastDNAml (26). A portion of the 23S rRNA gene, 1.5 Kb, of each of the four isolates was also amplified and sequenced. The amplified region corresponded to position numbers 2,821 to 4,306 in section 227 of the *V. cholerae* chromosome I sequence (21), a genus specific region for *Vibrio* species (27). Sequences

Dhaka University Institutional Repository

were manually aligned and phylogenetic trees were built using ClustalW (28) and Tree View (29).

The relatedness of the isolates to reference strains of *Vibrio* spp. was also determined by total DNA-DNA hybridization, using random-primer labeling and chemiluminescent detection with DIG High Prime DNA Labeling and Detection Starter Kit II (Roche, USA). Ten replicates of each strain (10 ng) were hybridized with genomic DNA from 41 reference type strains of *Vibrionaceae* species plus four additional strains of *V. cholerae* and two of *V. mimicus* and their relative binding ratios (RBR) were calculated. The reference strains were also probed against one another as controls. Immunological identification, using antiDIG-AP and CSPD (chemiluminescent substrate) was performed, according to manufacturer's protocol (Roche, USA). Blots were exposed to X-ray film for 20 min to 4 h, depending on the signal obtained. Developed film was scanned using a densitometer (Personal Densitometer SI) and dots quantitatively evaluated using ImageQuaNT software (for Windows NT, Molecular Dynamics, ver. 4.2a, Sunnyvale, CA). ERIC-PCR was used to elucidate relationships between isolates and reference strains. This method successfully differentiates the closely related species *V. cholerae* and *V. mimicus* (30, 31). DNA, enzymatically extracted, was amplified, using ERIC-2 primer. (Homologous sequence with the ERIC-1 primer has not been demonstrated to be present in Vibrios).

Genome sequencing. The genome of *Vibrio* sp. EX25 was sequenced in part by the Joint Genome Institute (JGI) and all general aspects of sequencing performed at the JGI can be found at http://www.jgi.doe.gov/. 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 was accomplished and optional paired end Sanger sequencing on 35kb fosmid libraries to 1-2X coverage (depending on repeat complexity). The Phred/Phrap/Consed software package (www.phrap.com) was used for sequence assembly and quality assessment (32, 33). After the shotgun stage, reads were assembled with parallel phrap (High Performance

Dhaka University Institutional Repository

Software, LLC). Draft assemblies were based on 39,974 total reads. Repeat resolution was performed using Dupfinisher (34). Gaps between contigs were closed by editing in Consed and several targeted finishing reactions, including transposon bombs (35), primer walks on clones, primer walks on PCR products, and adapter PCR reactions. Gene-finding and annotation were achieved using the RAST server (36). The completed genome sequences of *Vibrio* EX25 contain 42,569 reads, achieving an average of 8-fold sequence coverage per base, with an error rate of less than 1 in 100,000. 16S rRNA Genbank accession numbers for *Vibrio* sp. BB4, EX25 and EX97 and *Shewanella algae* A6 are AF 319768; AF319769, AF319770, and AF 319767, respectively.

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 of genomic sequences examined in this study. Firstly, nucleotide sequences as whole contigs were directly aligned using the MUMmer program (37). Secondly, ORFs of a given pair of genomes were reciprocally compared with 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 (38) 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 although ORF-independent method performed better for draft sequences of low quality; where sequencing errors, albeit rare, hampered identification of correct ORFs.

Determination of the average nucleotide identity (ANI). To determine the average nucleotide identity (ANI) as a measure of genetic distance between strains and

Dhaka University Institutional Repository

assignment of strains to species groups, a reciprocal best match BLASTN analysis was performed for each genome. The average similarity between genomes was measured as the ANI of all conserved genes, following the methods of Konstantinidis and Tiedje (39). With this method, an ANI >94% with >85% of protein-coding genes conserved between a pair of genomes is concluded to correspond to strains of the same species, whereas an ANI <94% and <85% conservation of protein-coding genes corresponds to strains of different species. These criteria were employed in our analysis.

Phylogenetic analyses based on genome sequences. A set of orthologous genes from the genomes of *Vibrio* EX25 and representatives of close relatives were identified and aligned using CLUSTALW (*dnaG, frr, infC, nusA, pgk, pyrG, rplA, rplB, rplC, rplD, rplE, rplF, rplK, rplL, rplM, rplN, rplP, rplS, rplT, rpmA, rpoB, rpsB, rpsC, rpsE, rpsI, rpsJ, rpsK, rpsM, rpsS, smpB, tsf*). The alignments were manually curated and concatenated. Evolutionary history was inferred by Maximum Likelihood method (with bootstrapping) using PHYML, applying the JTT model of substitution, with a gamma-distribution of rates of 5 categories of variable sites (with alpha estimated by the program).

Results and Discussion:

Phylogenetic analysis of the 16S rRNA revealed that three of the four isolated strains belonged to the genus *Vibrio* (BB4, EX25 and EX97) and the fourth, strain A6 was identified as *Shewanella* algae. One of the three *Vibrio* isolates, BB4 was identified *as V harveyi* (Figure 7.1). Sequence analysis of the *Vibrio*-specific region of the 23S rRNA gene confirmed identification of BB4 as *V. harveyi* (data not shown). All data, including phenotypic tests, DNA-DNA hybridization and ERIC-PCR confirmed identification of A6 and BB4 as *Shewanella algae* and *V. harveyi*, respectively.

Strains EX97 and EX25 clustered with *V. parahaemolyticus* and *V. alginolyticus* by 16S rRNA analysis; both were 99% identical to 16S rRNA of *V. parahaemolyticus*

Dhaka University Institutional Repository

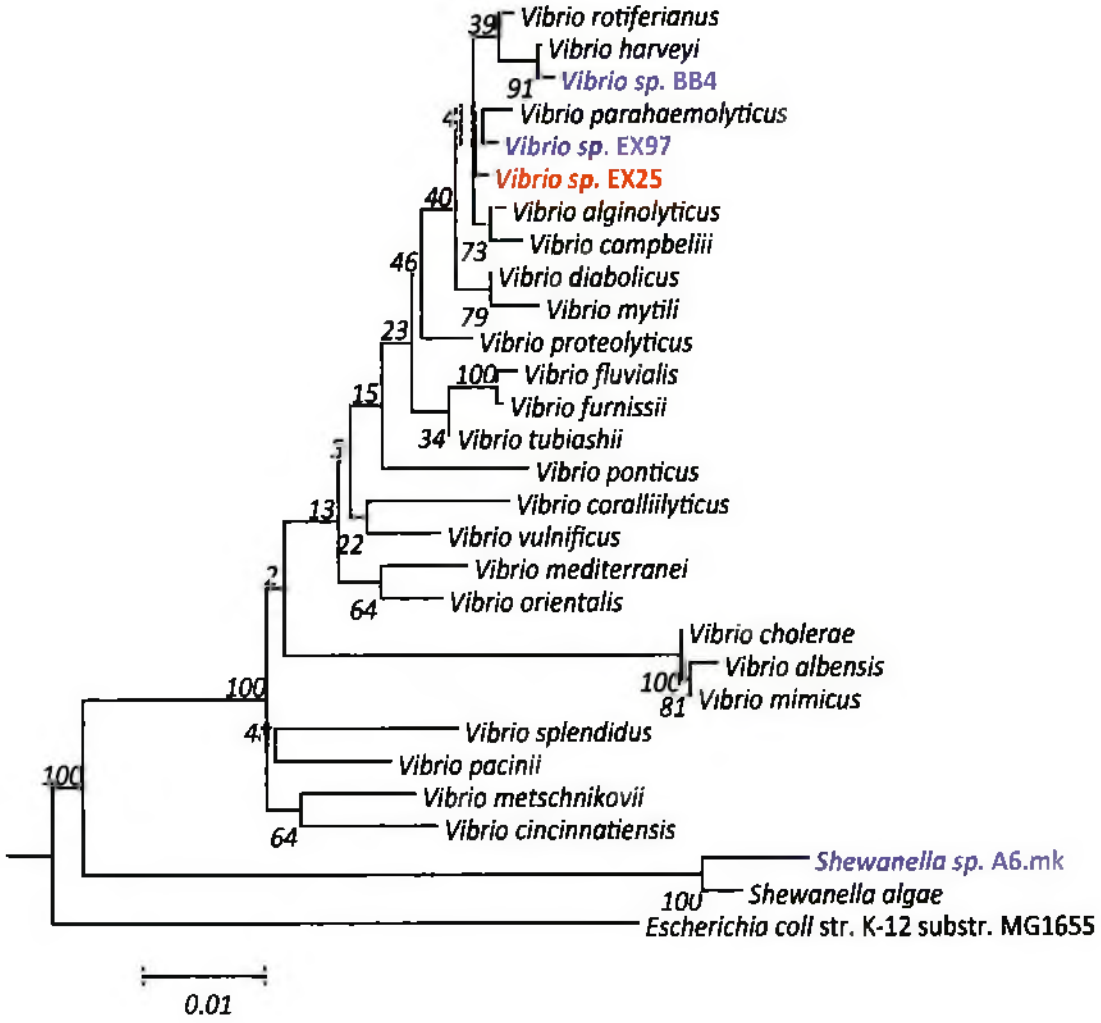
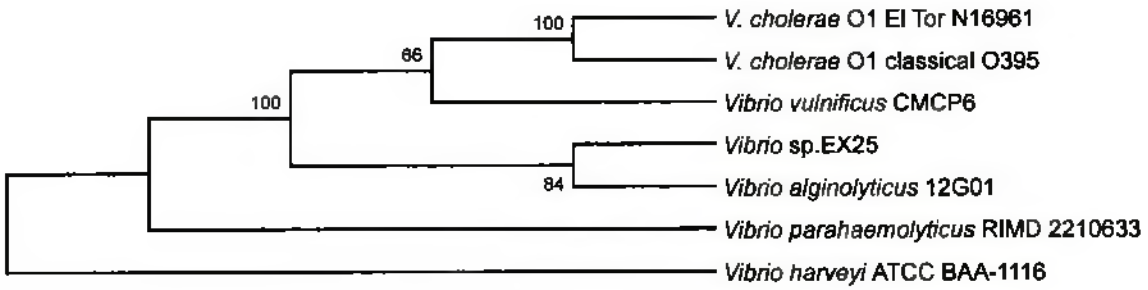


Figure 7.1: Evolutionary history was inferred using the Neighbor-Joining method employing 16S rDNA sequences obtained from the isolates from the East Pacific Rise (EPR) and other *Vibrio* species. The bootstrap consensus tree inferred from 5000 replicates is taken to represent the evolutionary history of the taxa analyzed. Scale bar represents 0.02 nucleotide substitutions per sequence position.

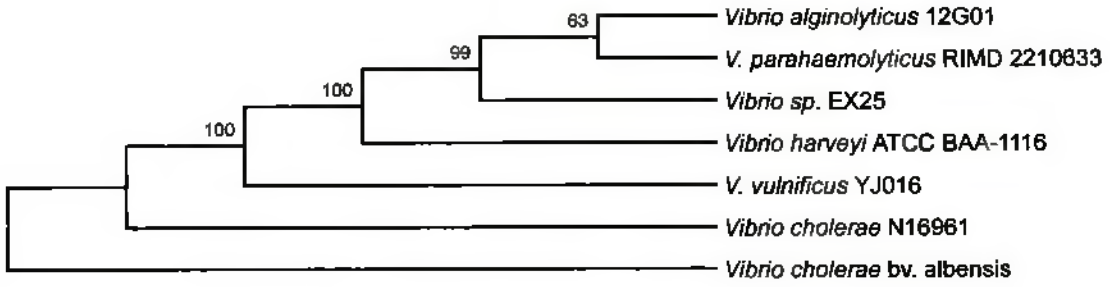
and *V. alginolyticus*. Since there was not enough variation in the 16S rRNA sequences to resolve the exact phylogeny of these strains further work was done to elucidate the identity of these two deep sea *Vibrio* strains. Both strains were sucrose-positive, oxidase positive, and required NaCl for growth. They grew well under both microaerophilic (Gas Pak anaerobic system (H_2 and CO_2), Beckton-Dickson) and strict anaerobic conditions

Dhaka University Institutional Repository

(80% N_2, 10% CO_2 and 10% H_2). The phylogenetic tree (Fig. 7.2a), based on sequences of the 23S rRNA genes, clustered *Vibrio* EX25 with *V. alginolyticus* and *V. parahaemolyticus* and *V. harveyi* branched separately. As protein-encoding genes have been reported to evolve much faster than rRNA genes and, therefore, are expected to provide higher resolution in phylogenetic analyses, another tree was constructed (Fig. 7.2b), based on *toxR* sequences. The results showed *Vibrio* EX25 branching separately from *V. alginolyticus* and *V. parahaemolyticus*.



A. 23S phylogeny: Neighbor Joining



B. toxR phylogeny: Neighbor Joining

Figure 7.2. Evolutionary history was inferred using the Neighbor-Joining method employing (A) 23S rDNA and (B) ToxR sequences obtained from the isolates from the East Pacific Rise (EPR) and other *Vibrio* species. The bootstrap consensus tree inferred from 500 replicates is taken to represent the evolutionary history of the taxa analyzed. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree.

Dhaka University Institutional Repository

To evaluate the genomic relatedness of these two isolates to other *Vibrio* species, DNA-DNA hybridization was conducted with 41 type strains of *Vibrionaceae* and their relative binding ratios (RBR) were calculated. The range of strains was selected (1) to cover the entire span of genera under *Vibrionaceae* and (2) to cover the entire span of species within the genus *Vibrio*. The first criterion was met by including *Grimontia hollisae*, *Photobacterium* spp, and *Salinovibrio* spp. According to the phylogenetic trees of 16S rRNA sequences, the new genus is located below the branching point of the four genera included in this study.

Vibrio EX25 yielded RBR values that were less than 15% when compared with *V. cholerae* and *V. mimicus,* with the conclusion that EX25 is not *V. cholerae* (intra-species level). The highest RBR values, 42%, were obtained when *Vibrio* EX25 was hybridized with DNA of the type strain of *V. alginolyticus* and *V. parahaemolyticus*. However, the lack of divergence in the 16S rRNA sequence of EX25 from *V. alginolyticus* and *V. parahaemolyticus* indicated that additional information was needed to determine the taxonomy of *Vibrio* EX25. Since this strain was both possibly a different species and also from a novel habitat we undertook whole genome sequencing

Genome Features. General features of the genome of *Vibrio* EX25 are summarized in Table 7.1 and Fig. 7.3. Examination of the gene content of each chromosome revealed that most of the essential genes required for growth and viability are on chromosome I, similar to *V parahaemolyticus*, while corresponding genes of *V cholerae* are on chromosome II. Genes for all the ribosomal proteins and at least one copy of all the tRNA genes were also found on *Vibrio* EX25 chromosome I. However, several genes involved in essential metabolic pathways (e.g., glycolysis) and virulence are located on chromosome II, suggesting that, as for both *V. cholerae* and *V. parahaemolyticus,* the *Vibrio* EX25 chromosome II is essential for growth and viability. Most of the housekeeping genes are on chromosome I as in other vibrios though some probably essential genes are found on chromosome II.

Dhaka University Institutional Repository

Table 7.1. General features of the *Vibrio antiquarius* EX25 genome

Feature	*V. antiquarius* EX25	*V. alginolyticus* 12G01	*V. parahaemolyticus* RIMD 2210633	*V. harveyii* ATCC BAA-1116	*V. vulnificus* CMCP6	*V. cholerae* N16961
Chromosome size (Mb)	5.09	5.16*	5.17	5.97	5.12	4.03
Chromosome I	3.26	5.16	3.29	3.77	3.28	2.96
Chromosome II	1.83		1.88	2.2	1.84	1.07
Avg. G+C content (mol%)	45	44	45	45	46.4	46.7
Chromosome I	45	44	45	45	46	47
Chromosome II	45		45	45	47	46
Protein coding sequences	4,529	4,666	4,832	5,920	4,488	3,835
Chromosome I	2,846	4,666	3,080	3,546	2,926	2,742
Chromosome II	1,683		1,752	2,374	1,562	1,093
Average CDS length (bp)	942	947	950	910	918	915
Chromosome I	925	947	952	923	887	948
Chromosome II	975		946	890	977	832
Percent of coding region	86.6	85.6	86	85.4	84	86
Chromosome I	87	85.6	86	85	83	87
Chromosome II	86		86	86	86	84
Ribosomal RNA operons	11	11	11	11	9	8
Chromosome I	10	11	10	10	8	8
Chromosome II	1		1	1	1	0
Number of tRNAs	124	69*	126	121	112	98
Chromosome I	111	69*	112	105	98	94
Chromosome II	13		14	16	13	4
Number of Integrons	1	0	1	?	1	1
Chromosome I	SI	0	1	?	SI	0
Chromosome II	0		0	0	0	SI

*Draft genome, Therefore, numbers are approximate; SI= Super Integron

Dhaka University Institutional Repository

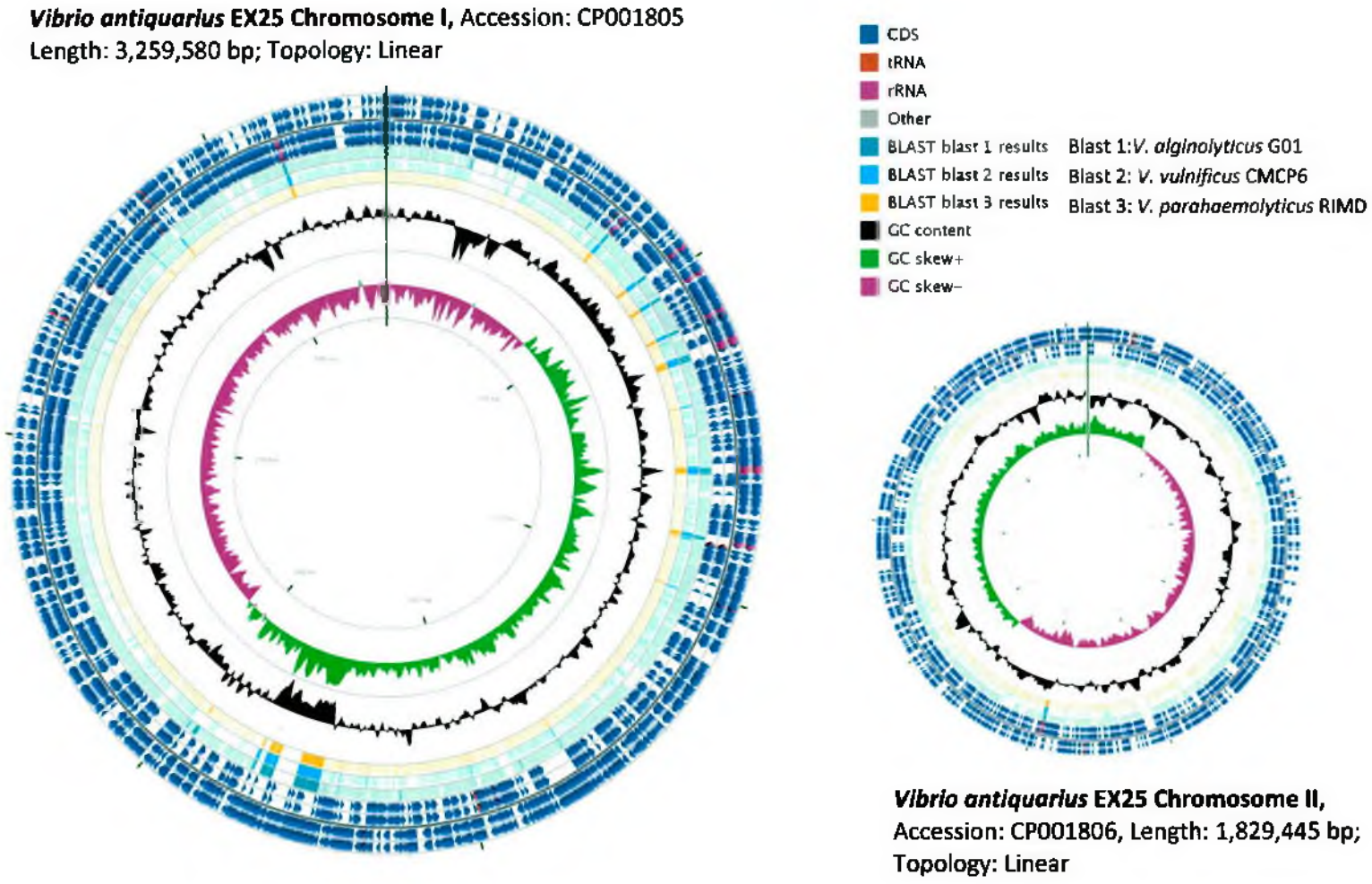


Figure 7.3: Circular maps of the chromosomes of *Vibrio antiquarius* EX25 genome generated by CGView. From the outside in, the first-second and third-fourth circles show CDSs transcribed clockwise and anticlockwise, respectively. Circles 5, 6 and 7 display Blast bi-directional hit against *V. alginolyticus* 12G01, *V. parahaemolyticus* RIMD 22110633 and *V. vulnificus* CMCP6 , circle 8 shows GC content, and circle 9, the GC skew.

Dhaka University Institutional Repository

Taxonomic standing of *Vibrio* EX25 by Average Nucleotide Identity (ANI). Species delineation as determined by the average nucleotide identity (ANI) between genomes has been employed with *Burkholderia, Escherichia coli, Salmonella* spp. and *Shewanella* spp. (46). Highest ANI was observed between *Vibrio* EX25 and *V. alginolyticus* 12G01 (92.07%) and *V. parahaemolyticus* RIMD 2210633(85%), below the threshold value of 95%, a hypothesized species demarcation for prokaryotes (47) (Table 7.2). These data along with finding form 16S rDNA, 23S rDNA, and *toxR* phylogeny and DNA–DNA hybridization led us to the conclusion that *Vibrio* EX25 is most closely related to *V. alginolyticus* and *V. parahaemolyticus,* but warrant that *Vibrio* EX25 is a novel *Vibrio* species, for which we propose the species name *Vibrio antiquarius*, a bacterium of the aquatic environment and derived from the antiquity of the deep sea.

V. antiquarius **EX25**	***V. alginolyticus*** **12G01**	***V. parahaemolyticus*** **RIMD2210633**	***V. harveyi*** **ATCC BAA-116**	***V. vulnificus*** **CMCP6**
ANI of conserved genes having >70% nucleotide identity	92.07%	84.99%	81.86%	78.23%
ANI of all conserved genes	91.25%	84.19%	80.25%	74.53%

Table 7.2. Average nucleotide identity (ANI) of conserved genes between *V. antiquarius* and *V. alginolyticus, V. parahaemolyticus, V harveyii,* and *V. vulnificus.*

To examine the predicted protein-coding sequences (CDS) of *Vibrio antiquarius* EX25 in finer detail, a multigenome comparison of *V. antiquarius* with eight genomes, all of which had been annotated or reannotated, based on subsystem technology (36), was performed. In agreement with results from ANI analysis, the *V. antiquarius* EX25 genome was more similar to that of *V. alginolyticus* and *V. parahaemolyticus* (Fig. 7.4). A Venn diagram (Fig. 7.5) depicts the *V. antiquarius* EX25 genome as sharing the largest number of CDSs with *V. parahaemolyticus* (3973, 87.7%) followed by *V. alginolyticus* (3943, 87%), *V. harveyi* (3567, 79%) and *V. vulnificus* (3309, 73%). The CDS shared with *V. parahaemolyticus* (3973) corresponds to 82% of the *V. parahaemolyticus*

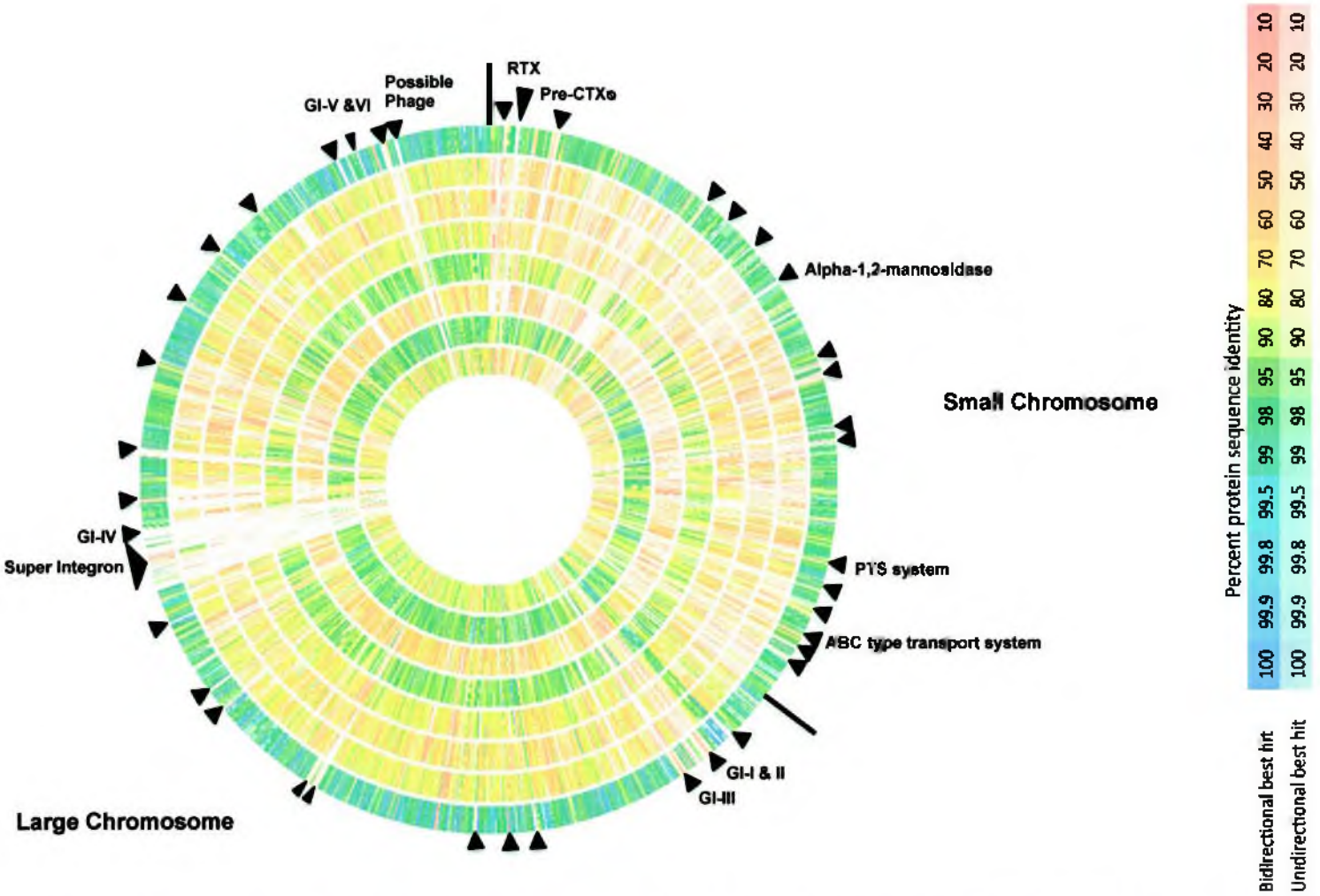


Figure 7.4: Genome comparison of *Vibrio antiquarius* EX25 with other *Vibrio* genomes. From outer ring to inner ring; *Vibrio alginolyticus* 12G01, *Vibrio cholerae* O1 biovar eltor str. N16961, *Vibrio fischeri* MJ11, *Vibrio furnissii* CIP 102972, *Vibrio harveyi* ATCC BAA-1116, *Vibrio hollisae* CIP 102972, *Vibrio parahaemolyticus* RIMD 2210633, *Vibrio vulnificus* CMCP6. Horizontal lines were drawn to separate chromosomes. Black Arrows indicates areas of rearrangement or insertion.

Dhaka University Institutional Repository

Dhaka University Institutional Repository

genome, whereas the number of shared CDS with *V. alginolyticus* (3943) constitutes 84.5% of the *V. alginolyticus* genome. Approximately 40 regions (including 17 in the small chromosome) of the genome of *V. antiquarius* encompassed at least three CDSs, has no reciprocal match in the *V. alginolyticus* genome (Fig. 7.4). These chromosomal regions contributed significantly to the 586 CDSs (13% of the *V. antiquarius* genome) with no reciprocal match in the *V. alginolyticus* genome. Forty three percent (256 CDSs) of these CDSs are either hypothetical or proteins of unknown function. The genome of *V. alginolyticus* contained 723 CDSs (15% of its genome) with no reciprocal match in the *V. antiquarius* genome (Fig. 7.5).

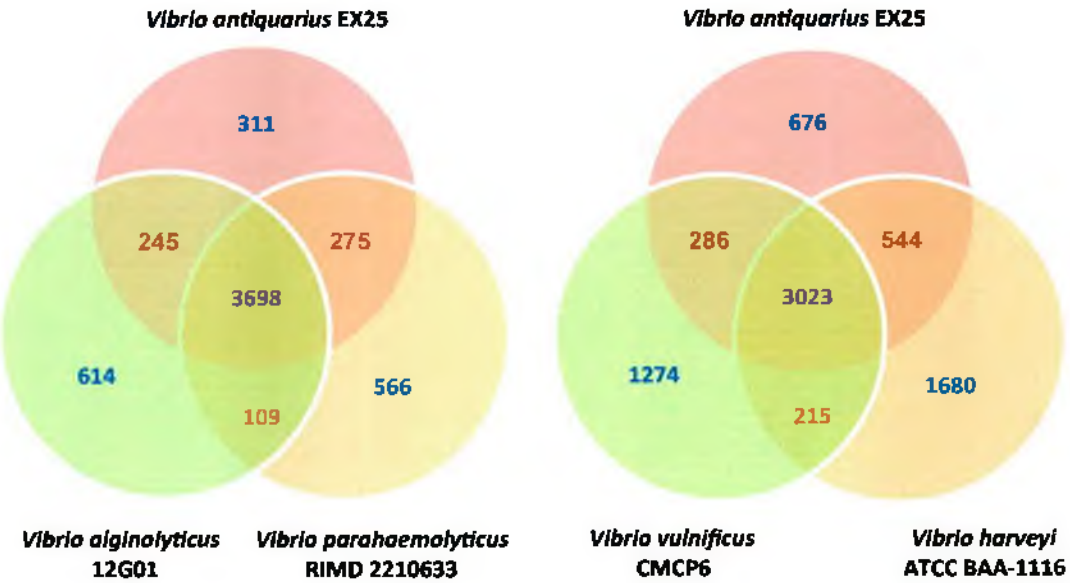


Figure 7.5: Venn diagrams showing number of unique and shared orthologs between *Vibrio antiquarius* EX25 and other *Vibrio* genomes.

Predicted biology of *V. antiquarius*

Response to Deep Sea Environment. For simplicity we present the discussion without saying "predicted" for all the metabolic properties, but it should be noted that most of what we are discussing is predicted properties and these have not been tested experimentally. The genome of *V. antiquarius* EX25 contains genes encoding properties,

Dhaka University Institutional Repository

which are commonly found in the deep-sea hydrothermal vent environment (Table 7.3). These include, cytochromes for reduction of O_2 to H_2O_2 and cytochrome C_{551} peroxidase that detoxifies peroxide, multiple catalase genes and a superoxide dismutase for tolerating high O2 concentrations, genes encoding alkyl hydroperoxide reductase that function to scavenge endogenous hydrogen peroxide. The ability to scavenge endogenous hydrogen peroxide is a trait that is absent in the other *Vibrio* genomes available for comparison and it is the major antioxidant enzyme for the endosymbiont of the tubeworm, *Riftia pachyptila,* inhabiting deep-sea vents (48). In addition, the presence of metalloendopeptidases and zinc-dependent carboxypeptidases in the genome of *V. antiquarius* can be interpreted as functioning in high concentrations of heavy metals, including zinc, in the hydrothermal vent environment (49).

Predicted Biology	Chromosome	Function
Response to Environment		
cytochrome *bd*	C-II (2)	Protection against O2 and H2O2
cytochrome C_{551} peroxidase	C-II	
catalase	C-II (2)	Tolerating high oxygen concentrations
superoxide dismutase	C-I (2)	
alkyl hydroperoxide reductase	C-II	*Scavenge endogenous hydrogen peroxide, a trait absent in the other Vibrio genomes*
methionine sulfoxide reductases	C-I	Oxidative damage repair
metalloendopeptidases	C-II (2)	Functioning in high concentrations of heavy metals, including zinc
zinc-dependent carboxypeptidases	C-II	
Cell sensing system		
LuxP & LuxQ	C-I	Autoinducer-2 (AI-2) mediated quorum sensing, biofilm formation, virulence, and other metabolic functions
LuxS & LuxN	C-II	
Biofilm-Related Pathways		
ornithine and arginine decarboxylase	C-I	Polyamine biosynthesis
C-di-GMP phosphodiesterase mbaA	C-I	Norspermidine
syp gene cluster	C-I	Gene clusters mediating biofilm formation
polar flagellum cluster, PF	C-I	Flagellar clusters
lateral flagellum cluster, LF	C-II	
Others		
rmf and *sulA*	C-I, C-II	persister cells
multicopper oxidase	C-I	*Mn(II) oxidation*
delta-9 fatty acid desaturase	C-II	*Fatty acid unsaturation, essential for growth under high pressure*
universal stress proteins, USPs	C-II (3)	function not defined

Table 7.3: Predicted biology of *V. antiquarius* EX25

Dhaka University Institutional Repository

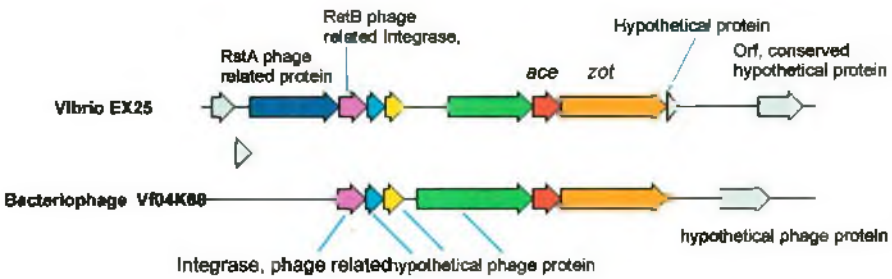
Deep-sea hydrothermal vents are a major source of dissolved elements, including Mn(II) in deep sea sediments of the world oceans (69). Mn has long been used as a reliable tracer of hydrothermal vent emissions (69-71). Unlike other vibrios, the genome of *V. antiquarius* contains a gene annotated as multicopper oxidase, an enzyme essential for manganese oxidation and laccase-like activity, all known to be involved in response to the presence of elements in the hydrothermal vent environment (72). Fatty acid unsaturation is a critical cellular process essential for growth under high pressure and genes such as delta-9 fatty acid desaturase are presumably up-regulated to increase membrane unsaturation (76). *V. antiquarius* contains delta-9 fatty acid desaturase (VEA.0292), facilitating the bacterium to withstand the high pressure of the deep sea.

Virulence Traits. The genome of *V. antiquarius* encodes a type III secretion system (TTSS), responsible for enabling injection of effector proteins directly into target host eukaryotic cells (79). Our investigation further revealed, unlike *V. parahaemolyticus, V. alginolyticus, V. harveyi,* and *V. vulnificus*, the genome of *V. antiquarius* contains two clusters of T6SS genes on the smaller chromosome. It was found to possess the thermolabile hemolysin (*tlh*) gene in the small chromosome, with 97% and 84% nucleotide similarity to that of *V. alginolyticus* and *V parahaemolyticus* respectively, but lacked the *tdh* and *trh* of *V parahaemolyticus.* The genome of *V. antiquarius* also harbors a pre-CTX prophage, evidenced by the presence of accessory cholera enterotoxin, *ace*, zonula occludence toxin, *zot*, RstA phage related protein, and RstB phage related integrase genes on the small chromosome. There was an insertion of bacteriophage genes between *rstA* and *ace*, with 96% homology to bacteriophage BfO4K68 (Fig. 7.6a). This region is most likely horizontally transferrable via the bacteriophage BfO4K68.

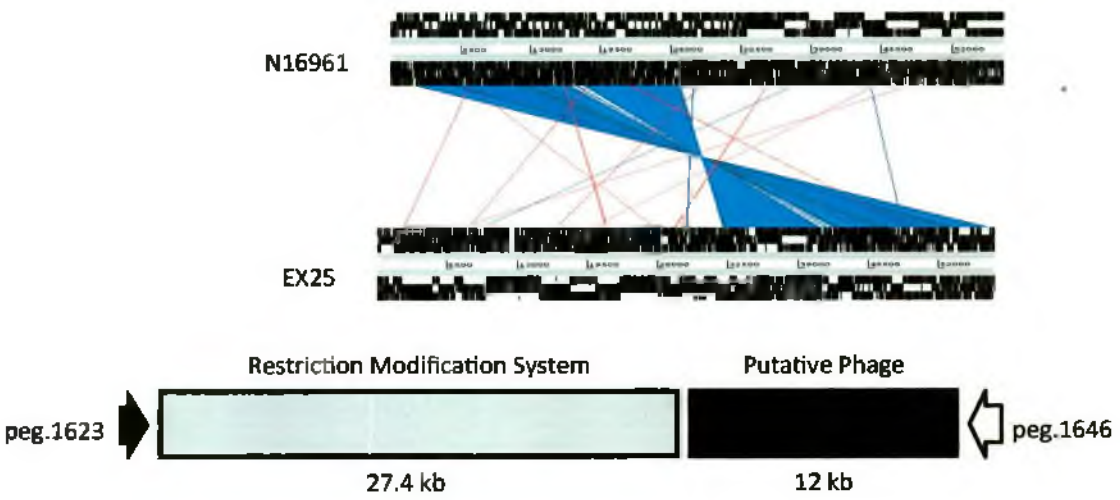
V. antiquarius also has a ca. 27.4 kb contiguous region on chromosome I (Fig. 7.6b), with 86% nucleotide sequence similarity to the *Vibrio* Pathogenicity Island 2 (VPI-2) found in both clinical and environmental *V. cholerae*. This region spans VC1758 to

Dhaka University Institutional Repository

VC1772 of the canonical VPI-2 and encodes a Type I restriction-modification system and five hypothetical proteins. However, sections VC1773 to VC1810 are not present; instead there is a 12 kb region that includes six hypothetical proteins and a phage integrase. The 12 kb insert does not have significant match with any sequences in the NCBI Genbank database, except for a 103 bp region with 78% sequence similarity to a transcriptional regulator of *V. parahaemolyticus*. Recent analyses of the VC1773 to VC1810 region of this island suggest that this section is a hot-spot for novel DNA insertion (101). The presence of virulence factors, including two primarily found in *V cholerae*, namely pre-CTXΦ and VPI-2 in the non-cholera *Vibrio antiquarius* from a deep-sea environment suggests a multifaceted role outside the human host.



A. Pre-CTXphi



B. VPI-2 like element in EX25

Figure 7.6. Schematic representation of a) Pre-CTXphi and b) VPI-2 like element identified in the genome of *Vibrio antiquarius* EX25.

Dhaka University Institutional Repository

Genomic Islands. Using Island Viewer (111) to study the genome of *V. antiquarius*, twenty-seven genomic islands were predicted (Fig. 8). Of these, 22 islands are on chromosome I and the other five on chromosome II. The average size of the islands is 13.3 Kb, with a G+C content ranging from 36 to 40%, much lower than the overall 45% chromosomal G+C content. The five GIs on chromosome II had a G+C content ranging from 38 to 42% and average size of 7 Kb. GI-I encodes seven proteins (Fig. 9) and 7% of the islands displayed homology (74%) with the O-antigen gene cluster of *Escherichia coli* serotype O98: K?:H8 and *Shigella dysenteriae* strain M13547. GI-II encodes 12 proteins (VEA.1776 – VEA.1787), many of which are capsular polysaccharide biosynthesis proteins (Fig. 9). The island exhibited 68% sequence similarity with the *Escherichia coli* capsule transport proteins, but only 10% query coverage. Other genomic islands inferred by these methods are shown below. A total of 222 strain-specific CDSs in the genome of *V. antiquarius* were identified (4.9% of the total CDSs), with no reciprocal match in the genomes of *V. parahaemolyticus*, *V. vulnificus*, *V. harveyi*, and *V. cholerae*. Among *V. antiquarius* specific protein-coding sequences, 116 (52.25%) are hypothetical or proteins of unknown function.

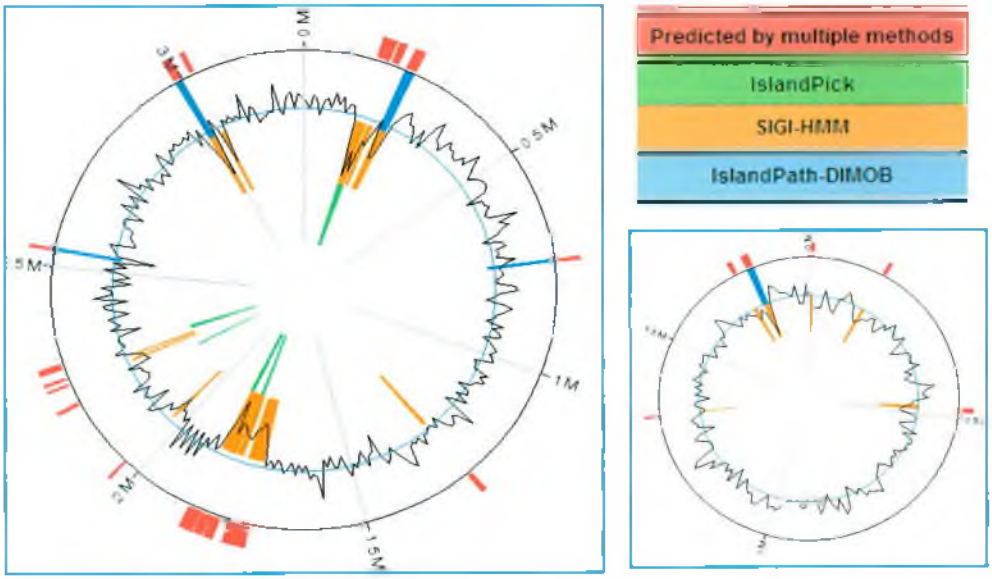


Figure 7.8. Graphical view of the distribution of genomic islands (GI's) in a) the large chromosome (C-I); and b) the small chromosome (C-II) of *Vibrio antiquarius* EX25. Color corresponds to the GI prediction methods.

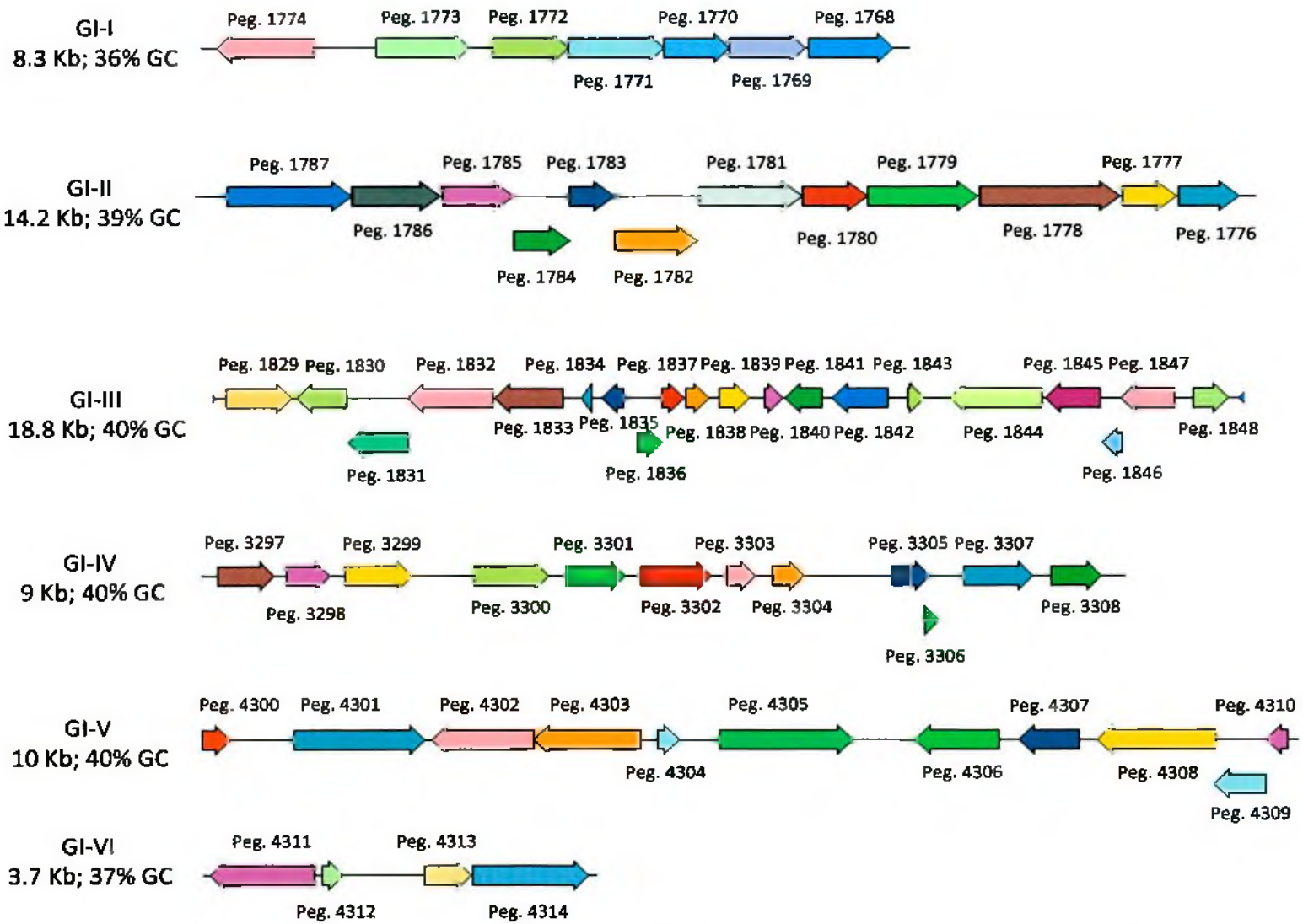


Figure 7. 9: Schematic representation of the putative genomic islands identified in *Vibrio antiquarius* EX25

Dhaka University Institutional Repository

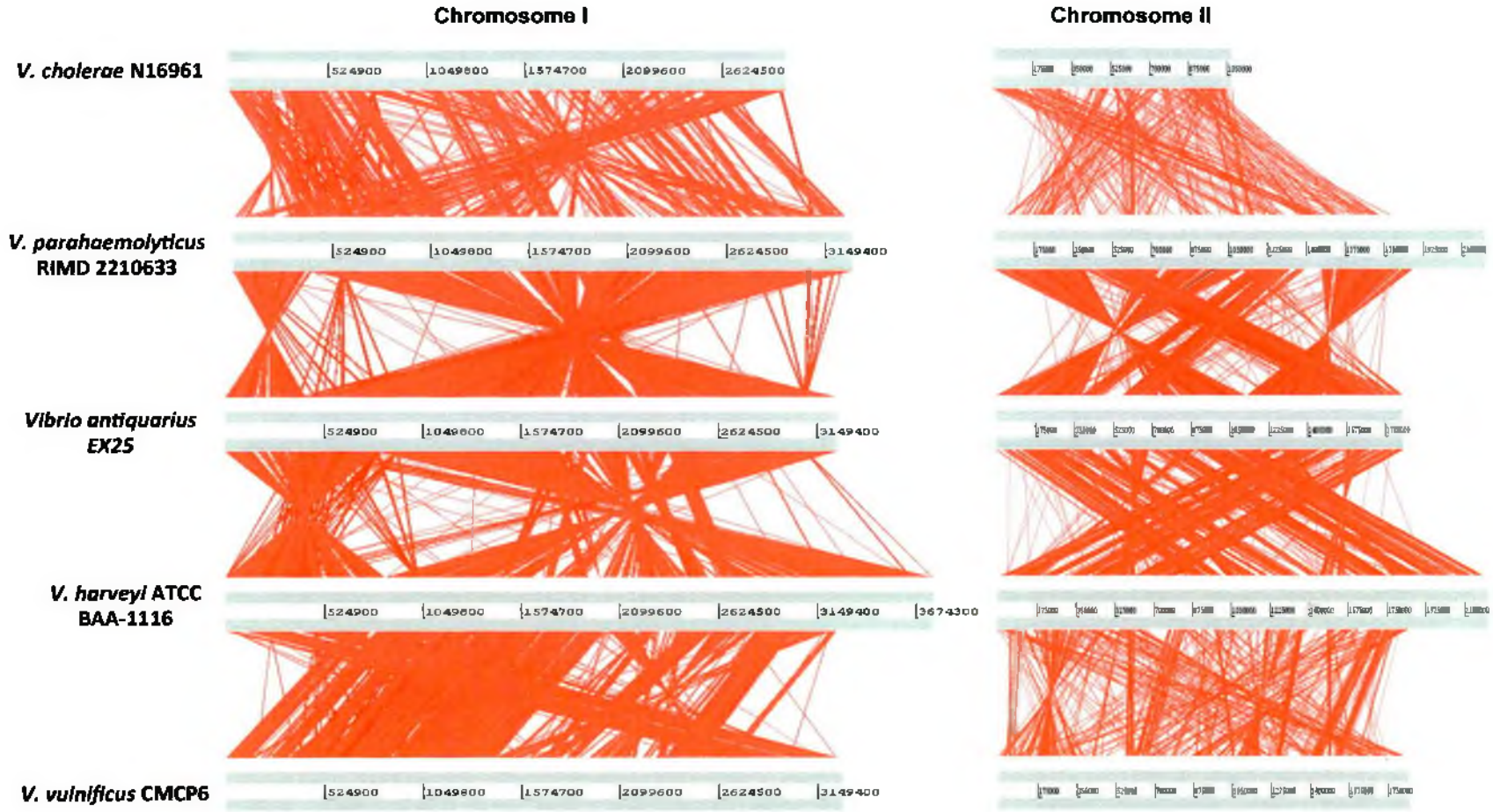


Figure 7.10: Linear pairwise comparison of *Vibrio antiquarius* EX25 genome. The analysis was performed using Artemis Comparison Tool (ACT). Connecting red lines between the genomes highlights regions with similarity; red lines indicate homologous blocks of sequence and inversions whereas gaps are indicative of unique DNA. The grey bars represent the forward and reverse strands. EX25 has high degree of similarity with *V. parahaemolyticus* and *V. harveyii*, despite of genome wise rearrangements.

Dhaka University Institutional Repository

Genome Rearrangement. The evolution of *V. antiquarius* EX25 genome structure is marked by several intra and inter-chromosomal rearrangements, compared to pathogenic *Vibrio* spp., namely *V. parahaemolyticus*, *V. harveyi*, *V. cholerae*, and *V. vulnificus* (Fig. 7.10). The presence of several large, inverted locally-collinear blocks indicates that inversions of large sections of the genome have occurred in the genomes of the pathogenic *Vibrio* spp., compared to *V. antiquarius,* supporting the notion that extensive genome plasticity is common in *Vibrio* species (112). Conservation of gene order and sequence similarity exists between the large chromosomes of *V. antiquarius, V. parahaemolyticus,* and *V. harveyi.* However, genomic comparisons of the large chromosome of *V. parahaemolyticus* and *V. harveyi* with *V. antiquarius* also reveal rearrangements and deletions, exhibiting a recognizable pattern (Fig. 7.10). Only a limited sequence similarity was found with other species for chromosome II, except for *V parahaemolyticus*, which exhibited a pattern of large blocks of inversions, rearrangements and deletions as in its large chromosome (Fig. 7.10). Moreover, whole genome sequence alignment analysis, using the MUMmer program, revealed extensive genomic rearrangements in chromosome II of *V. antiquarius*, compared with *V. parahaemolyticus* and *V. harveyi* (Fig. 7.11).

Using the tandem repeat finder program (114), we identified 64 tandem repeats with period lengths of 6 to 417 and 20 with period lengths of 6 to 429, respectively, throughout chromosome I and Chromosome II of *V. antiquarius*. In general, tandem repeats are generated by duplication, which changes the genome structure. Thus, it may be one of the important processes of evolution for *V. antiquarius* in its hydrothermal vent habitat. Interestingly, many of the tandem repeats are in genes encoding proteins and exhibiting high mutation rates.

Dhaka University Institutional Repository

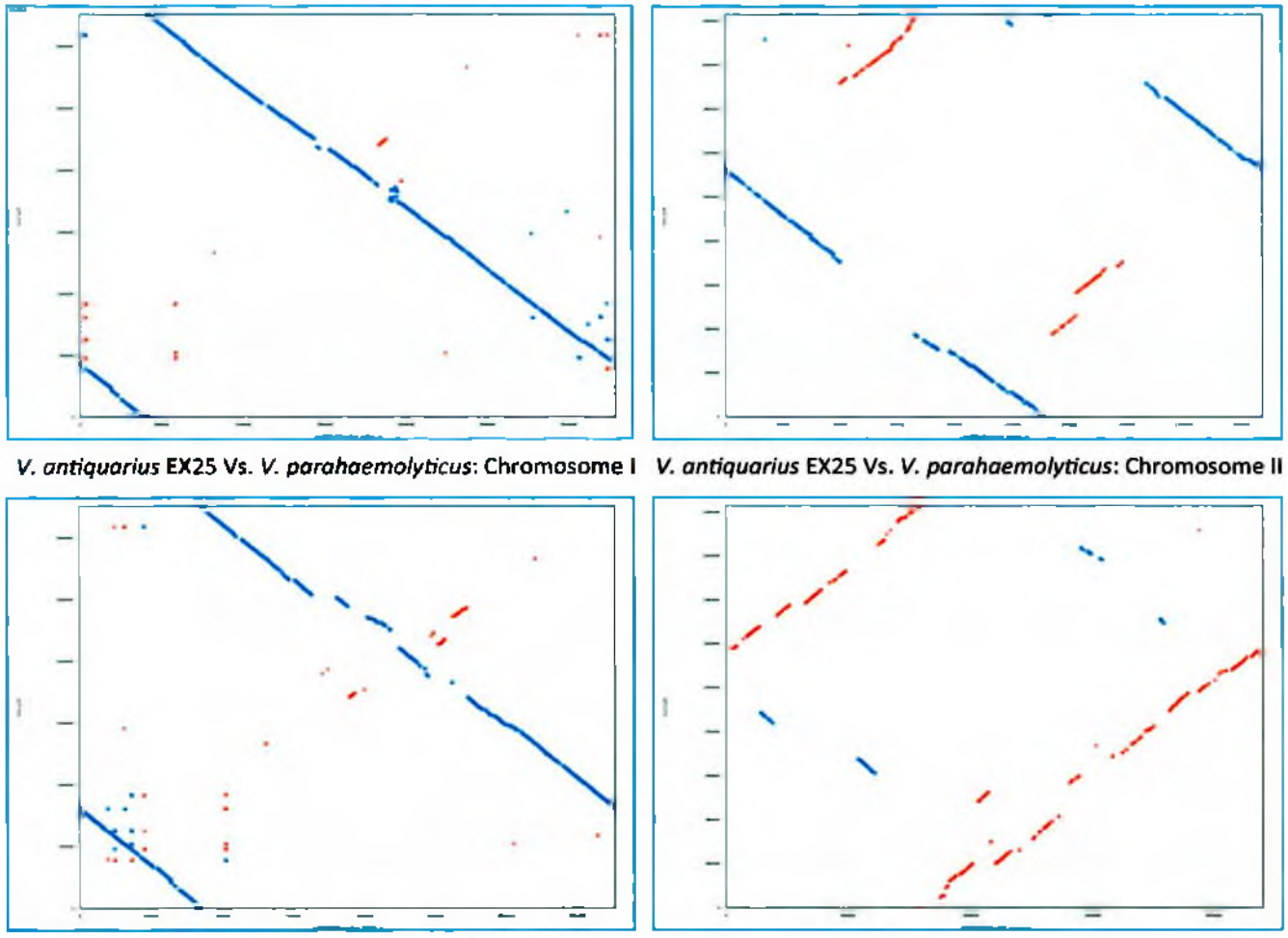


Figure 7.11: Chromosome wise alignment of *Vibrio antiquarius* EX25 genome with *V. parahaemolyticus* and *V. harveyi* genomes by MUMmer demonstrates evidence of genomic inversion and rearrangement.

Dhaka University Institutional Repository

Phylogeny

For more precise phylogenetic inference, a Maximum Likelihood Tree was constructed, using homologous alignment of 31 housekeeping genes present in the genomes of *V. antiquarius* and related vibrios (Fig. 7.12). *V. antiquarius, V. parahaemolyticus,* and *V. harveyi* formed a monophyletic group, with *V. antiquarius* clustered with *V. parahaemolyticus* and *V. harveyi* as an outgroup to both of them. Strains of *V. cholerae* and *V. vulnificus* were each monophyletic within the species. The bootstrap support for the position of *V. antiquarius* with *V. parahaemolyticus* and *V. harveyi* is strong indicating that at least the core of these genomes shares a common ancestry that excludes *V. cholerae*.

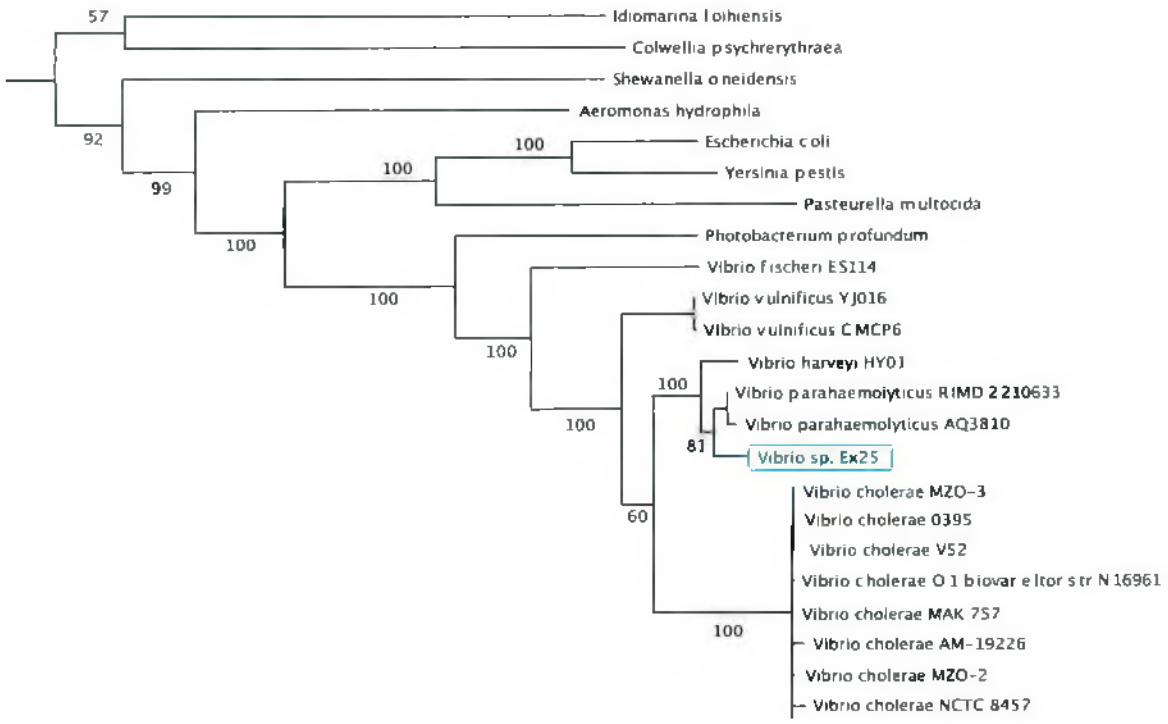


Figure 7.12. Protein sequences of 31 housekeeping genes *(dnaG, frr, infC, nusA, pgk, pyrG, rplA, rplB, rplC, rplD, rplE, rplF, rplK, rplL, rplM, rplN, rplP, rplS, rplT, rpmA, rpoB, rpsB, rpsC, rpsE, rpsI, rpsJ, rpsK, rpsM, rpsS, smpB, tsf)* from genomes of *Vibrio* spp. and representatives of their close relatives were identified and aligned using CLUSTALW. The aligments were manually curated and then concatenated. Maximum likelihood trees (with bootstrapping) were made using PHYML, applying the JTT model of substitution, with a gamma-distribution of rates of 5 categories of variable sites.

Dhaka University Institutional Repository

Conclusion:

Because *V. cholerae* is a marine/estuarine bacterium, based on its halophilic nature and, like all vibrios, is autochthonous to the aquatic environment, it was of interest to pursue the potential origin and evolution of vibrios from the deep sea environment. When samples collected from deep sea hydrothermal vents revealed the presence of vibrios phenotypically similar to *Vibrio* species pathogenic for humans, the genomes of one of these deep sea isolates was sequenced. In the course of our studies, we determined the complete genome sequence of the deep sea *Vibrio* EX25. Genomic and other comparisons demonstrate significant distance from any other species and thus we conclude this is a new *Vibrio* species. We propose the name *Vibrio antiquarius*. The presence of many homologs of *Vibrio* virulence genes in this deep sea species raises important questions about these genes serving much more basic purposes in the environment, rather than specificity in pathogenicity for humans. It is possible that some of these genes are more connected to commensal relationships, as it is known that *V. cholerae* is commensal with zooplankton, notably copepods (117). The presence of these genes in diverse *Vibrio* species (diverse phylogenetically and ecologically) suggests that these affiliations between commensals are likely very old and might help explain the common evolution of *Vibrio* species into pathogens of humans and marine animals.

Dhaka University Institutional Repository

References:

1. Jannasch HW (1995) Microbial interaction with hydrothermal fluids. in *Sea floor hydrothermal systems: physical, chemical, biological, and geological interactions* ed S.E. Humphris RAZ, L.S. Mullineaux, R.E. Thomson (American Geophysical Union, Washington, D.C.), pp 273-296.
2. Moyer CL, Dobbs FC, & Karl DM (1995) Phylogenetic diversity of the bacterial community from a microbial mat at an active, hydrothermal vent system, Loihi Seamount, Hawaii. *Appl Environ Microbiol* 61(4):1555-1562.
3. Raguenes G, Christen R, Guezennec J, Pignet P, & Barbier G (1997) Vibrio diabolicus sp. nov., a new polysaccharide-secreting organism isolated from a deep-sea hydrothermal vent polychaete annelid, Alvinella pompejana. *Int J Syst Bacteriol* 47(4):989-995.
4. Reysenbach AL, Longnecker K, & Kirshtein J (2000) Novel bacterial and archaeal lineages from an in situ growth chamber deployed at a Mid-Atlantic Ridge hydrothermal vent. *Appl Environ Microbiol* 66(9):3798-3806.
5. Sievert SM, Kuever J, & Muyzer G (2000) Identification of 16S ribosomal DNA-defined bacterial populations at a shallow submarine hydrothermal vent near Milos Island (Greece). *Appl Environ Microbiol* 66(7):3102-3109.
6. Karl DM, Wirsen CO, & Jannasch HW (1980) Deep-Sea Primary Production at the Galapagos Hydrothermal Vents. *Science* 207(4437):1345-1347.
7. Schwarz JR, Walder JD, & Colwell RR (1974) Deep-sea bacteria: growth and utilization of hydrocarbons at ambient and in situ pressure. *Appl Microbiol* 28(6):982-986.
8. Barbieri E, *et al.* (1999) Occurrence, diversity, and pathogenicity of halophilic Vibrio spp. and non-O1 Vibrio cholerae from estuarine waters along the Italian Adriatic coast. *Appl Environ Microbiol* 65(6):2748-2753.
9. DeLoney-Marino CR, Wolfe AJ, & Visick KL (2003) Chemoattraction of Vibrio fischeri to serine, nucleosides, and N-acetylneuraminic acid, a component of squid light-organ mucus. *Appl Environ Microbiol* 69(12):7527-7530.
10. Heidelberg JF, Heidelberg KB, & Colwell RR (2002) Bacteria of the gamma-subclass Proteobacteria associated with zooplankton in Chesapeake Bay. *Appl Environ Microbiol* 68(11):5498-5507.
11. Reidl J & Klose KE (2002) *Vibrio cholerae* and cholera: out of the water and into the host. *FEMS Microbiol Rev* 26(2):125-139.
12. Urakawa H, Kita-Tsukamoto K, & Ohwada K (1999) Reassessment of the taxonomic position of *Vibrio iliopiscarius* (Onarheim et al. 1994) and proposal for *Photobacterium iliopiscarium* comb. nov. *Int J Syst Bacteriol* 49 Pt 1:257-260.
13. Olive DM & Bean P (1999) Principles and applications of methods for DNA-based typing of microbial organisms. *J Clin Microbiol* 37(6):1661-1669.
14. Roque A, Molina-Aja A, Bolan-Mejia C, & Gomez-Gil B (2001) In vitro susceptibility to 15 antibiotics of vibrios isolated from penaeid shrimps in Northwestern Mexico. *Int J Antimicrob Agents* 17(5):383-387.
15. Arias CR, Garay E, & Aznar R (1995) Nested PCR method for rapid and sensitive detection of *Vibrio vulnificus* in fish, sediments, and water. *Appl Environ Microbiol* 61(9):3476-3478.
16. Thompson JR, *et al.* (2004) Diversity and dynamics of a north Atlantic coastal *Vibrio* community. *Appl Environ Micbrobiol* 70(7):4103-4110 .

Dhaka University Institutional Repository

17. Grisez L, Chair M, Sorgeloos P, & Ollevier F (1996) Mode of infection and spread of *Vibrio* anguillarum in turbot Scophthalmus maximus larvae after oral challenge through live feed. *Diseases of Aquatic Organisms* 26(3):181-187 .

18. Sawabe T, *et al.* (2002) Fluorescent amplified fragment length polymorphism and repetitive extragenic palindrome-PCR fingerprinting reveal host-specific genetic diversity of *Vibrio halioticoli*-like strains isolated from the gut of Japanese abalone. *Appl Environ Microbiol* 68(8):4140-4144.

19. Vandenberghe J, *et al.* (1999) Vibrios associated with Litopenaeus vannamei larvae, postlarvae, broodstock, and hatchery probionts. *Appl Environ Microbiol* 65(6):2592-2597.

20. Tamplin M, Gauzens A, Huq A, Sack D, & Colwell R (1990) Attachment of *Vibrio cholerae* Serogroup O1 to Zooplankton and Phytoplankton of Bangladesh Waters. *Appl. Environ. Miobiol.* 56(6):1977-1980.

21. Heidelberg JF, *et al.* (2000) DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* 406:477-483.

22. Ohwada K, Tabor PS, & Colwell RR (1980) Species Composition and Barotolerance of Gut Microflora of Deep-Sea Benthic Macrofauna Collected at Various Depths in the Atlantic Ocean. *Appl Environ Microbiol* 40(4):746-755.

23. Tabor PS, Ohwada K, & Colwell RR (1981) Filterable marine bacteria found in the deep sea: distribution, taxonomy, and response to starvation. *Microbial Ecology* 7:67-83.

24. Boone DR, Johnson RL, & Liu Y (1989) Diffusion of the interspecies electron carriers H2 and formate in methanogenic ecosystems and its implications in the measurement of Km for H2 or formate uptake. *Appl Environ Micbrobiol* 55(7):1735.

25. Reysenbach AL, Wickham GS, & Pace NR (1994) Phylogenetic analysis of the hyperthermophilic pink filament community in Octopus Spring, Yellowstone National Park. *Appl Environ Microbiol* 60(6):2113-2119.

26. Olsen GJ, Matsuda H, Hagstrom R, & Overbeek R (1994) fastDNAmL: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. *Comput Appl Biosci* 10(1):41-48.

27. Lipp EK, Huq A, & Colwell RR (2002) Effects of global climate on infectious disease: the cholera model. *Clin Microbiol Rev* 15(4):757-770.

28. Thompson JD, Higgins DG, & Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22(22):4673-4680.

29. Page RDM (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 12:357-358.

30. Versalovic J, Koeuth T, & Lupski R (1991) Distribution of repetitive DNA sequences in eubacteria and application to finerpriting of bacterial enomes. *Nucleic acids research* 19(24):6823.

31. Rivera IG, *et al.* (1995) Enterobacterial repetitive intergenic consensus sequences and the PCR to generate fingerprints of genomic DNAs from *Vibrio cholerae* O1, O139, and non-O1 strains. *Appl Environ Microbiol* 61(8):2898-2904.

32. Ewing B & Greene P (1998) Base-Calling of Automated Sequencer Traces Using Phred. II. Error Probabilities. *Genome Research* 8:186-194.

33. Gordon D, Abajian C, & Green P (1998) Consed: A Graphical Tool for Sequence Finishing. *Genome Research* 1998(8).

Dhaka University Institutional Repository

34. Han CS & Chain P (2006) Finishing repeat regions automatically with Dupfinisher. *International Conference on Bioinformatics and Computational Biology*, eds Arabnia HR & Valafar H (CSREA Press), pp 141-146.

35. Goryshin IY & Reznikoff WS (1998) Tn5 in vitro transposition. *J Biol Chem* 273(13):7367-7374.

36. Aziz R, *et al.* (2008) The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9(1):75.

37. Kurtz S, *et al.* (2004) Versatile and open software for comparing large genomes. *Genome Biol* 5(2):R12.

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

39. Konstantinidis KT & Tiedje JM (2005) Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci U S A* 102(7):2567-2572.

40. Zillig W, *et al.* (1992) RNA polymerases and transcription in archaebacteria. *Biochem Soc Symp* 58:79-88.

41. Condon C, Liveris D, Squires C, Schwartz I, & Squires CL (1995) rRNA operon multiplicity in *Escherichia coli* and the physiological implications of rrn inactivation. *J Bacteriol* 177(14):4152-4156.

42. Pisabarro A, Correia A, & Martin JF (1998) Characterization of the rrnB operon of the plant pathogen Rhodococcus fascians and targeted integrations of exogenous genes at rrn loci. *Appl Environ Microbiol* 64(4):1276-1282.

43. Makino K, *et al.* (2003) Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism distinct from that of *V. cholerae*. *The Lancet* 361(9359):743-749.

44. Kim YK & McCarter LL (2000) Analysis of the polar flagellar gene system of Vibrio parahaemolyticus. *J Bacteriol* 182(13):3693-3704.

45. Stewart BJ & McCarter LL (2003) Lateral flagellar gene system of Vibrio parahaemolyticus. *J Bacteriol* 185(15):4508-4518.

46. Konstantinidis KT, Ramette A, & Tiedje JM (2006) Toward a more robust assessment of intraspecies diversity, using fewer genetic markers. *Appl Environ Microbiol* 72(11):7286-7293.

47. Konstantinidis KT & Tiedje JM (2005) Towards a genome-based taxonomy for prokaryotes. *J Bacteriol* 187(18):6258-6264.

48. Markert S, *et al.* (2007) Physiological proteomics of the uncultured endosymbiont of Riftia pachyptila. *Science* 315(5809):247-250.

49. Karl DM, *et al.* (1988) A Microbiological Study of Guaymas Basin High-Temperature Hydrothermal Vents. *Deep-Sea Res* 35(5):777-791 .

50. Surette MG & Bassler BL (1998) Quorum sensing in *Escherichia coli* and *Salmonella typhimurium*. *Proc Natl Acad Sci U S A* 95(12):7046-7050.

51. Surette MG, Miller MB, & Bassler BL (1999) Quorum sensing in *Escherichia coli*, *Salmonella typhimurium*, and *Vibrio harveyi*: a new family of genes responsible for autoinducer production. *Proc Natl Acad Sci U S A* 96(4):1639-1644.

52. Higgins DA, *et al.* (2007) The major *Vibrio cholerae* autoinducer and its role in virulence factor production. *Nature* 450(7171):883-886.

53. Patel CN, *et al.* (2006) Polyamines are essential for the formation of plague biofilm. *J Bacteriol* 188(7):2355-2363.

Dhaka University Institutional Repository

54. Tabor CW & Tabor H (1985) Polyamines in microorganisms. *Microbiol Rev* 49(1):81-99.

55. Karatan E, Duncan TR, & Watnick PI (2005) NspS, a predicted polyamine sensor, mediates activation of *Vibrio cholerae* biofilm formation by norspermidine. *J Bacteriol* 187(21):7434-7443.

56. Hamana K & Nakagawa Y (2001) Polyamine distribution profiles in newly validated genera and species within the Flavobacterium-Flexibacter-Cytophaga-Sphingobacterium complex. *Microbios* 106 Suppl 2:105-116.

57. Hamana K & Nakagawa Y (2001) Polyamine distribution profiles in the eighteen genera phylogeneticaiiy located within the Flavobacterium-Flexibacter-Cytophaga complex. *Microbios* 106(413):7-17.

58. Hosoya R, Hamana K, Niitsu M, & Itoh T (2004) Polyamine analysis for chemotaxonomy of thermophilic eubacteria: Polyamine distribution profiles within the orders Aquificales, Thermotogales, Thermodesulfobacteriales, Thermales, Thermoanaerobacteriales, Clostridiales and Bacillales. *J Gen Appl Microbiol* 50(5):271-287.

59. Bomchil N, Watnick P, & Kolter R (2003) Identification and characterization of a *Vibrio cholerae* gene, mbaA, involved in maintenance of biofilm architecture. *J Bacterial* 185(4):1384-1390.

60. Niba ET, Naka Y, Nagase M, Mori H, & Kitakawa M (2007) A genome-wide approach to identify the genes involved in biofilm formation in E. coli. *DNA Res* 14(6):237-246.

61. Reeser RJ, Medler RT, Billington SJ, Jost BH, & Joens LA (2007) Characterization of Campylobacter jejuni biofilms under defined growth conditions. *Appl Environ Microbiol* 73(6):1908-1913.

62. Joshua GW, Guthrie-Irons C, Karlyshev AV, & Wren BW (2006) Biofilm formation in Campylobacter jejuni. *Microbiology* 152(Pt 2):387-396.

63. Kalmokoff M, *et al.* (2006) Proteomic analysis of Campylobacter jejuni 11168 biofilms reveals a role for the motility complex in biofilm formation. *J Bacteriol* 188(12):4312-4320.

64. Yip ES, Grublesky BT, Hussa EA, & Visick KL (2005) A novel, conserved cluster of genes promotes symbiotic colonization and sigma-dependent biofilm formation by *Vibrio fischeri*. *Mol Microbiol* 57(5):1485-1498.

65. Yildiz FH & Schoolnik GK (1999) *Vibrio cholerae* O1 El Tor: identification of a gene cluster required for the rugose colony type, exopolysaccharide production, chlorine resistance, and biofilm formation. *Proc Natl Acad Sci U S A* 96(7):4028-4033.

66. Walker GC (1996) The SOS response of Escherichia coli. *Escherichia coli and Salmonella: cellular and molecular biology*, ed Neidhardt FC (ASM Press, Washington, D.C.), pp 1400-1416.

67. Yoshida H, *et al.* (2002) The ribosome modulation factor (RMF) binding site on the 100S ribosome of Escherichia coli. *J Biochem* 132(6):983-989.

68. Jenal U & Malone J (2006) Mechanisms of cyclic-di-GMP signaling in bacteria. *Annu Rev Genet* 40:385-407.

69. Dick GJ, Lee YE, & Tebo BM (2006) Manganese(II)-oxidizing Bacillus spores in Guaymas Basin hydrothermal sediments and plumes. *Appl Environ Microbiol* 72(5):3184-3190.

70. Webb SM, Dick GJ, Bargar JR, & Tebo BM (2005) Evidence for the presence of Mn(III) intermediates in the bacterial oxidation of Mn(II). *Proc Natl Acad Sci U S A* 102(15):5558-5563.

71. Edmonds HN, *et al.* (2003) Discovery of abundant hydrothermal venting on the ultraslow-spreading Gakkel ridge in the Arctic Ocean. *Nature* 421(6920):252-256.

72. Ridge JP, *et al.* (2007) A multicopper oxidase is essential for manganese oxidation and laccase-like activity in Pedomicrobium sp. ACM 3067. *Environ Microbiol* 9(4):944-953.

Dhaka University Institutional Repository

73. Campanaro S, *et al.* (2005) Laterally transferred elements and high pressure adaptation in Photobacterium profundum strains. *BMC Genomics* 6:122.

74. Macdonald AG & Cossins AR (1985) The theory of homeoviscous adaptation of membranes applied to deep-sea animals. *Symp Soc Exp Biol* 39:301-322.

75. Hazell SL & Graham DY (1990) Unsaturated fatty acids and viability of Helicobacter (Campylobacter) pylori. *J Clin Microbiol* 28(5):1060-1061.

76. Vezzi A, *et al.* (2005) Life at depth: Photobacterium profundum genome sequence and expression analysis. *Science* 307(5714):1459-1461.

77. Nachln L, Nannmark U, & Nystrom T (2005) Differential roles of the universal stress proteins of Escherichia coli in oxidative stress resistance, adhesion, and motility. *J Bacteriol* 187(18):6265-6272.

78. Campbell BJ, *et al.* (2009) Adaptations to submarine hydrothermal environments exemplified by the genome of Nautilia profundicola. *PLoS Genet* 5(2):e1000362.

79. Huber KE & Waldor MK (2002) Filamentous phage integration requires the host recombinases XerC and XerD. *Nature* 417(6889):656-659.

80. Hedlund BP & Staley JT (2001) *Vibrio cyclotrophicus* sp. nov., a polycyclic aromatic hydrocarbon (PAH)-degrading marine bacterium. *Int J Syst Evol Microbiol* 51(Pt 1):61-66.

81. Yahr TL (2006) A critical new pathway for toxin secretion? *N Engl J Med* 355(11):1171-1172.

82. Bingle LE, Bailey CM, & Pallen MJ (2008) Type VI secretion: a beginner's guide. *Curr Opin Microbiol* 11(1):3-8.

83. Cascales E (2008) The type VI secretion toolkit. *EMBO Rep* 9(8):735-741.

84. Filloux A, Hachani A, & Bleves S (2008) The bacterial type VI secretion machine: yet another player for protein transport across membranes. *Microbiology* 154(Pt 6):1570-1583.

85. Weber B, Hasic M, Chen C, Wai SN, & Milton DL (2009) Type VI secretion modulates quorum sensing and stress response in Vibrio anguiliarum. *Environ Microbiol* .

86. Planet PJ, Kachlany SC, Fine DH, DeSalle R, & Figurski DH (2003) The Widespread Colonization Island of Actinobacillus actinomycetemcomitans. *Nat Genet* 34(2):193-198.

87. Iida T, *et al.* (1998) Close proximity of the tdh, trh and ure genes on the chromosome of *Vibrio parahaemolyticus*. *Microbiology* 144 (Pt 9):2517-2523.

88. McCarthy SA, DePaola A, Cook DW, Kaysner CA, & Hill WE (1999) Evaluation of alkaline phosphatase- and digoxigenin-labelled probes for detection of the thermolabile hemolysin (tih) gene of *Vibrio parahaemolyticus*. *Lett Appl Microbiol* 28(1):66-70.

89. Waldor MK & Mekalanos JJ (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* 272:1910-1914.

90. Zhang D, Rajanna C, Sun W, & Karaolis DK (2003) Analysis of the Vibrio pathogenicity island-encoded Mop protein suggests a pleiotropic role in the virulence of epidemic *Vibrio cholerae*. *FEMS Microbiol Lett* 225(2):311-318.

91. Xie ZY, Hu CQ, Chen C, Zhang LP, & Ren CH (2005) Investigation of seven Vibrio virulence genes among *Vibrio alginolyticus* and *Vibrio parahaemolyticus* strains from the coastal mariculture systems in Guangdong, China. *Lett Appl Microbiol* 41(2):202-207.

Dhaka University Institutional Repository

92. Nishibuchi M, Janda JM, & Ezaki T (1996) The thermostable direct hemolysin gene (tdh) of *Vibrio hollisae* is dissimilar in prevalence to and phyiogeneticaliy distant from the tdh genes of other vibrios: implications in the horizontal transfer of the tdh gene. *Microbiol Immunol* 40(1):59-65.

93. Boyd EF, Moyer KE, Shi L, & Waldor MK (2000) Infectious CTXPhi 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.

94. Hentschel U, Steinert M, & Hacker J (2000) Common molecular mechanisms of symbiosis and pathogenesis. *Trends microbiol* 8(5):226-231.

95. Faruque SM & Nair GB (2002) Molecular ecology of toxigenic *Vibrio cholerae*. *Microbiol Immunol* 46(2):59-66.

96. Lin Z, Kumagai K, Baba K, Mekalanos JJ, & Nishibuchi M (1993) *Vibrio parahaemolyticus* has a homolog of the *Vibrio cholerae* toxRS operon that mediates environmentally induced regulation of the thermostable direct hemolysin gene. *J Bacteriol* 175(12):3844-3855.

97. Reich KA & Schoolnik GK (1994) The light organ symbiont *Vibrio fischeri* possesses a homolog of the *Vibrio cholerae* transmembrane transcriptional activator ToxR. *J Bacteriol* 176(10):3085.

98. Lee SE, *et al.* (2000) *Vibrio vulnificus* has the transmembrane transcription activator ToxRS stimulating the expression of the hemolysin gene *vvhA*. *J of Bacteriol* 182(12):3405.

99. Wang SY, Lauritz J, Jass J, & Milton DL (2002) A ToxR homolog from *Vibrio anguillarum* serotype O1 regulates its own production, bile resistance, and biofilm formation. *J Bacteriol* 184(6):1630.

100. Crawford JA, Krukonis ES, & DiRita VJ (2003) Membrane localization of the ToxR winged-helix domain is required for TcpP-mediated virulence gene activation in *Vibrio cholerae*. *Mol Microbiol* 47(5):1459-1473.

101. Murphy RA & Boyd EF (2008) Three pathogenicity islands of Vibrio cholerae can excise from the chromosome and form circular intermediates. *J Bacteriol* 190(2):636-647.

102. Smith B & Oliver JD (2006) In situ and in vitro gene expression by *Vibrio vulnificus* during entry into, persistence within, and resuscitation from the viable but nonculturable state. *Appl Environ Microbiol* 72(2):1445-1451.

103. Kirn TJ, Jude BA, & Taylor RK (2005) A colonization factor links *Vibrio cholerae* environmental survival and human infection. *Nature* 438(7069):863-866.

104. Zampini M, *et al.* (2005) *Vibrio cholerae* persistence in aquatic environments and colonization of intestinal cells: involvement of a common adhesion mechanism. *FEMS Microbiol Lett* 244(2):267-273.

105. Singleton FL, Attwell R, Jangi S, & Colwell RR (1982) Effects of temperature and salinity on *Vibrio cholerae* growth. *Appl Environ Microbiol* 44(5):1047-1058.

106. Jump A & Tuttle JH (1981) Characterization of the thiosulfate-oxidizing bacteria isolated from marine environments. *Annu. Gen. Meet. ASM*, (ASM), p 187.

107. Boucher Y & Doolittle WF (2000) The role of lateral gene transfer in the evolution of isoprenoid biosynthesis pathways. *Mol Microbiol* 37(4):703-716.

108. Singleton FL, Attwell R, Jangi S, & Colwell RR (1982) Effects of temperature and salinity on *Vibrio cholerae* growth. *Appl Environ Microbiol* 44(5):1047.

109. Colwell R, Kaper J, & Joseph S (1977) *Vibrio cholerae, Vibrio parahaemolyticus*, and other Vibrios: occurrence and distribution in Chesapeake Bay. *Science* 198:394-396.

Dhaka University Institutional Repository

110. Singleton FL, Atwell RW, Jangi MS, & Colwell RR (1982) Influence of salinity and organic nutrient concentration on survival and growth of *Vibrio cholerae* in aquatic microcosms. *Appl Environ Micbrobiol* 43(5):1080-1085.

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

112. Haihong H, *et al.* (Genome plasticity of *Vibrio parahaemolyticus*: microevolution of the'pandemic group'. *BMC Genomics* 9.

113. Venkateswaran K, Dollhopf ME, Aller R, Stackebrandt E, & Nealson KH (1998) *Shewanella amazonensis* sp. nov., a novel metal-reducing facultative anaerobe from Amazonian shelf muds. *Int J Sys Evol Microbiol* 48(3):965.

114. Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* 27(2):573-580.

115. Nishibuchi M (1996) Origin of the pathogenic vibrios in the environment: inference from the studies on the molecular genetics of *Vibrio cholerae* and *Vibrio parahaemolyticus*. *Nippon Saikingaku Zasshi* 51(3):823-832.

116. Sechi L, Dupre I, Deriu A, Fadda G, & Zanetti S (2000) Distribution of *Vibrio cholerae* virulence genes among different Vibrio species isolated in Sardinia, Italy. *J Appl Microbiol* 88(3):475-481.

117. Huq A, *et al.* (1983) Ecological relationship between *Vibrio cholerae* and planktonic copepods. *Appl. Environ. Microbiol.* 45:275-283.

118. Colwell R & Spira W (1992) The ecology of *Vibrio cholerae*. Cholera, eds Barua D & Greenough Wi (Plenum Medical Book Co., New York), pp 107-127.

Dhaka University Institutional Repository

CHAPTER 8:

GENOMIC DIVERSITY OF 2010 HAITIAN CHOLERA OUTBREAK STRAINS

Dhaka University Institutional Repository

Abstract:

The millions of deaths from cholera over the past 200 years, coupled with the morbidity and mortality of cholera in Haiti since October, 2010, are grim reminders that *Vibrio cholerae*, the etiologic agent of cholera, remains a scourge. We report the isolation of both *V. cholerae* O1 and non-O1/O139 early in the Haiti cholera epidemic from samples collected from victims in 18 towns across eight Arrondissements of Haiti. The results showed two distinct populations of *Vibrio cholerae* coexisted in Haiti early in the epidemic. As non-O1/O139 *V. cholerae* was the sole pathogen isolated from 21% of the clinical specimens, its role in this epidemic, either alone or in concert with *V. cholerae* O1, cannot be dismissed. A genomic approach was used to examine similarities and differences among the Haitian *V. cholerae* O1 and *V. cholerae* non-O1/O139 strains. A total of 47 *V. cholerae* O1 and 29 *V. cholerae* non-O1/O139 isolates from patients and the environment were sequenced. Comparative genome analyses of the 76 genomes and eight reference strains of *V. cholerae* isolated in concurrent epidemics outside Haiti and 27 *V. cholerae* genomes available in the public database demonstrated substantial diversity of *Vibrio cholerae* and ongoing flux within its genome.

Introduction:

Cholera is a disease endemic in many developing countries and the causative agent, *Vibrio cholerae,* is autochthonous to estuaries and river systems world-wide (1). The recent outbreak of cholera in Haiti is a stark reminder of how rapidly a cholera epidemic can emerge and spread through a population. On January 12, 2010, a devastating 7.0 M_w earthquake struck Haiti and, in the ensuing 12 days approximately 52 aftershocks with a magnitude of 4.5 M_w or greater were recorded. It is estimated that more than 250,000 people died, approximately 300,000 were injured and one million rendered homeless. A very hot summer season followed. Amidst the clearing of rubble and rebuilding of infrastructure, the first case of cholera was diagnosed on October 21, 2010, and then the disease quickly erupted throughout Haiti. On November 5, 2010, the full force of Hurricane Tomas skirted the island causing torrential rains and severe

Dhaka University Institutional Repository

flooding. Cholera, not officially reported in Haiti since 1960 (2), erupted in epidemic proportions (3) and in just one year after the initial cases, cholera has claimed more than 6500 lives and sickened ca. 500,000 (4). The series of events of 2010 clearly created a "perfect storm" for cholera, with assignment of attribution remaining controversial (4).

Identical pulsed field gel electrophoresis (PFGE) patterns for 13 bacterial isolates recovered from Haitian cholera patients led to a conclusion that the Haitian isolates derived from a strain similar to *V. cholerae* isolated in South Asia (3). From a limited comparison of genomes, Chin *et al.* concluded the Haitian cholera epidemic was most likely a result of an introduction of a strain from a distant geographic location and derived from human activity (5). A UN Independent Panel of Experts in a report on the 2010 Cholera Outbreak in Haiti concluded the source of the Haiti cholera outbreak was due to contamination of water upstream in the Artibonite River with a pathogenic strain of a South Asian type of *V. cholerae* (6). Another report (7) employing whole-genome sequence typing of Nepalese isolates and comparing these with a few previously sequenced *V. cholerae* strains, interpreted the results as Nepal being the likely origin of the Haitian outbreak. A rebuttal, however, was published which pointed out that South Asia, not Nepal, might be the origin of the Haitian cholera outbreak strains, stating correctly that proper attribution cannot be accomplished by an exclusionary approach (8). Notably, none of the Nepalese soldiers had been recorded to be ill during the outbreak in Haiti (4), therefore involvement of Nepalese peacekeepers remains controversial (9). A further study employing whole-genome phylogeny and core genome single-nucleotide polymorphisms (8) showed the Haiti outbreak strain to be genetically related to strains originating in India and Cameroon. Like earlier studies, however, the authors concluded that a definitive genetic origin for the outbreak in Haiti must remain speculative (8).

Although the molecular and genomic data in the previous reports cited above have been interpreted as being consistent with *V. cholerae* being introduced to Haiti by outside visitors (5, 7, 8), a definitive statement of source attribution cannot yet be made. To assess source attribution, a qualified reference database (including recent strains from Nepal, India, Cameroon, and related *V. cholerae* strains from concurrent epidemics)

Dhaka University Institutional Repository

representative of global phylogenetic diversity as well as well-documented population diversity within the heterogeneous clade and its near neighbor comparators, is critical. In this study, the genomic diversity of Haitian isolates was examined and these data are presented in order to provide a broader perspective of the complexity of the species, *V. cholerae*, particularly of those isolates from Haiti and, perhaps, contribute to the understanding of the origin of the outbreak as either indigenous or introduced.

Methods:

Bacteriological Analysis. Traditional methods (39, 40) for isolating and identifying *V. cholerae* from environmental water samples were used for bacteriological analysis of all samples. Environmental concentrated water samples were added to an enrichment broth, Alkaline Pepton Water (APW) and incubated overnight. Following enrichment, aliquots were sub-cultured onto selective media, thiosulfate citrate bile salts sucrose (TCBS), for presumptive identification of *V. cholerae*. The identity of presumptive colonies is confirmed by specific biochemical tests designed to differentiate *V. cholerae* from other species and other members of the *Vibrionaceae*, and the *Enterobacteriaceae*. Confirmed *V. cholerae* strains are then serogrouped as O1 or O139 by simple slide agglutination using specific antisera. Strains of *V. cholerae* that do not agglutinate with specific O1 and O139 antisera are designated as non-O1/non-O139. Stock cultures of confirmed *V. cholerae* are maintained in nutrient agar containing 0.5%NaCl overlaid with oil or as glycerol stocks at -70°C for future analysis and reference.

Sequencing, assembly and annotation. Genomic DNA of twenty *V. cholerae* strains was subjected to next generation whole genome Illumina and hybrid Ilumina/454 sequencing and closure strategies as previously described (41). Libraries were constructed with target insert sizes of 3 kb and paired-end 100 bp. Hybrid and Illumina sequences were assembled using the Celera and Velvet assemblers, respectively (42, 43), and all chromosomes were manually annotated using the Manatee system

Dhaka University Institutional Repository

(http://manatee.sourceforge.net/). Accession numbers of the genomes sequenced for this study are listed in supplementary Table 8.1.

SNP-discovery. A nucmer-based (44) in-house bioinformatics pipeline that facilitates discovery of fine single nucleotide polymorphisms (SNPs) was used to examine genetic relatedness among the analyzed, completed, and draft *V. cholerae* O1 genomes. SNP calls were manually curated for false positive calls that originated from misalignment, paralogous genes, and insufficient genome coverage. A comprehensive all versus all BLASTN (45) search among the reference and query genomes of the sequence surrounding each SNP (41 nt) resulted in multiple ambiguous and paralogous hits in some of the query molecules and these too were excluded, as well as repetitive regions prone to nucmer misalignment, were excluded from further phylogenetic analysis. Genomic regions with identities over 98% on chromosomes I and chromosome II were excluded, and regions that were not conserved among the strains were excluded from further analysis.

Identification and annotation of genomic islands. We defined genomic islands (GIs) as a continuous array of five or more CDSs that were discontinuously distributed among genomes of test strains. Correct transfer or insertion of GIs was readily differentiated from deletion event by comparing genome-based phylogenetic tree and full matrices showing pairwise detection of orthologous genes between test strains. Identified GIs were designated and annotated using the BLASTP search of its member CDSs against GenBank NR database as described elsewhere (46).

Phylogenomics based on genome sequences. Orthologous regions to *V. cholerae* N16961 were identified by comparisons based on similarity and were used to generate phylogenetic trees. The set of orthologous regions for each CDS of a reference genome was identified according to nucleotide similarity and aligned using CLUSTALW2. The resultant multiple alignments were concatenated to form genome scale alignments, which were then used to generate the neighbor-joining (47) phylogenetic trees.

Dhaka University Institutional Repository

SL #	Strain	Serotype	Geographical Origin	Source	Year of Isolation	GenBank Accession
1	CP1048	*V. cholerae* O1 OGET	Bangladesh	C	2010	SRA037369
2	CP1042	*V. cholerae* O1 OGET	Thailand	C	2010	SRA037382
3	CP1041	*V. cholerae* O1 OGET	Zambia	C	2004	SRA037374
4	CP1038	*V. cholerae* O1 OGET	Zimbabwe	C	2009	SRA037375
5	CP1032	*V. cholerae* O1 OGET	Mexico	C	1991	SRA037376
6	HC-06A1	*V. cholerae* O1 OGET	Saint-Marc, Artibonite , Haiti	C	2010	SRA037376
7	HC-7A1	*V. cholerae* O1 OGET	Dessalines, Artibonite , Haiti	C	2010	SRA035959
8	HC-17A1	*V. cholerae* O1 OGET	Gonaïves, Artibonite , Haiti	C	2010	SRA046087
9	HC-19A1	*V. cholerae* O1 OGET	Saint-Marc, Artibonite , Haiti	C	2010	SRA035994
10	HC-21A1	*V. cholerae* O1 OGET	Saint-Marc, Artibonite , Haiti	C	2010	SRA035997
11	HC-22A1	*V. cholerae* O1 OGET	Saint-Marc, Artibonite , Haiti	C	2010	SRA035991
12	HC-23A1	*V. cholerae* O1 OGET	Bassin Bleu, Nord-Ouest, Haiti	C	2010	SRA030719
13	HC-28A1	*V. cholerae* O1 OGET	Gonaïves, Ouest, Haiti	C	2010	SRA035987
14	HC-32A1	*V. cholerae* O1 OGET	Port-au-Prince, Ouest , Haiti	C	2010	SRA035998
15	HC-33A2	*V. cholerae* O1 OGET	Port-au-Prince, Ouest , Haiti	C	2010	SRA035995
16	HC-38A1	*V. cholerae* O1 OGET	Croix-des-Bouquets, Ouest, Haiti	C	2010	SRA030739
17	HC-40A1	*V. cholerae* O1 OGET	Arcahaie, Ouest, Haiti	C	2010	SRA030718
18	HC-43A1	*V. cholerae* O1 OGET	Port-au-Prince, Ouest, Haiti	C	2010	SRA030737
19	HC-48A1	*V. cholerae* O1 OGET	Port-au-Prince, Ouest, Haiti	C	2010	SRA030742
20	HC-48B2	*V. cholerae* O1 OGET	Port-au-Prince, Ouest, Haiti	C	2010	SRA035993
21	HC-49A2	*V. cholerae* O1 OGET	Port-au-Prince, Ouest, Haiti	C	2010	SRA030717
22	HC-50A2	*V. cholerae* O1 OGET	Port-au-Prince, Ouest, Haiti	C	2010	SRA046626
23	HC-61A1	*V. cholerae* O1 OGET	Croix-des-Bouquets, Ouest, Haiti	C	2010	SRA030716
24	HC-62A1	*V. cholerae* O1 OGET	Port-au-Prince, Ouest, Haiti	C	2010	SRA046089
25	HC-64A1	*V. cholerae* O1 OGET	Port-au-Prince, Ouest, Haiti	C	2010	SRA035979
26	HC-65A1	*V. cholerae* O1 OGET	Port-au-Prince, Ouest, Haiti	C	2010	SRA035971
27	HC-67A1	*V. cholerae* O1 OGET	Port-au-Prince, Ouest, Haiti	C	2010	SRA035983
28	HC-68A1	*V. cholerae* O1 OGET	Port-au-Prince, Ouest, Haiti	C	2010	SRA035984
29	HC-70A1	*V. cholerae* O1 OGET	Montrouis, Ouest, Haiti	C	2010	SRA037382
30	HC-71A1	*V. cholerae* O1 OGET	Port-au-Prince, Ouest, Haiti	C	2010	SRA035981
31	HC-72A2	*V. cholerae* O1 OGET	Arcahaie, Ouest, Haiti	C	2010	SRA036000
32	HC-77A1	*V. cholerae* O1 OGET	Gonaïves, Ouest, Haiti	C	2010	SRA046088
33	HC-80A1	*V. cholerae* O1 OGET	Port-au-Prince, Ouest, Haiti	C	2010	SRA035982
34	HC-81A1	*V. cholerae* O1 OGET	Port-au-Prince, Ouest, Haiti	C	2010	SRA035973
35	HCEL1786	*V. cholerae* O1 OGET	Haiti	C	2010	AELH01000000
36	HCEL1792	*V. cholerae* O1 OGET	Haiti	C	2010	AELJ01000000
37	HCEL1798	*V. cholerae* O1 OGET	Haiti	C	2010	AELI01000000
38	HCUF01	*V. cholerae* O1 OGET	Saint-Marc, Artibonite, Haiti	C	2010	SRA030713
39	HFU-02	*V. cholerae* O1 OGET	Saint-Marc, Artibonite, Haiti	C	2010	SRA030741
40	HC-1A2	*V. cholerae* Non-O1/O139	Saint-Marc, Artibonite , Haiti	C	2010	SRA030738
41	HC-2A1	*V. cholerae* Non-O1/O139	Saint-Marc, Artibonite, Haiti	C	2010	SRA030736
42	HC-36A1	*V. cholerae* Non-O1/O139	Port-au-Prince, Ouest, Haiti	C	2010	SRA035976
43	HC-41B1	*V. cholerae* Non-O1/O139	Haiti, Haiti	C	2010	SRA046401
44	HC-55C2	*V. cholerae* Non-O1/O139	Port-au-Prince, Ouest, Haiti	C	2010	SRA046086
45	HC-59A1	*V. cholerae* Non-O1/O139	Port-au-Prince, Ouest, Haiti	C	2010	SRA045689

Dhaka University Institutional Repository

SL #	Strain	Serotype	Geographical Origin	Source	Year of Isolation	GenBank Accession
46	HC-60A1	*V. cholerae* Non-O1/O139	Port-au-Prince, Ouest, Haiti	C	2010	SRA046381
47	HC-61A2	*V. cholerae* Non-O1/O139	Croix-des-Bouquets, Ouest, Haiti	C	2010	SRA045873
48	HC-78A1	*V. cholerae* Non-O1/O139	Croix-des-Bouquets, Ouest , Haiti	C	2010	SRA035966
49	HE-09	*V. cholerae* Non-O1/O139	Cange, Centre, Haiti	E	2010	SRA030714
50	HE39	*V. cholerae* Non-O1/O139	Cange, Centre, Haiti	E	2010	SRA030720
51	HE40	*V. cholerae* Non-O1/O139	Hospital waste, Cange, Haiti	E	2010	SRA037283
52	HE46	*V. cholerae* Non-O1/O139	Gray water, Cange, Haiti	E	2010	SRA046084
53	HE48	*V. cholerae* Non-O1/O139	Cange, Centre, Haiti	E	2010	SRA030715
54	CIRS101	*V. cholerae* O1 INET	Bangladesh	C	2002	ACVW00000000
55	MJ-1236	*V. cholerae* O1 INET	Matlab, Bangladesh	C	1994	CP001485-6
56	B33	*V. cholerae* O1 OGET	Beira, Mozambique	C	2004	ACHZ00000000
57	MO10	*V. cholerae* O139	Madras,India	C	1992	AAKF03000000
58	RC9	*V. cholerae* O1 OGET	Kenya	C	1985	ACHX00000000
59	N16961	*V. cholerae* O1 INET	Bangladesh	C	1975	AE003852-3
60	INDRE 91/1	*V. cholerae* O1 INET	Mexico	C	1991	ADAK00000000
61	NCTC 8457	*V. cholerae* O1 INET	Saudi Arabia	C	1910	AAWD01000000
62	BX330286	*V. cholerae* O1 INET	Australia	E	1986	ACIA00000000
63	M66-2	*V. cholerae* O1	Makassar, Indonesia	C	1937	CP001233-4
64	MAK757	*V. cholerae* O1 OGET	Celebes Islands	C	1937	AAUS00000000
65	2740-80	*V. cholerae* O1 INET	US Gulf Coast	E	1980	AAUT01000000
66	V52	*V. cholerae* O37	Sudan	C	1968	AAKJ02000000
67	O395	*V. cholerae* O1 OGCL	India	C	1965	CP000626-7
68	RC27	*V. cholerae* O1 ET	Indonesia	C	1991	ADAI00000000
69	12129(1)	*V. cholerae* O1 INET	Australia	E	1985	ACFQ00000000
70	AM-19226	*V. cholerae* O39	Bangladesh	C	2001	AATY01000000
71	MZO-3	*V. cholerae* O37	Bangladesh	C	2001	AAUU01000000
72	TMA21	*V. cholerae* Non-O1/O139	Brazil	E	1982	ACHY00000000
73	1587	*V. cholerae* O12	Lima, Peru	C	1994	AAUR01000000
74	623-39	*V. cholerae* Non-O1/O139	Bangladesh	E	2002	AAWG00000000
75	MZO-2	*V. cholerae* O14	Bangladesh	C	2001	AAWF01000000
76	V51	*V. cholerae* O141	USA	C	1987	AAKI02000000
77	CT 5369-93	*V. cholerae* Non-O1/O139	Brazil	S	1993	ADAL00000000
78	RC385	*V. cholerae* O135	Cheasapeake bay, USA	E	1998	AAKH02000000
79	TM11079-80	*V. cholerae* O1 OGET	Brazil	S	1980	ACHW00000000
80	VL426	*V. cholerae* Non-O1/O139	Maidstone, Kent, UK	E	Unknown	ACHV00000000

Table 8.1: List of strains used in this study. C; Clinical, E; Environmental S; Sewage

Multilocus variable number of tandem repeats (MLVA). Genomic DNA was prepared from LB agar cultures using the Prepman Ultra kit (Applied Biosystems). Five loci for MLVA were amplified using the primers and PCR conditions described in

Dhaka University Institutional Repository

previous studies (48-50). For strains not in our culture collection, VNTR loci were extracted from published sequences for comparison. Sequence data for each isolate were added to a group of known sequences that were aligned simultaneously and edited using jPHYDIT (51).

Principal component analysis (PCA). PCA analysis was conducted based on the average nucleotide identity (ANI) (52) matrix of *V. cholerae* strain samples using R. To determine ANI, tetra nucleotide frequency and genetic distance between strains (53), a reciprocal best match BLASTN analysis (ANIb) was performed for each genome using JSpecies (54), and divergences were determined to reflect genomic distance between genomes. The divergence matrix provides a measure of genetic distance between *V. cholerae* strains of all samples and can be used evaluate the relatedness between each sample.

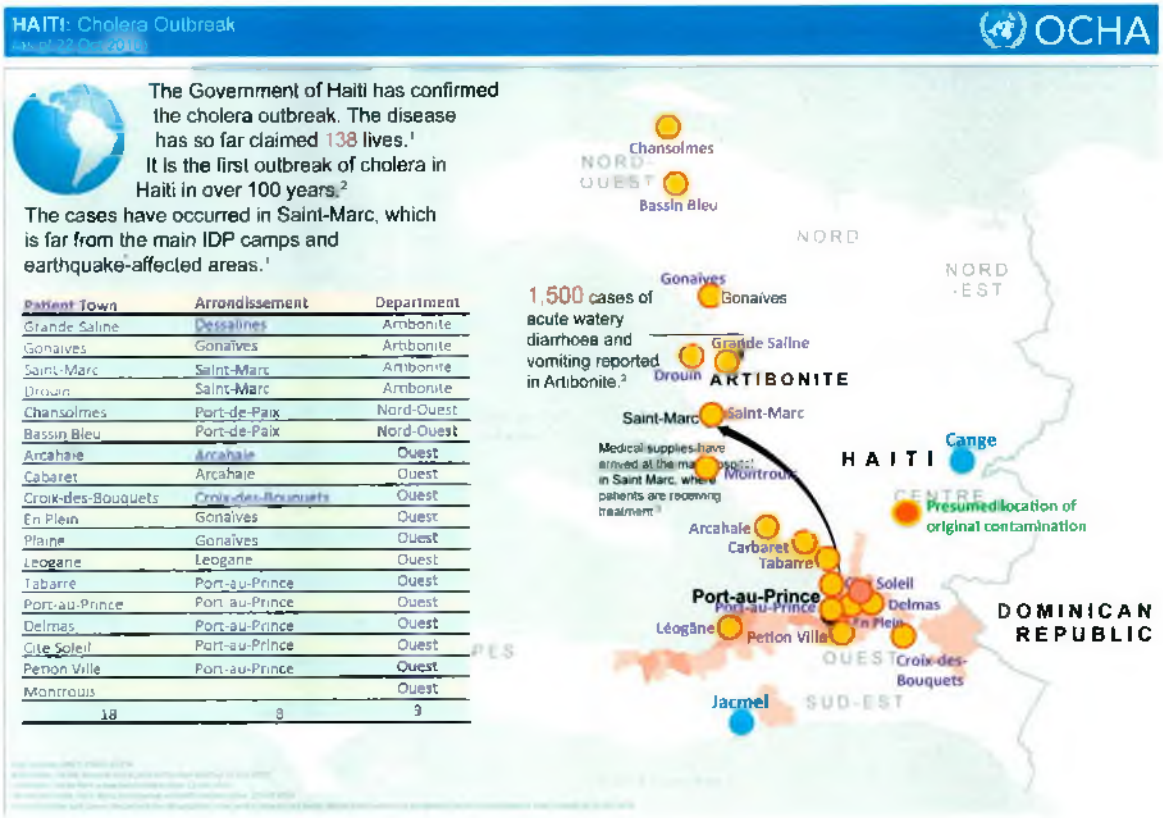
Whole genome mapping. Maps were generated, facilitating closure and assembly, thus allowing a detailed study of polymorphisms and their respective genome localization in the studied population. Whole genome maps were prepared by OpGen, Gaithersburg MD, USA. Following gentle lysis and dilution, high-molecular mass genomic DNA molecules were spread and immobilized on derivatized glass slides and digested with *Nhe*I. The DNA digests were stained with YOYO-1 fluorescent dye, and photographed using a fluorescent microscope interfaced with a digital camera. Automated image-analysis software located and sized fragments, and assembled fragments from multiple scans into whole-chromosome whole genome maps.

Results and Discussion:

Clinical samples from 81 patients suffering symptoms of cholera with varying severity were collected from 18 towns across eight Arrondissements of Haiti (Fig. 8.1, Table 8.1), spanning three weeks in November, 2010, very early in the outbreak. Bacteriological analysis identified *V. cholerae* O1 and *V. cholerae* non-O1/O139 as the sole pathogen in 48% (n=39) and 21% (n=17) of the samples, respectively. *V. cholerae*

Dhaka University Institutional Repository

O1 and Non-O1/O139 were co-cultured from six of the samples (7%). Isolation of *V. cholerae* Non-O1/O139, in addition to toxigenic *V. cholerae* O1, indicated that at least two disparate *V. cholerae* populations coexisted early in the outbreak. From the remaining 31% (n=25) of the clinical samples, other enteric pathogens were also cultured and subsequent analysis of a portion of the latter indicated the presence of opportunistic pathogens including *Aeromonas caviae, Aeromonas hydrophila,* and *Proteus* mirabilis (Clark *et al.* manuscript in preparation). Environmental samples were also collected form two different towns in Haiti concurrent with the epidemics; bacteriological analysis of these samples identified *V. cholerae* non-O1/O139 from each of those samples.



Patient Town	Arrondissement	Department
Grande Saline	Dessalines	Artibonite
Gonaives	Gonaives	Artibonite
Saint-Marc	Saint-Marc	Artibonite
Drouin	Saint-Marc	Artibonite
Chansolmes	Port-de-Paix	Nord-Ouest
Bassin Bleu	Port-de-Paix	Nord-Ouest
Arcahaie	Arcahaie	Ouest
Cabaret	Arcahaie	Ouest
Croix-des-Bouquets	Croix-des-Bouquets	Ouest
En Plein	Gonaives	Ouest
Plaine	Gonaives	Ouest
Leogane	Leogane	Ouest
Tabarre	Port-au-Prince	Ouest
Port-au-Prince	Port-au-Prince	Ouest
Delmas	Port-au-Prince	Ouest
Cite Soleil	Port-au-Prince	Ouest
Petion Ville	Port-au-Prince	Ouest
Montrouis		Ouest
18	8	9

Figure 8.1: Source and distribution of isolates collected during the 2010 Haitian cholera outbreak. Yellow and Blue colored circles indicate sites where clinical and environmental strains were collected, respectively. The panel on the left shows distribution of strains with respect to different locations of sample collection in Haiti.

Dhaka University Institutional Repository

Multilocus variable-number tandem repeat (VNTR) analysis (MLVA). MLVA, a useful tool for tracking *V. cholerae* (10) and resolving distinct populations of clinical isolates from different geographic regions (11, 12) was used to examine the clinical and environmental *V. cholerae* O1 isolates from Haiti and other geographical locations (Table 8.2). Altered El Tor *V. cholerae* circulating in South East Asia has been reported to be a stable VNTR genotype over the years {Bangladesh 2001-2005 (12), India 2004-2007 (11)}, and our typing of the 9.3.6.n.n VNTR genotype supported the stability of that genotype. Haitian *V. cholerae* O1 strains, on the other hand, displayed a nearly identical VNTR genotype (Table 8.2), forming a tight clonal complex, as reported previously (13). Investigation of *V. cholerae* strains from Zimbabwe, Zambia, Mexico, and Thailand confirmed the Haitian *V. cholerae* strains contained a genetically distinct genotype (8.3.6.n.n) not observed previously and representing a new VNTR genotype (Table 8.2, Fig. 8.2). The MLVA data provide two valuable pieces of information. First, clonal transmission of VNTR alleles, or the limited VNTR diversity among the Haitian cholera outbreak strains, indicates either a possible single source of contamination, in agreement with other findings (13), or a single clone of *V. cholerae* responsible for the epidemic. Secondly, Haitian strains exhibit a genetically distinct MLVA profile. Environmental *V. cholerae* non-O1/O139 isolates demonstrated heterogeneous VNTR genotypes, with variability at all five loci (Table 8.2). Notable, however, was the observation that the clinical *V. cholerae* non-O1/O139 isolates revealed an identical VNTR genotype, akin to, but distinct from, the clinical *V. cholerae* O1 isolates.

Genomic relatedness of Haitian strains with reference *Vibrio cholerae* strains

Whole genome sequencing of 47 Haitian *V. cholerae* O1, 29 *V. cholerae* non-O1/O139, and eight reference strains of *V. cholerae* isolated outside Haiti in concurrent epidemics including Bangladesh (CP1048, 2010) and Thailand (CP1042, 2010) as well as older epidemics (Table 8.1) was performed for this study. Genome comparisons of these 76 genomes and an additional 27 *V. cholerae* genomes available in the public database were conducted to ascertain the degree of genomic diversity within *V. cholera* isolates from the same epidemics.

Dhaka University Institutional Repository

ID	Origin, Year	Source	Type	VC0147	VC0436-7*	VC1650	VCA0171	VCA0283
O395	India, 1965	C	O1 OGCL	7	4	3	24	14
N16961	Bangladesh, 1975	C	O1 INET	9	7	7	23	14
CIRS101	Bangladesh, 2002	C	O1 OGET	9	3	6	16	11
CP1032	Mexico, 1991	C	O1 OGET	8	7	8	13	17
CP1033	Mexico, 2000	C	O1 OGET	8	7	8	13	24
CP1038	Zimbabwe, 2009	C	O1 OGET	8	3	3	18	22
CP1041	Zambia, 2004	C	O1 OGET	8	7	6	9	10
CP1042	Thailand, 2010	C	O1 OGET	9	6	7	17	17
CP1048	Bangladesh, 2010	C	O1 OGET	9	3	6	22	17
AE-1786	Haiti, 2010	C	O1 OGET	8	3	6	14	20
AE-1792	Haiti, 2010	C	O1 OGET	8	3	6	14	N.D.
AE-1798	Haiti, 2010	C	O1 OGET	8	3	6	14	20
HC-7A1	Grande Saline, Haiti, 2010	C	O1 OGET	8	3	6	14	20
HC-20A2	Saint-Marc, Haiti, 2010	C	O1 OGET	8	3	6	14	20
HC-23A1	Bassin Bleu, Haiti, 2010	C	O1 OGET	8	3	6	14	20
HC-38A1	Croix-des-Bouquets, Haiti, 2010	C	O1 OGET	8	3	6	12	16
HC-40A1	Arcahaie, Haiti, 2010	C	O1 OGET	8	3	6	14	20
HC-43A1	Port-au-Prince, Haiti, 2010	C	O1 OGET	8	3	6	14	20
HC-48A1	Port-au-Prince, Haiti, 2010	C	O1 OGET	8	3	6	14	20
HC-49A2	Port-au-Prince, Haiti, 2010	C	O1 OGET	8	3	6	14	19
HC-61A1	Croix-des-Bouquets, Haiti, 2010	C	O1 OGET	8	3	6	14	20
HC-61A1	Croix-des-Bouquets, Haiti, 2010	C	O1 OGET	8	3	6	14	N.D.
HC-70A1	Montrouis, Haiti, 2010	C	O1 OGET	8	3	6	14	12
HC-81A1	Cite Soleil, Haiti, 2010	C	O1 OGET	8	3	6	14	20
HCUF01	Saint-Marc, Haiti, 2010	C	O1 OGET	8	3	6	14	20
HFU-02	Saint-Marc, Haiti, 2010	C	O1 OGET	8	3	6	14	19
HC-1A2	Saint-Marc, Haiti, 2010	C	Non O1/O139	12	8	7	16	NA
HC-2A1	Saint-Marc, Haiti, 2010	C	Non O1/O139	12	8	7	16	NA
HC-36A1	Tabarre, Haiti, 2010	C	Non-O1/O139	12	8	7	15	NA
HE-09	Cange, Haiti, 2010	E	Non O1/O139	11	3	5	13	NA
HE39	Cange, Haiti, 2010	E	Non O1/O139	8	3	2	13	NA
HE48	Cange, Haiti, 2010	E	Non O1/O139	7	6	6	6	NA

N.D.: Not determined, C; Clinical, E; Environmental, INET; Inaba El Tor, OGET; Ogawa El Tor, OGCL: Ogawa Classical; *: Intergenic; NA: Not Applicable

Table 8.2: VNTR genotypes of a subset of *V. cholerae* strains isolated from Haiti and other geographic locations. Allele number indicates the number of repeat units of each locus. The repeat units of the MLVA loci are: AACAGA at VC0147, GACCCTA at VC0436-7, GATAATCCA at VC1650, TGCTGT at VCA0171, and ACCAGA at VCA0283.

Dhaka University Institutional Repository

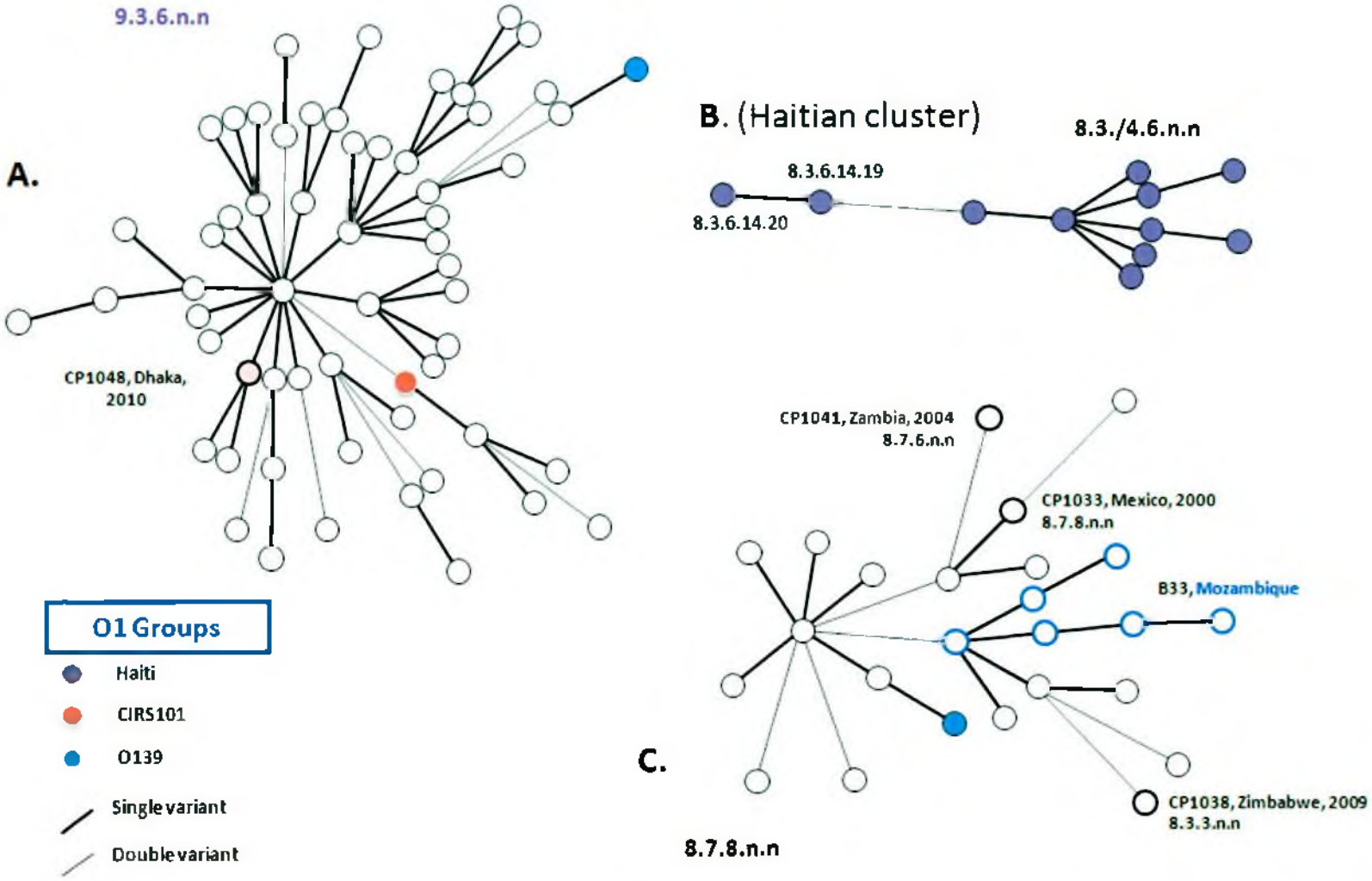
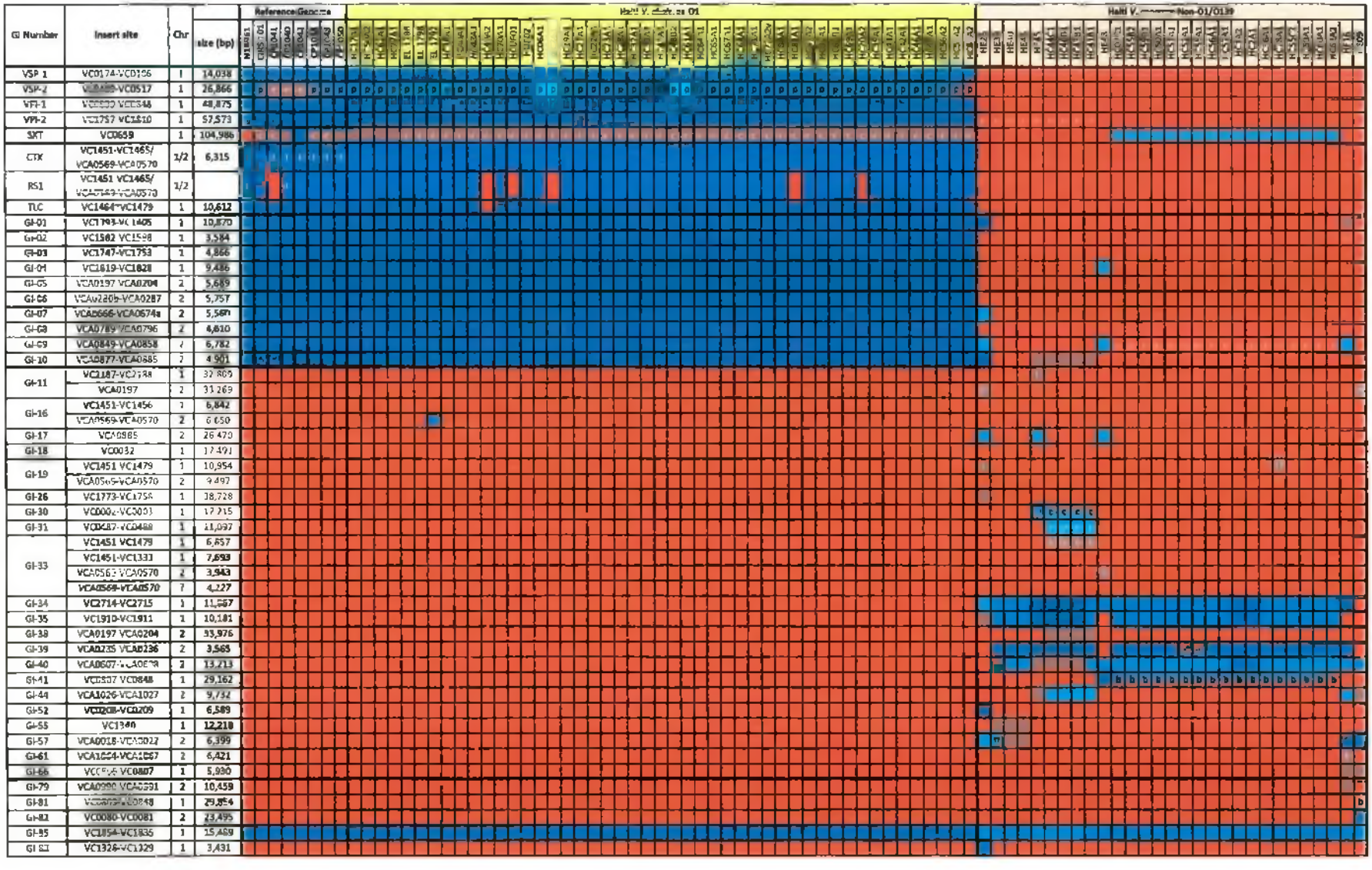


Figure 8.2: Minimum spanning tree based on MLVA clustered *V. cholerae* O1 strains into three distinct groups

Dhaka University Institutional Repository

GI Number	Insert site	Chr	size (bp)
VSP-1	VC0174-VC0196	1	14,038
VSP-2	VC0488-VC0517	1	26,866
VPI-1	VC0320-VC0348	1	48,875
VPI-2	VC1757-VC1810	1	57,573
SXT	VC0659	1	104,986
CTX	VC1451-VC1465/ VCA0569-VCA0570	1/2	6,315
RS1	VC1451-VC1465/ VCA0569-VCA0570	1/2	
TLC	VC1464-VC1479	1	10,612
GI-01	VC1793-VC1805	1	10,870
GI-02	VC1582-VC1598	1	3,584
GI-03	VC1747-VC1753	1	4,866
GI-04	VC1819-VC1828	1	9,486
GI-05	VCA0197-VCA0204	2	5,689
GI-06	VCA0280-VCA0287	2	5,757
GI-07	VCA0666-VCA0674	2	5,569
GI-08	VCA0789-VCA0796	2	4,810
GI-09	VCA0849-VCA0858	2	6,782
GI-10	VCA0877-VCA0885	2	4,901
GI-11	VC1187-VC1248	1	37,869
	VCA0197	2	31,269
GI-16	VC1451-VC1456	1	6,842
	VCA0569-VCA0570	2	6,050
GI-17	VCA0985	2	26,470
GI-18	VC0032	1	17,491
GI-19	VC1451-VC1479	1	10,954
	VCA0565-VCA0570	2	9,497
GI-26	VC1773-VC1758	1	38,728
GI-30	VC0002-VC0031	1	17,215
GI-31	VC0487-VC0488	1	11,097
GI-33	VC1451-VC1479	1	6,657
	VC1451-VC1333	1	7,693
	VCA0566-VCA0570	2	3,943
	VCA0569-VCA0570	2	4,227
GI-34	VC2714-VC2715	1	11,567
GI-35	VC1910-VC1911	1	10,181
GI-38	VCA0197-VCA0204	2	33,976
GI-39	VCA0235-VCA0236	2	3,565
GI-40	VCA0607-VCA0678	2	13,213
GI-41	VC0837-VC0848	1	29,162
GI-44	VCA1026-VCA1027	2	9,732
GI-52	VC0208-VC0209	1	6,589
GI-55	VC1340	1	12,210
GI-57	VCA0018-VCA0022	2	6,399
GI-61	VCA1024-VCA1067	2	6,421
GI-66	VC0795-VC0807	1	5,930
GI-79	VCA0280-VCA0291	2	10,459
GI-81	VC0837-VC0848	1	29,854
GI-82	VC0080-VC0081	2	23,495
GI-85	VC1854-VC1835	1	15,469
GI-87	VC1326-VC1329	1	3,431



Dhaka University Institutional Repository

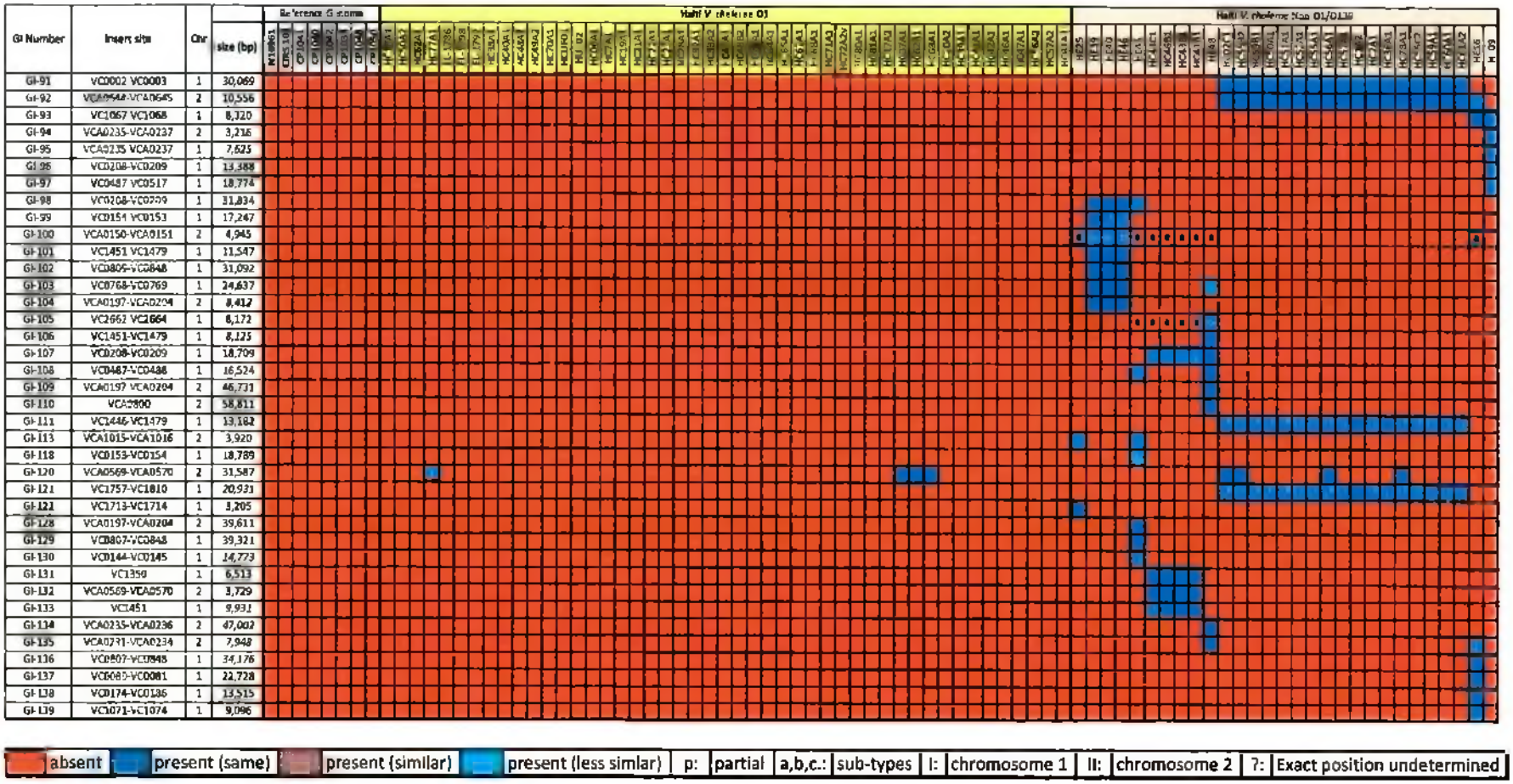


Table 8.3: Occurrence and distribution of mobile genetic elements (MGE's) in Haitian Strains. MGE's found are color-coded according to nucleotide sequence similarity values to the reference ORFs. Major genomic islands are indicated at the top row.

Dhaka University Institutional Repository

The Haitian genome sequences were found to be genetically similar to the recently emerged *V. cholerae* Altered El Tor strains CP1038 (Zimbabwe, 2009), CP1048 and CP1050 (Bangladesh, 2010), CP1041 (Zambia, 2004), and CP1042 (Thailand, 2010). These genomes belong to the 7th-pandemic clade, possessing a genomic backbone representative of *V. cholerae* O1 El Tor and *V. cholerae* O139, *i.e.*, phylocore genome 1 (PG-1) clade (14, 15) with variations in their mobile genetic elements (MGEs) (Table 8.3). MGE encode toxins, antibiotic resistance, virulence factors, and a wide range of proteins with as yet unknown function. MGEs are resident in the natural genetic reservoir of *V. cholerae* and related *Vibrio* spp., including *V. parahaemolyticus* and *V. mimicus* (14, 15). Thus, within a dynamic microbial gene pool, where horizontal gene transfer is a significant driver of microbial evolution (16), MGEs play a pivotal role in the genesis of diversity and hyper-variability of *Vibrio* genomes (14, 15). The MGE catalogue of the Haitian *V. cholerae* O1 strains demonstrates that they share genetically similar genomic and pathogenic islands (n=18) with the recently isolated Altered El Tor *V. cholerae* strains from South Asia and Africa (Table 8.3). Like the concurrent epidemic strains, CTXΦ genes were located in the large chromosome (C-I) of the Haitian *V. cholerae* O1 (HCO1) isolates (Table 8.3). RS1 was detected in most of the HCO1 strains, but four strains appeared to lack RS1 (Table 8.3) as does strain (CP1041) from a 2004 epidemic in Zambia. In addition, all HCO1 isolates, except HC49A2, which also lacked RS1, harbor TLC (Toxin-linked cryptic plasmid, a genetic element that plays some role in the biology of CTXΦ, perhaps facilitating either its acquisition or its replication). As anticipated, none of the Haitian *V. cholerae* non-O1/O139 strains had CTXΦ, RS1, or TLC.

Two genomic islands (GIs) in two other HCO1 strains (GI-16 in AELH.1792, and GI-120 in HC-77A1), absent in all concurrent epidemic strains examined in this study, were located in the same insertion site (VCA0569-VCA0570) in the small chromosome (C-II). GI-16 is a 6.6 kb island engendering *Vibrio* VSK prophage, a pre-CTX prophage, integration of which has previously been described in *V. cholerae* O139 MO10 (C-I) (15) and non-O1/O139 AM-19226 and 1587 (in C-II) (16), but heretofore not reported in *V. cholerae* O1. Further inspection identified GI-16, with a genetically similar set of coding sequences (CDS) on C-I (VC1451-VC1456), in six of the Haitian non-O1/O139 strains

Dhaka University Institutional Repository

including one strain (HC-2A1) containing GI-16 at both the C-I and C-II insertion sites (Table 8.3). GI-120, inserted at the same insertion site in HC-77A1, is a 31.5 kb island with 30 CDS, encoding different transporters and stress response protein. Our search for this GI identified another Haitian strain, HC-78A1, a clinical non-O1/O139 isolate, harboring this GI at the same chromosomal location. Subsequent analysis indicated GI-120 might be the result of rearrangement of the C-I loci VC1416 to VC1446 into C-II. In addition, all HCO1 strains contained the 10 GIs (GI-01 to GI-10) previously reported in *V. cholerae* O1 (Table S2). In contrast, the 29 clinical and environmental *V. cholerae* non-O1/O139 strains revealed a number of GIs (n=40) encoding genes for virulence, survival, stress response, metabolism etc. and distributed across both chromosomes. The environmental strains possessed a substantially larger repertoire of GIs (Table 8.3).

Variations in CTXΦ-RS1. CTXΦ present in toxigenic *V. cholerae* consists of two functionally distinct gene clusters, core and RS2 region (17, 18). The core contains *ctx*AB, encoding cholera toxin (CT), together with five other genes, *psh*, *cep*, *pIIICTX*, *ace*, and *zot,* required for phage morphogenesis. RS2 encodes proteins with roles in replication (RstA), integration (RstB), and regulation (RstR) of site-specific recombination of CTXΦ (17, 19). In toxigenic *V. cholerae* El Tor strains, a satellite phage, RS1, invariably is present and located adjacent to CTXΦ. It is similar to RS2 except that it contains an additional gene, *rstC*, that encodes an antirepressor protein, promoting transcription of CTXΦ genes required for generating infectious particles (19). The CTXΦ and RS1 element array in the Haitian cholera isolates are similar, but not identical to those found in the Altered El Tor *V. cholerae* strains, i.e., CIRS101.

We observed a single amino acid deletion (delta S), a GTA deletion at nucleotides 77-79 in *rst*B of the Haitian *V. cholerae* O1 strains (Table 8.4). This deletion, a novel feature for a classical biotype of *V. cholerae* (20), was not detected in any of the Altered El Tor, Hybrid, or El Tor variants of *V. cholerae*, except *V. cholerae* CP1048 (Bangladesh 2010) and CP1032 (Mexico, 1991). Interestingly, *V. cholerae* INDRE 91/1 (Mexico, 1991) and RC27 (Indonesia, 1991) contained the GTA deletion, underscoring this deletion is not novel to Western Hemisphere isolates. Moreover, PCR amplification

Dhaka University Institutional Repository

and sequence analysis of *rstB* gene identified novel polymorphisms (AèG, at nt. 192, and GèA, at nt. 108). The former was found only in the HCO1 strains, whereas the latter was also detected in the Mexico strains, *V. cholerae* INDRE91/1 and CP1032, suggesting the GèA polymorphism may be specific to Western Hemisphere strains, a conclusion requiring additional data.

Strain	Gene	CTX																														
		rstR	*rstA*																			*rstB*					*cep-zot*	*zot-ctxA**	*ctxB*			
	Pos	-	27	51	52	162	183	258	345	354	414	459	516	540	558	579	606	774	77-79	90	96	288	291	344	80 sites	1197 - 74	-					
N16961	CTX^{ET}	ET	C	T	T	C	C	G	G	C	T	C	G	A	C	T	T	C	GTA	A	T	A	C	G	E	4	E					
O395	CTX^{CL}	CL	T	.	.	T	A	C	T	.	.	.	A	G	.	C	C	T	-	T	C	G	T	.	C	7	C					
IB1346	CTX^{MOZ}	CL	T	.	.	T	A	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	E	3, 4	C					
IB1617	CTX^{ENV}	ENV	T	C	G	T	.	.	.	T	A	T	A	G	T	C	.	.	.	.	.	.	.	T	E	3, 4	C					
CIRS101	CTX^{HYB}	ET	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	E	3	C					
HCO1	CTX^{HYB}	ET	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	_	.	.	.	.	.	E	5	C^{6}					

Table 8.4: Sites of nucleotide variation in CTX prophages and RS1 elements. Dots indicate sequence identical to that of *V. cholerae* N16961. ET or E: El Tor type, C: Classical type, and Env: Environmental type.

Variation was also detected in the number of ToxR binding repeats (TTTTGAT) between *zot* and *ctx*A. *V. cholerae.* ToxR is a global transcriptional regulator of virulence gene expression and this repeated sequence is required for ToxR binding and activation of the *ctx*AB promoter. The ToxR-binding site is located immediately upstream of *ctx*AB and affinity of ToxR binding is influenced by the repeat sequences (21). The HCO1 strains contained five copies of this repeat (Table 8.4), compared to three or four copies in other Altered El Tor, Hybrid or El Tor variants of *V. cholera,* including recent epidemic strains (CP1048, CP1038, CP1032), and seven copies were found in the classical biotype of *V. cholerae* O395. The presence of an increased number of ToxR binding repeats has been hypothesized to correlate with a severe form of cholera (22, 23), which may be the case in the Haitian epidemic (24).

A further feature of the Haitian strains was the presence of three mutations in *ctx*B compared with two in Altered El Tor Hybrid or El Tor variant strains of *V. cholerae* O1

Dhaka University Institutional Repository

(24). The same *ctxB* polymorphisms in *V. cholerae* have previously been reported for strains from Kolkata and Orissa, India, and most recently in *V. cholerae* isolates from Nigeria, Cameroon, Kenya, Sudan, Ethiopia and Nepal (25, 26). Taken together, these data suggest that the ancestor to the strains responsible for the Haitian outbreak successfully utilized genetic changes in CTXΦ and RS1 element to affect disease severity, an attribute of the classical biotype and thereby achieve better environmental fitness, an attribute of the El Tor biotype—traits that *V. cholerae* has adopted in its recent evolution (5, 9-11).

***Vibrio* Pathogenicity Island-1 (VPI-1).** VPI-1 codes for the toxin co-regulated pilus involved in biofilm formation, attachment to chitinous organisms, and as the receptor for CTXφ (27). This island was found in all HCO1 isolates but not in clinical or environmental *V. cholerae* non-O1/O139 isolates. Interestingly, although devoid of VPI-1, three different GIs were identified at the VPI-1 site in 85% of non-O1/O139 isolates (Table S3). GI-41 was identified at the VPI-1 locus in one environmental and 15 clinical isolates, whereas other GIs, GI-102 and GI-81, were identified in other environmental isolates (Table S3). It is of interest that GI-81 represents a part of the *Vibrio* Seventh Pandemic Island II (VSP-II) but is inserted at the VPI-1 site (Table 8.3, 8.5).

***Vibrio* Pathogenicity Island-2 (VPI-2).** Most of the genomes have encoded sequences with at least some homology to the VPI-2 inserted at the canonical locus for this island (VC1758 and VC1810) (15). Whereas all *V. cholerae* O1 strains contained a complete VPI-2, many non-O1/O139 strains, both clinical and environmental, contained only the sialidase sequence of the sialic acid metabolism region of VPI-2. That is, the VPI-2 locus in non-O1/O139 strains lack most of the VPI-2 ORFs, including the type 1 restriction modification (T1RMS) system and mu-like phage regions (Table 8.5). Interestingly, the sialidase has been shown to unmask the GM1 gangliosides of human intestinal epithelial cells, thus making them more available to cholera toxin (28). Yet, as many of these strains do not encode cholera toxin, the high level of conservation of sialidase among these strains suggests an additional role for sialidase in the life cycle of *Vibrio*. Notably, the clinical non-O1/O139 isolate HC1A2 contained GI-121 inserted at

Dhaka University Institutional Repository

Table 8.5. Distributions and genomic variations of other known pathogenicity islands in outbreak and reference isolates. "+"= island is present; "-"= island is absent; Sialidase += only sialidase positive; -(T3SS)= type 3 secretion system inserted in this region, GI number in the parenthesis indicates particular GI located at this site.

Strain	Serotype	VPI-1	VPI-2				VSP-I	VSP-II	ICE
			VPI-2	T1RM5	S.A.M.	Mu-Phage			
HC17A1	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC17A2	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC20A2	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC23A1	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC37A1	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC38A1	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC39A1	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC41A1	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC42A1	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC46A1	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC47A1	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC48A1	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC50A2	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC56A2	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC57A2	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC61A1	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC62A1	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC62B1	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC69A1	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC70A1	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC77A1	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
HC81A2	O1	+	+	+	+	+	+	CIRS101_VSP-II	ICEVchBan5
AELH 1792	O1	+	+	+	+	+	+	ND	ND
HC49A2	O1	+	+	+	+	+	+	ND	ND
AELH 1786	O1	+	+	+	+	+	+	ND	ND
AELH 1798	O1	+	+	+	+	+	+	ND	ND
HC40A1	O1	+	+	-	+	-	+	ND	ND
HC43A1	O1	+	+	+	+	+	+	ND	ND
HCUF01	O1	+	+	+	+	+	+	ND	ND
HFU0	O1	+	+	+	+	+	+	ND	ND

HC02C1	Non-O1/O139	- (GI-41)	+	-	+	-	-	-	Uncharacterized ICE
HC41B1	Non-O1/O139	-	+	-	+	-	-	-	-
HC43B1	Non-O1/O139	-	+	-	+	-	-	-	-
HC46B1	Non-O1/O139	-	+	-	+	-	-	-	-
HC50A1	Non-O1/O139	- (GI-41)	+	-	Sialidase +	-	-	-	Uncharacterized ICE
HC51A1	Non-O1/O139	- (GI-41)	+	-	Sialidase +	-	-	-	Uncharacterized ICE
HC52A1	Non-O1/O139	- (GI-41)	+	-	Sialidase +	-	-	-	Uncharacterized ICE
HC55A1	Non-O1/O139	- (GI-41)	+	-	Sialidase +	-	-	-	Uncharacterized ICE
HC55B2	Non-O1/O139	- (GI-41)	+	-	Sialidase +	-	-	-	Uncharacterized ICE
HC55C2	Non-O1/O139	- (GI-41)	+	-	Sialidase +	-	-	-	Uncharacterized ICE
HC56A1	Non-O1/O139	- (GI-41)	+	-	Sialidase +	-	-	-	Uncharacterized ICE
HC57A1	Non-O1/O139	- (GI-41)	+	-	Sialidase +	-	-	-	Uncharacterized ICE
HC59A1	Non-O1/O139	- (GI-41)	+	-	Sialidase +	-	-	-	Uncharacterized ICE
HC59B1	Non-O1/O139	- (GI-41)	+	-	Sialidase +	-	-	-	Uncharacterized ICE
HC61A2	Non-O1/O139	- (GI-41)	+	-	Sialidase +	-	-	-	Uncharacterized ICE
HC60A1	Non-O1/O139	- (GI-41)	+	-	Sialidase +	-	-	-	Uncharacterized ICE
HC1A2	Non-O1/O139	- (GI-41)	+	-(GI-121)	Sialidase +	-	-	ND	ND
HC2A1	Non-O1/O139	- (GI-41)	+	-	Sialidase +	-	-	ND	ND
HC44C1	Non-O1/O139	-	+	-	+	-	-	ND	ND
HE09	Non-O1/O139	- (GI-81)	-	-	-	-	-	ND	ND
HE16	Non-O1/O139	- (GI-102)	-	-	-	-	- (New GI)	ND	ND
HE25	Non-O1/O139	-	+	- (T3SS)	+	-	-	ND	ND
HE39	Non-O1/O139	- (GI-102)	+	-	+	-	-	ND	ND
HE40	Non-O1/O139	- (GI-102)	+	-	+	-	-	ND	ND
HE45	Non-O1/O139	- (GI-102)	+	-	+	-	-	ND	ND
HE46	Non-O1/O139	- (GI-102)	+	-	+	-	-	ND	ND
HE48	Non-O1/O139	- (GI-41)	-	-	-	-	-	ND	ND

"+" = Island is present, "-" = island is absent; "(GI-41)" = GI-41 is located at this island site; "(GI-81)" = GI-81 is located at this island site; "(GI-102)" = GI-102 is located at this island site; "Sialidase +" = only sialidase ORF is present; "- (T3SS)" = type 3 secretion system inserted in this region.

Dhaka University Institutional Repository

VPI-2 site and part of GI-09 encoding genes responsible for response regulator, autolysin sensor kinase, and ABC-type transport system. An environmental isolate, *V. cholerae* HE25, encoded a type III secretion system within the VPI-2 region. The Type III secretion systems most frequently found in *V. cholerae* serogroups non-O1/non-O139 induce severe diarrhea in cholera infection models (29) and this secretion system has been reported in *V. cholerae* strains which encode and express the O1-antigen (15). Our results are consistent with previous studies showing VPI-2 to be highly heterogeneous in both clinical and environmental strains (15, 28).

***Vibrio* Seventh Pandemic (VSP) Islands.** All HCO1 but none of the Haitian *V. cholerae* non-O1/O139 genomes examined in this study carried VSP-I. Interestingly an environmental *V. cholerae* non-O1/O139, HE16, carried a previously uncharacterized GI at the VSP-I insertion loci. Similarly, all HCO1 carried a variant VSP-II, first reported in *V. cholerae* O1 CIRS101, though none of the *V. cholerae* non-O1/O139 were found to carry this or any other variant at the VSP-II locus (Table 8.5). The VSP-II island in O1 strains was inserted at the usual insertion locus in the genome and, as observed in *V. cholerae* CIRS101, revealed many deletions compared to the VSP-II in other *V. cholerae* O1 seventh pandemic strains described previously (30).

Integrated Conjugated Elements (ICE). ICE found in HCO1 strains were compared with ICE*Vch*Ban5 and ICE*Vch*Moz in *V. cholerae* O1 CIRS101 and B33, respectively. Analysis of their genetic organization revealed HCO1 ICE is >99% similar to ICE*Vch*Ban5, differ only by 5SNPS (31). In contrast, the majority of the Haitian *V. cholerae* non-O1/O139 strains carry an ICE sequence variant—only three strains lacked an insertion at this site (Table 8.5). All of the ICE insertions encode an integrase gene of the P4 integrase family and integrate at *prfC* on C-I. Minor differences between these elements and ICE*Vch*Ban5 were observed at the junctions between sequence contigs. In contrast, the ICE found in Haitian clinical *V. cholerae* non-O1/O139 was distinct from any ICE reported to date, notably lacking the cluster of CDSs encoding antibiotic resistance. Although this might be expected due to the higher variability observed in ICE present in *V. cholerae* non-O1/O139 from other geographical areas (32), it is interesting

Dhaka University Institutional Repository

that the same element is present in each of the Haitian *V. cholerae* non-O1/O139 strains. Even more intriguing is the fact that these *V. cholerae* non-O1/O139 strains were of clinical origin and in 21% the only isolate cultured from the patient.

Mutations in *gyrA* and *parC*. The Haitian *V. cholerae* isolates were resistant to sulfisoxazole, trimethoprim-sulfamethoxazole, furazolidone, streptomycin, and nalidixic acid, and displayed reduced susceptibility to ciprofloxacin (3, 5). The strains were sensitive to azithromycin and tetracycline. Analysis of the *gyrA* and *parC* (VC1258, VC2430, respectively, in Table 8.5) genes revealed two point mutations in the Haitian strains, *i.e.*, a ser83ile substitution in *gyrA* and a ser85leu substitution in *parC*. Both point mutations are associated with quinolone resistance in clinical *V. cholerae* and have been reported in India and most recently in Nigeria and Cameroon (33-37). We observed the same point mutations in *gyrA* and *parC* in recent isolates from Zimbabwe (CP1038), Thailand (CP1042), and Bangladesh (CP1048). In contrast, whereas *V. cholerae* O1 CIRS101 (Bangladesh, 2002) and CP1041 (Zambia, 2004) contained the same point mutation at *gyrA*, neither strain harbored the *parC* mutation, offering evidence of similarity with, but not identity to *V. cholerae* in Asia and Africa.

Phylogenomics. The phylogeny of the Haitian *V. cholerae* isolates was inferred by constructing a genome-relatedness neighbor-joining tree employing homologous alignment of 992 orthologous protein-coding genes (~989552 bp) of 96 *V. cholerae* genomes, as a strict measure of the core *V. cholerae* genome. The evolutionary tree (Fig. 8.3a) corroborated the previously described 1.7 million bp phylogeny of *V. cholerae* (15), showing serogroups of *V. cholerae* evolving into distinct evolutionary lineages. Significantly, 26 of 29 *V. cholerae* non-O1/O139 strains included in our phylogenetic analysis clustered with toxigenic *V. cholerae* O1 strains (i.e., Amazonia and TM11079-80) (Fig. 8.3a) implicated in sporadic cases and severe epidemics originating in the Western Hemisphere. The phylogeny is concordant with the observation that non-O1/O139 strains isolated from the Haitian clinical samples harbor a battery of virulence/fitness factors organized on MGEs. That is, despite being non-O1/O139, these strains harbor the genomic backbone similar to some toxigenic *V. cholerae* O1,

Dhaka University Institutional Repository

explaining the coexistence of non-O1/O139 with *V. cholerae* O1. Furthermore, the clinical non-O1/O139 isolates formed two closely related clusters (HC-1 and HC-2), with the majority of the non-O1/O139 isolates clustering together in HC-2 (Fig. 8.3b). Both clinical and environmental isolates clustered together, along with 1991 and 1980 *V. cholerae* O1 isolates from Brazil, indicating potential indigenous origin of these strains. The Haitian *V. cholerae* O1 strains, on the other hand, clustered with other 7th pandemic *V. cholerae* strains in a single monophyletic clade (Fig. 2a). A tree constructed with only 7th pandemic strains (Fig. 8.3c) showed clustering of the Haitian strains together, as well as with clinical isolates from Zimbabwe (CP1038) and Bangladesh (CP1048, and CP1050), an observation in agreement with the recent report of shared ancestry of Haiti, India and Cameroon isolates (8). Haitian *V. cholerae* O1 isolates subdivided into several closely related clusters within a tight monophyletic clade indicting stochastic biological transmission events resulting in genetic drift among the strains isolated during the first three weeks of the outbreak. Genetic drift would account for radiation of ~50% of the isolates away from the basal Haitian isolates. To better demonstrate the heterogeneity of these clonal subpopulations, principal component analysis (PCA), a multivariate method useful for exploratory cluster visualization, was employed. Based on the divergence of average nucleotide and tetra nucleotide frequency matrix using R, PCA analysis showed the Haitian strains formed a heterogeneous cloud of clonal subpopulations (Fig. 8.4), indicating that two closely related isolates from Bangladesh (CP1048) and Zimbabwe (CP1048) hard to exclude as maximum variations observed among the various Haitian strains must be included (Fig. 8.4).

Relatedness of outbreak and reference isolates by Whole Genome Mapping

Whole genome maps were created for a set of Haitian *V. cholerae* isolates and isolates from concurrent epidemics elsewhere. Comparison of the whole genome maps showed all Haitian *V. cholerae* O1 outbreak strains to be identical within the limits of the mapping technology resolution, suggesting the outbreak to be clonal. Detailed optical map comparisons of outbreak isolates with closely related reference isolates CP1048 and CP1038 revealed substantial similarity, yet differences were observed. Comparison of these regions with annotated sequences of a previously studied isolate, AE-1786, showed

Dhaka University Institutional Repository

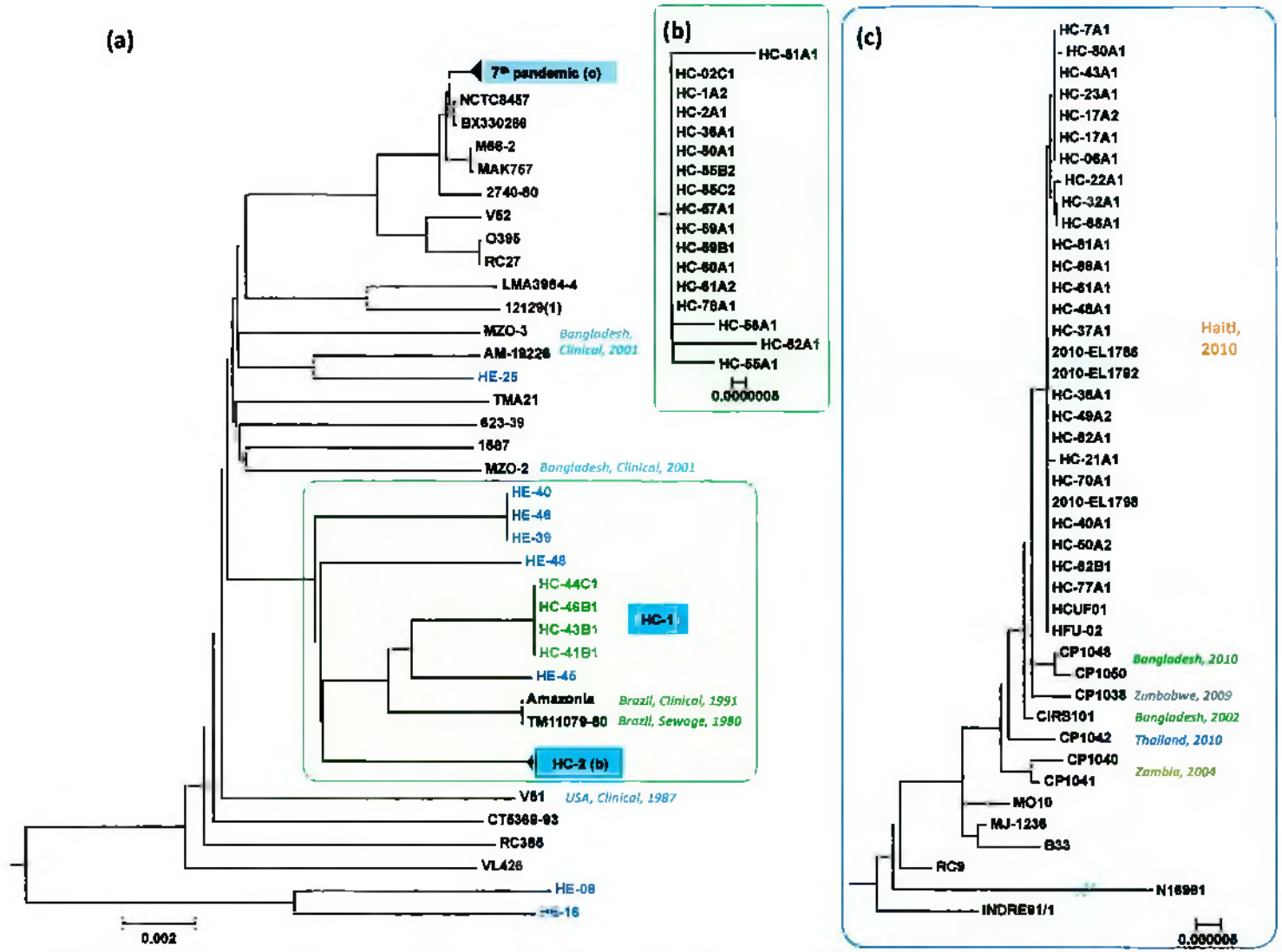
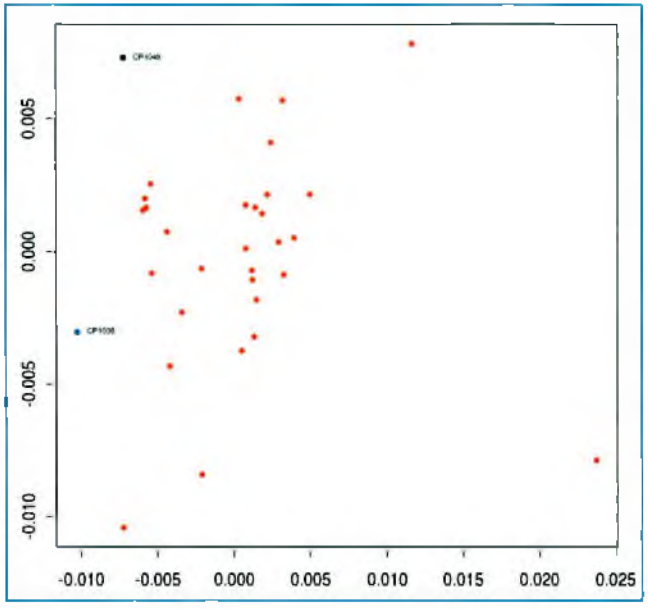
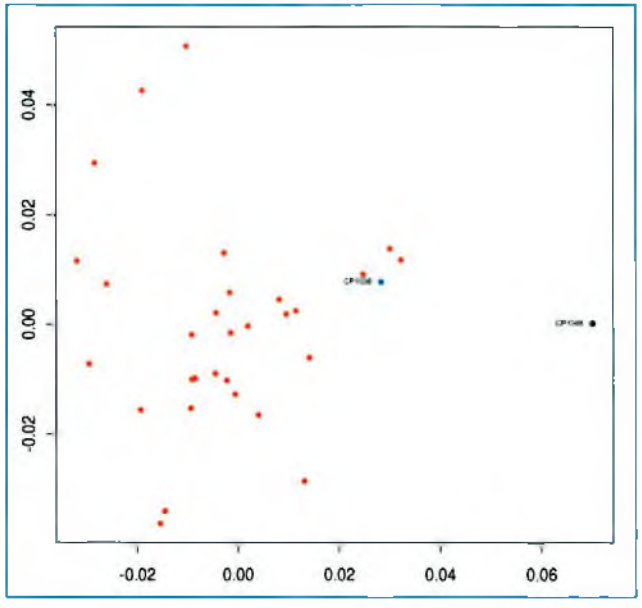


Figure 8.3: Neighbor-joining trees showing phylogenetic relationships based on 992 orthologous protein-coding genes (~989552 bp) of (a) 96 *V. cholerae* genomes, representing diverse serogroups. Strains with blue and green color represent Haitian environmental and clinical isolates, respectively; (b) *V. cholerae* non-O1/O139 strains representing clinical Non-O1/O139 group 2; (c) *V. cholerae* O1 and O139 strains representing the 7th pandemic cluster.





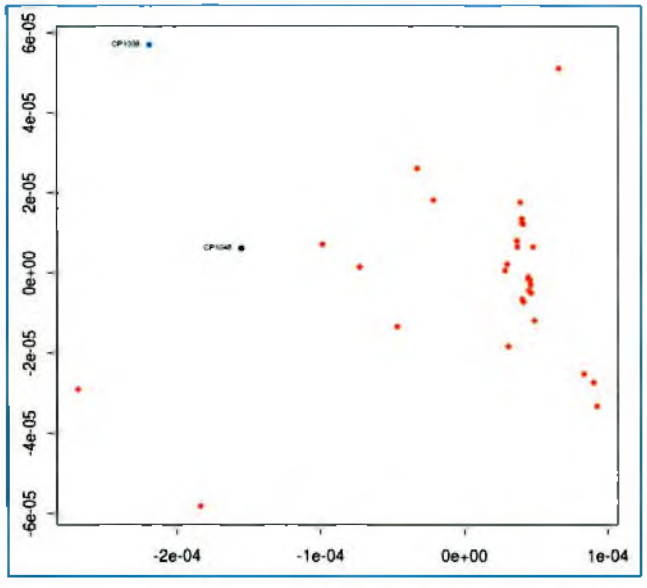


Figure 8.4: The two-dimensional PCA projection plots based on divergence of average nucleotide identity and tetra nucleotide frequency The figures showed the diversity among the Haitian strains (red) as opposed to two closely related reference strains CP1048 (black) and CP1038 (blue).

Dhaka University Institutional Repository

that while most of the HCO1 and CP1048 (Bangladesh, 2010) contained TLC of approximately 18Kb, two strains AE-1786 (Haiti, 2010) and CP1038 (Zimbabwe, 2009), displayed a TLC of approximately 13kb (Fig. 8.5a). Distinct polymorphism was also observed in the mannose sensitive hemagglutinin (MSHA) (Fig. 8.5b), a surface pilus contributing the ability of *V. cholerae* to adhere to plankton (38). Detailed whole genome map comparisons demonstrated three Haitian strains, AE-1786, HC61A1, and HC46A1, harbored an ~18kb tRNA cluster in MSHA, whereas other HCO1 and two closely related reference isolates CP1048 and CP103 contained ca. 24kb cluster (Fig. 8.5b). A larger insert (~27Kb) was identified at the ICE locus in Haitian strain HC20A2, compared with the ~18Kb insert found in other HCO1 strains (Fig. 8.5d) as well as CP1048 and CP1038.

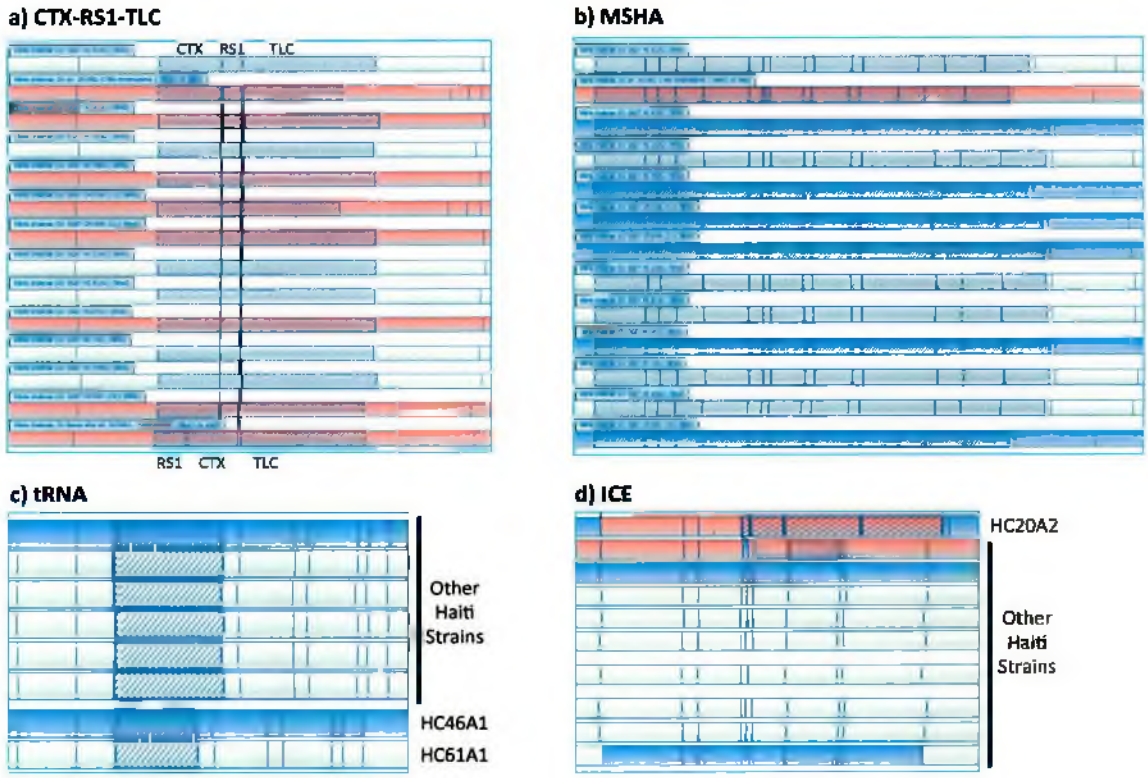


Figure 8.5: Polymorphisms in Haitian outbreak and reference isolates as determined by whole genome mapping. Polymorphism detected in a) TLC element, b) MSHA gene clusters, c) tRNA region, and d) ICE element. Highlighted by the blue colored crossed line box are regions where polymorphisms were detected.

Dhaka University Institutional Repository

Genome similarity clustering (Fig. 8.6) of whole genome maps for both chromosomes showed the outbreak isolates did not differ substantially from CP1038 and CP1048. Yet, other strains from Thailand (CP1042), Mexico (CP1032, 1033), and Zambia (CP1041) differed by >2% and therefore clustered separately. All clinical and environmental non-O1/O139 strains clustered in a monophyletic group. The genome similarity clustering confirmed heterogeneity of the *V. cholerae* non-O1/O139, as reported previously (38).

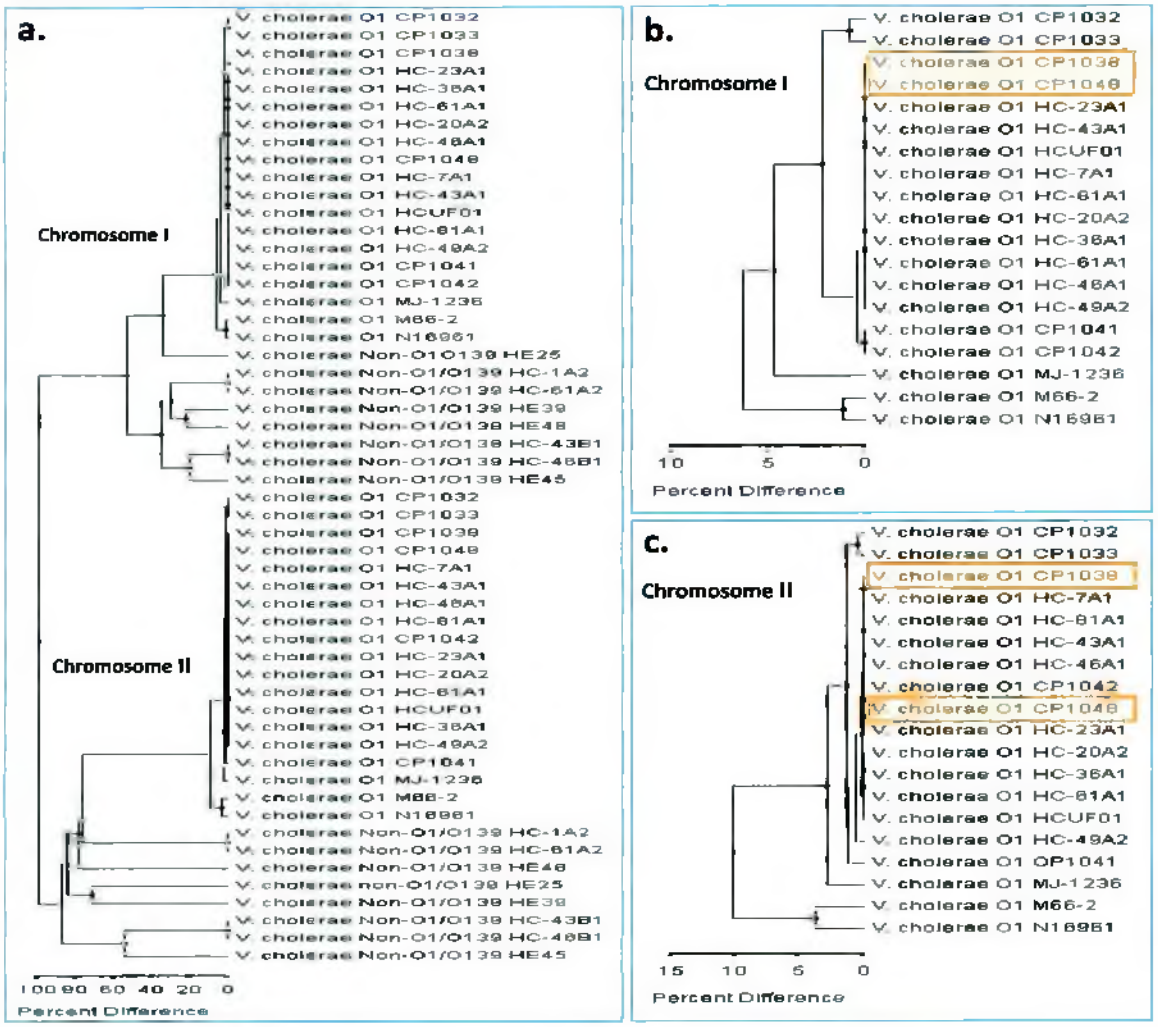


Figure 8.6: Whole genome map similarity cluster of Haitian outbreak isolates with reference strains using UPGMA: a) Chromosome I and II of each isolates clustered together; b) similarity cluster of *V. cholerae* O1 based on chromosome I; and c) similarity cluster of *V. cholerae* O1 based on chromosome II.

Dhaka University Institutional Repository

SNP-derived phylogeny. To establish a high-resolution phylogenomic framework, only 16 genomes were subjected to single nucleotide polymorphism SNP-discovery and SNP-validation. Our developed bioinformatics pipeline takes into account genome coverage and corrects for paralogous genomic regions. This analysis yielded a panel of 277 high-quality SNPs (70 synonymous, 169 non-synonymous, and 38 intergenic SNPs) distributed stochastically within the chromosome I (213 SNPs) and II (64 SNPs), without indication of mutational hotspots (Fig. 8.7, Fig. 8.8). The SNPs showed only two different alleles for every strain at each assayed position. The high incidence of bi-allelism suggests limited diversity among the analyzed genomes. The high-resolution of the SNP-derived phylogenetic scheme allowed accurate placement of the Haitian *V. cholerae* strains, separated by twelve Haiti specific SNPs from the rest of the strains included in the analysis (Fig. 8.9). Haitian *V. cholerae* strains formed a closely related cluster, with Haitian *V. cholerae* 1798 placing in the center of the radial expansion. The phylogeny shows genetic relationship of Haitian *V. cholerae* strains with *V. cholerae* CP1048, CP1038, CP1041 and CP1042 (Table 8.1). The Haitian *V. cholerae* strains feature up to five isolate-specific SNPs indicative of micro-evolution within the timeframe of the Haitian cholera epidemic. Only a single shared SNP is carried in the Haitian *V. cholerae* from *V. cholerae* HC38A1 to HC48A1.

Dhaka University Institutional Repository

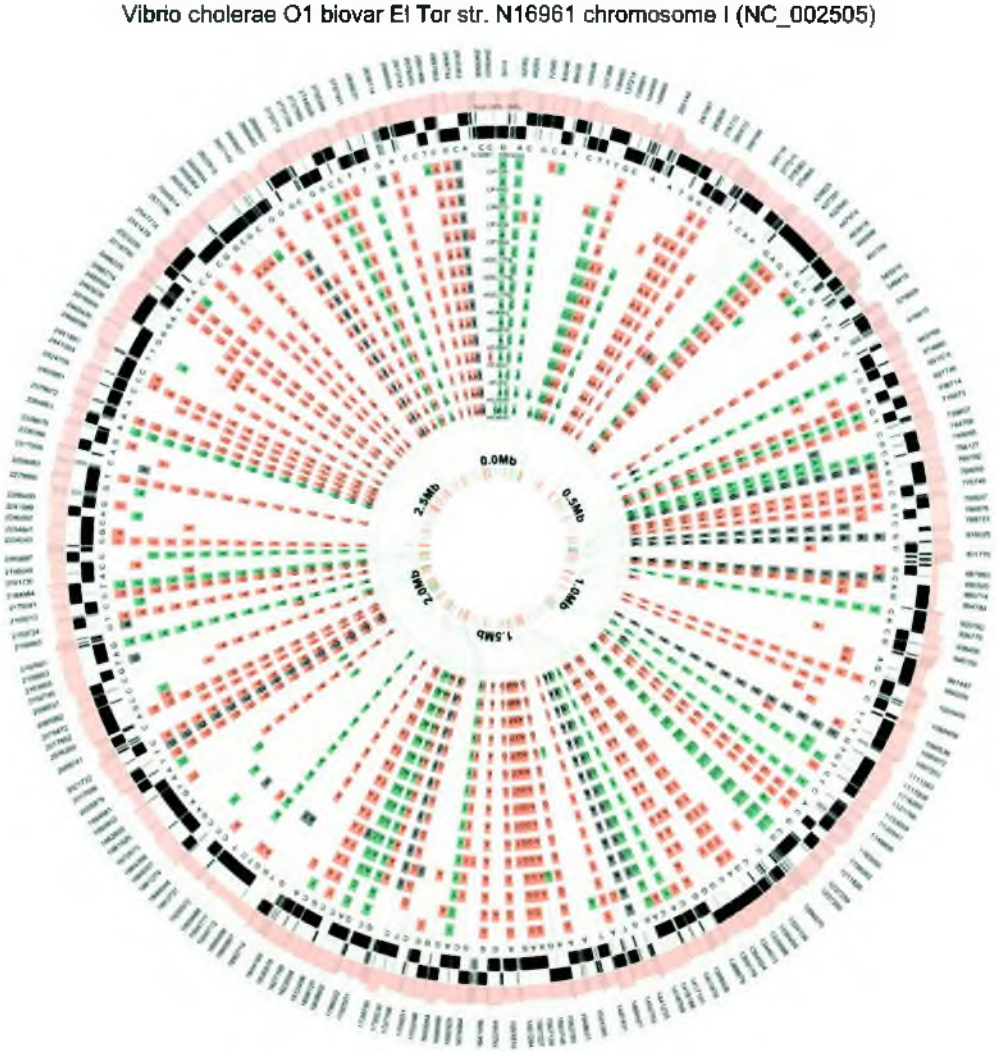


Figure 8.7: SNP-discovery on chromosome I in *V. cholerae* O1. Circles from outer to inner circle: Identified SNP positions referenced to N16961 positions on chromosome I (circle 1). GC-skew using a 5 kb window size (circle 2). Predicted open reading frames of strain N16961 encoded on the plus (circle 3) and minus strands (circle 4) (S4). SNP-backbone of *V. cholerae* O1 comprising 216 SNPs; 57 sSNPs (green), 134 nsSNPs (red) and 25 intergenic SNPs (grey). Individual SNP patterns for the 15 query genomes from outer to inner circle showing the reference base sequence (circle 5) and the query sequence base (circles 6 to 20) whenever it differs from the reference.

Dhaka University Institutional Repository

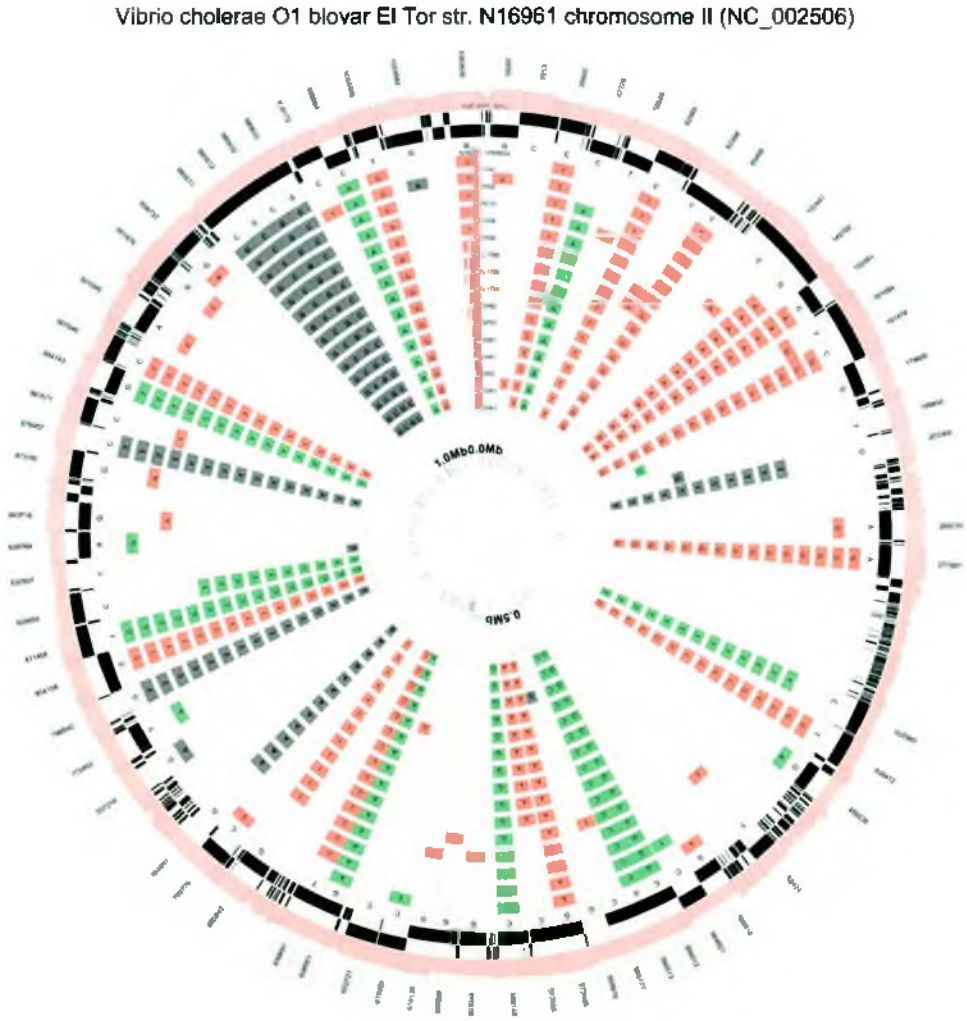


Figure 8.8: SNP-discovery on chromosome II in *V. cholerae* O1. Circles from outer to inner circle: Identified SNP positions referenced to N16961 positions on chromosome II (circle 1). GC-skew using a 5 kb window size (circle 2). Predicted open reading frames of strain N16961 encoded on the plus (circle 3) and minus strands (circle 4) (19). SNP-backbone of *V. cholerae* O1 comprising 67 SNPs; 16 sSNPs (green), 38 nsSNPs (red) and 13 intergenic SNPs (grey). Individual SNP patterns for the 15 query genomes from outer to inner circle showing the reference base sequence (circle 5) and the query sequence base (circles 6 to 20) whenever it differs from the reference.

Dhaka University Institutional Repository

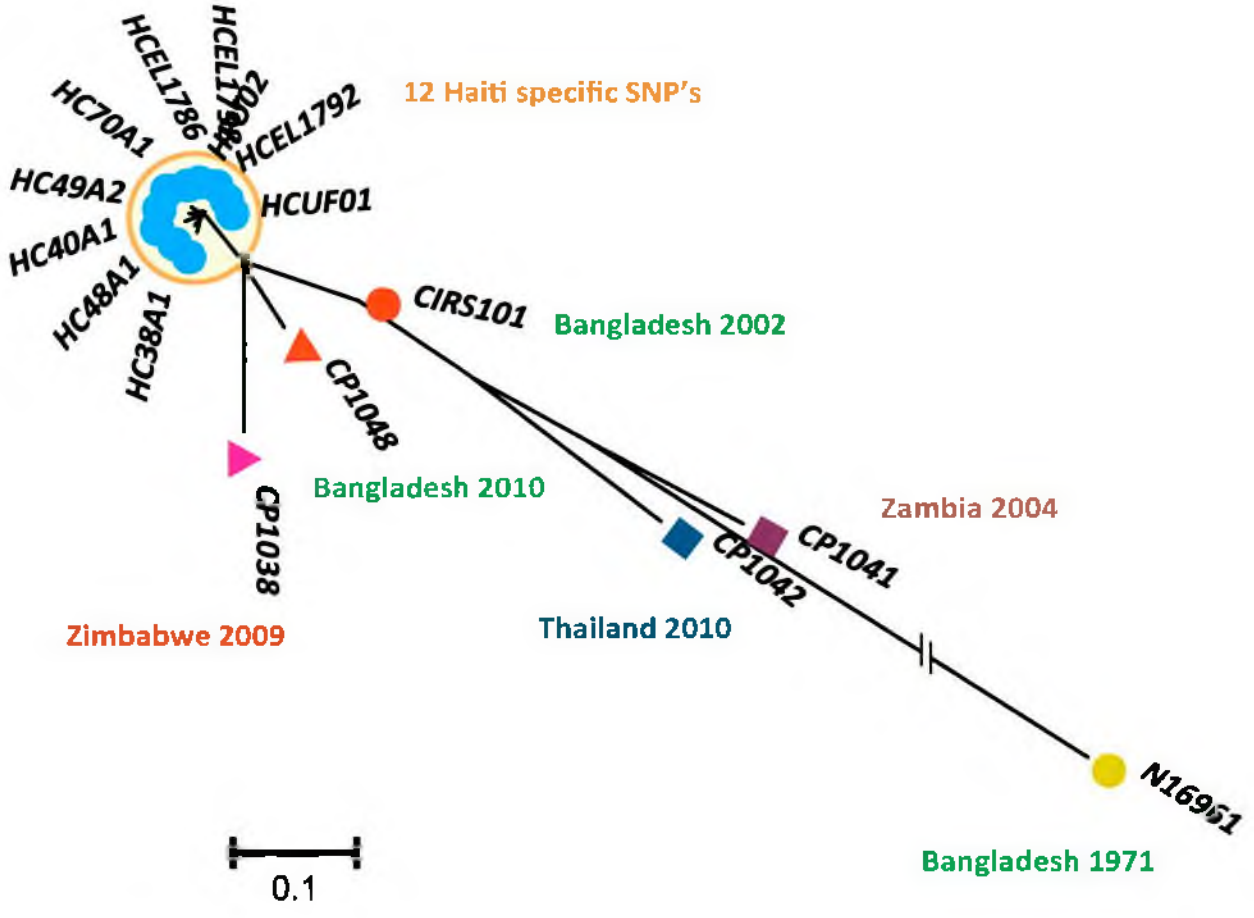


Figure 8.9: SNP-based phylogenetic tree for *V. cholerae* O1, constructed using 277 intra- and intergenic SNPs identified in the *V. cholerae* O1 population, revealing both clustering and radial expansion of the Haitian *V. cholerae* strains.

Dhaka University Institutional Repository

Conclusion:

Genomic analysis of *Vibrio cholerae* isolates obtained early in the Haitian cholera epidemic has provided evidence that two distinct *Vibrio* populations, *V. cholerae* O1 and *V. cholerae* non-O1/O139, contributed to the cholera epidemic in Haiti. Results of comprehensive genomic analysis showed that *V. cholerae* O1 populations were clonal, resembling epidemic isolates from South Asia and Africa. *V. cholerae* non-O1/O139 populations were not clonal but most probably serve as a reservoir for genomic and pathogenicity islands. Results of our analysis demonstrated that *V. cholerae* non-O1/O139 populations in Haiti harbor a genomic backbone similar to that of toxigenic *V. cholerae* O1 circulating in the Western hemisphere. Core gene and SNP-derived phylogenies suggest, and PCA findings reinforce, that quite quickly, *i.e.*, within a three-week period early in the cholera epidemic, substantial genomic diversity accumulated in the circulating population. Our study speaks to the critical need for an up-to-date, properly curated, and publicly available reference genomic database (including geographically diverse and circulating strains from cholera endemic regions) reflecting quality coverage of global phylogenetic diversity, as well as diversity within this heterogeneous clade and its near neighbors if investigations of future *V. cholerae* epidemics are to be effective and useful for public health applications. Moreover, it may also be necessary to generate the database at the time of a study (Hasan et al., unpublished data) that offers evidence of genetic drift during the short time frame of routine laboratory passage of *V. cholerae*. Therefore, a need to re-sequence reference genome, as the apposite comparator, is also underscored if SNP data are to be used to assign attribution in outbreak investigations.

References:

1. Colwell RR (1996) Global climate and infectious disease: the cholera paradigm *Science* 274, 2025-2031.
2. Rottmann S (2010) in *BBC News*.
3. CDC (2010). Laboratory Test Results of Cholera Outbreak Strain in Haiti Announced.
4. Farmer PE & Ivers LC (2012) Cholera in Haiti: the equity agenda and the future of tropical medicine *Am J Trop Med Hyg* 86, 7-8.
5. Chin CS, Sorenson J, Harris JB, Robins WP, Charles RC, Jean-Charles RR, Bullard J, Webster DR, Kasarskis A, Peluso P, *et al.* (2011) The origin of the Haitian cholera outbreak strain *N Engl J Med* 364, 33-42.
6. UN (2011). Final Report of the Independent Panel of Experts on the Cholera Outbreak in Haiti.
7. Hendriksen RS, Price LB, Schupp JM, Gillece JD, Kaas RS, Engelthaler DM, Bortolaia V, Pearson T, Waters AE, Upadhyay BP, *et al.* (2011) Population Genetics of *Vibrio cholerae* from Nepal in 2010: Evidence on the Origin of the Haitian Outbreak *mBio* 2, e00157-00111.
8. Reimer AR, Van Domselaar G, Stroika S, Walker M, Kent H, Tarr C, Talkington D, Rowe L, Olsen-Rasmussen M, Frace M, *et al.* (2011) Comparative genomics of *Vibrio cholerae* from Haiti, Asia, and Africa *Emerg Infect Dis* 17, 2113-2121.
9. Pun SB (2011) South Asia instead of Nepal may be the origin of the Haitian cholera outbreak strain. *mBio* 2, e00219-00211.
10. Kendall EA, Chowdhury F, Begum Y, Khan AI, Li S, Thierer JH, Bailey J, Kreisel K, Tacket CO, LaRocque RC, *et al.* (2010) Relatedness of *Vibrio cholerae* O1/O139 isolates from patients and their household contacts, determined by multilocus variable-number tandem-repeat analysis *J Bacteriol* 192, 4367-4376.
11. Ghosh R, Nair GB, Tang L, Morris JG, Sharma NC, Ballal M, Garg P, Ramamurthy T, & Stine OC (2008) Epidemiological study of *Vibrio cholerae* using variable number of tandem repeats *FEMS Microbiol Lett* 288, 196-201.
12. Stine OC, Alam M, Tang L, Nair GB, Siddique AK, Faruque SM, Huq A, Colwell R, Sack RB, & Morris JG, Jr. (2008) Seasonal cholera from multiple small outbreaks, rural Bangladesh *Emerg Infect Dis* 14, 831-833.
13. Ali A, Chen Y, Johnson JA, Redden E, Mayette Y, Rashid MH, Stine OC, & Morris JG, Jr. (2011) Recent clonal origin of cholera in Haiti *Emerg Infect Dis* 17, 699-701.
14. Hasan NA, Grim CJ, Haley BJ, Chun J, Alam M, Taviani E, Hoq M, Munk AC, Saunders E, Brettin TS, *et al.* (2010) Comparative genomics of clinical and environmental *Vibrio mimicus Proc Natl Acad Sci U S A* 107, 21134-21139.
15. Chun J, Grim CJ, Hasan NA, Lee JH, Choi SY, Haley BJ, Taviani E, Jeon YS, Kim DW, Brettin TS, *et al.* (2009) Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae Proc Natl Acad Sci U S A* 106, 15442-15447.
16. Treangen TJ & Rocha EP (2011) Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes *PLoS Genet* 7, e1001284.
17. Safa A, Nair GB, & Kong RY (2010) Evolution of new variants of *Vibrio cholerae* O1 *Trends Microbiol* 18, 46-54.
18. Waldor MK & Mekalanos JJ (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin *Science* 272, 1910-1914.
19. Waldor MK, Rubin EJ, Pearson GD, Kimsey H, & Mekalanos JJ (1997) Regulation, replication, and integration functions of the *Vibrio cholerae* CTXphi are encoded by region RS2 *Mol Microbiol* 24, 917-926.
20. Choi SY, Lee JH, Kim EJ, Lee HR, Jeon YS, von Seidlein L, Deen J, Ansaruzzaman M, Lucas GM, Barreto A, *et al.* (2010) Classical RS1 and environmental RS1 elements in *Vibrio cholerae* O1 El Tor strains harbouring a tandem repeat of CTX prophage: revisiting Mozambique in 2005 *J Med Microbiol* 59, 302-308.

Dhaka University Institutional Repository

21. Pfau JD & Taylor RK (1996) Genetic footprint on the ToxR-binding site in the promoter for cholera toxin *Mol Microbiol* 20, 213-222.
22. Kaper JB, Morris JG, Jr., & Levine MM (1995) Cholera *Clin Microbiol Rev* 8, 48-86.
23. Kay AB (1994) in *ASM press*, pp. 3–24.
24. Nair GB, Qadri F, Holmgren J, Svennerholm AM, Safa A, Bhuiyan NA, Ahmad QS, Faruque SM, Faruque AS, Takeda Y, *et al.* (2006) Cholera due to altered El Tor strains of *Vibrio cholerae* O1 in Bangladesh *J Clin Microbiol* 44, 4211-4213.
25. M. Humphrys CB, M. Parsons, K. Greene, N. Garrett, G. Gomez, G. Dahourou, J. Boncy, S. A. Balajee, D. Talkington (2011) in *ASM General Meeting*, p. 966.
26. Kumar P, Jain M, Goel AK, Bhadauria S, Sharma SK, Kamboj DV, Singh L, Ramamurthy T, & Nair GB (2009) A large cholera outbreak due to a new cholera toxin variant of the *Vibrio cholerae* O1 El Tor biotype in Orissa, Eastern India *J Med Microbiol* 58, 234-238.
27. Boyd EF, Moyer KE, Shi L, & Waldor MK (2000) Infectious CTXPhi and the vibrio pathogenicity island prophage in *Vibrio mimicus*: evidence for recent horizontal transfer between *V. mimicus* and *V. cholerae Infect Immun* 68, 1507-1513.
28. Almagro-Moreno S & Boyd EF (2009) Sialic acid catabolism confers a competitive advantage to pathogenic *Vibrio cholerae* in the mouse intestine *Infect Immun* 77, 3807-3816.
29. Shin OS, Tam VC, Suzuki M, Ritchie JM, Bronson RT, Waldor MK, & Mekalanos JJ (2011) Type III secretion is essential for the rapidly fatal diarrheal disease caused by non-O1, non-O139 *Vibrio cholerae MBio* 2, e00106-00111.
30. Taviani E, Grim CJ, Choi J, Chun J, Haley B, Hasan NA, Huq A, & Colwell RR (2010) Discovery of novel *Vibrio cholerae* VSP-II genomic islands using comparative genomic analysis *FEMS Microbiol Lett* 308, 130-137.
31. Sjolund-Karlsson M, Reimer A, Folster JP, Walker M, Dahourou GA, Batra DG, Martin I, Joyce K, Parsons MB, Boncy J, *et al.* (2010) Drug-resistance mechanisms in *Vibrio cholerae* O1 outbreak strain, Haiti, 2010 *Emerg Infect Dis* 17, 2151-2154.
32. Taviani E, Grim CJ, Chun J, Huq A, & Colwell RR (2009) Genomic analysis of a novel integrative conjugative element in *Vibrio cholerae FEBS Lett* 583, 3630-3636.
33. Baranwal S, Dey K, Ramamurthy T, Nair GB, & Kundu M (2002) Role of active efflux in association with target gene mutations in fluoroquinolone resistance in clinical isolates of *Vibrio cholerae Antimicrob Agents Chemother* 46, 2676-2678.
34. Garg P, Chakraborty S, Basu I, Datta S, Rajendran K, Bhattacharya T, Yamasaki S, Bhattacharya SK, Takeda Y, Nair GB, *et al.* (2000) Expanding multiple antibiotic resistance among clinical strains of *Vibrio cholerae* isolated from 1992-7 in Calcutta, India *Epidemiol Infect* 124, 393-399.
35. Krishna BV, Patil AB, & Chandrasekhar MR (2006) Fluoroquinolone-resistant *Vibrio cholerae* isolated during a cholera outbreak in India *Trans R Soc Trop Med Hyg* 100, 224-226.
36. Islam MS, Midzi SM, Charimari L, Cravioto A, & Endtz HP (2009) Susceptibility to fluoroquinolones of *Vibrio cholerae* O1 isolated from diarrheal patients in Zimbabwe *JAMA* 302, 2321-2322.
37. Quilici ML, Massenet D, Gake B, Bwalki B, & Olson DM (2010) *Vibrio cholerae* O1 variant with reduced susceptibility to ciprofloxacin, Western Africa *Emerg Infect Dis* 16, 1804-1805.
38. Chiavelli DA, Marsh JW, & Taylor RK (2001) The mannose-sensitive hemagglutinin of *Vibrio cholerae* promotes adherence to zooplankton *Appl Environ Microbiol* 67, 3220-3225.
39. Huq A, *et al.* (2006) Detection, isolation, and identification of *Vibrio cholerae* from the environment *Curr Protoc Microbiol* Chapter 6, Unit6A 5.
40. National Center for Infectious Diseases (U.S.) PAHO (1994) *Laboratory methods for the diagnosis of Vibrio cholerae. Atlanta, Ga: Centers for Disease Control and Prevention* (CDC/NCID, Atlanta, Ga).
41 Eppinger M, *et al.* (2010) Genome sequence of the deep-rooted *Yersinia pestis* strain Angola reveals new insights into the evolution and pangenome of the plague bacterium *J Bacteriol* 192, 1685-1699.

Dhaka University Institutional Repository

42. Huson DH, *et al.* (2001) Design of a compartmentalized shotgun assembler for the human genome *Bioinformatics* 17 Suppl 1, S132-139.
43. Zerbino DR (2010) Using the Velvet de novo assembler for short-read sequencing technologies *Curr Protoc Bioinformatics* Chapter 11, Unit 11 15.
44. Kurtz S, *et al.* (2004) Versatile and open software for comparing large genomes *Genome Biol* 5, R12.
45. Karlin S, *et al.* (1990) Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes *Proc Natl Acad Sci U S A* 87, 2264-2268.
46. Chun J, *et al.* (2009) Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae Proc Natl Acad Sci U S A* 106, 15442-15447.
47. Saitou N, *et al.* (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees *Mol Biol Evol* 4, 406-425.
48. Choi SY, *et al.* (2010) Multilocus variable-number tandem repeat analysis of *Vibrio cholerae* O1 El Tor strains harbouring classical toxin B *J Med Microbiol* 59, 763-769.
49. Stine OC, *et al.* (2008) Seasonal cholera from multiple small outbreaks, rural Bangladesh *Emerg Infect Dis* 14, 831-833.
50. Ghosh R, *et al.* (2008) Epidemiological study of *Vibrio cholerae* using variable number of tandem repeats *FEMS Microbiol Lett* 288, 196-201.
51. Jeon YS, *et al.* (2005) jPHYDIT: a JAVA-based integrated environment for molecular phylogeny of ribosomal RNA sequences *Bioinformatics* 21, 3171-3173.
52. Goris J, *et al.* (2007) DNA-DNA hybridization values and their relationship to whole-genome sequence similarities *Int J Syst Evol Microbiol* 57, 81-91.
53. Teeling H, *et al.* (2004) Application of tetranucleotide frequencies for the assignment of genomic fragments *Environ Microbiol* 6, 938-947.
54. Richter M, *et al.* (2009) Shifting the genomic gold standard for the prokaryotic species definition *Proc Natl Acad Sci U S A* 106, 19126-19131.

Dhaka University Institutional Repository

CHPATER 9:

CONCLUSION

Dhaka University Institutional Repository

Some three billion years past, and before plants and animals began their evolution here on Earth, the microbial world thrived and formed Earth's biosphere. The enormous evolutionary potential of these organisms is illustrated by the fact that the innumerable species currently living differ in many properties including metabolic capacities, cell surface compositions, life styles, ecological niches and host specificities. Alone, they established distinct niches, developing survival strategies thus enabling survival in the manifold conditions of the ever-changing landscape of Earth. Then, as other life forms evolved, the microbes took advantage of these survival factors and used them to adapt, survive, and colonize the various flora and fauna that began to appear and evolve. Here, we studied *Vibrio cholerae* as a classical example of such evolutionary process and uncovered that being an efficient pathogen, *V. cholerae* has evolved special mechanisms, which include the plasticity of the genome and the rate of phenotype generation, whereas survival/virulence factors conferring its fitness to withstand the selective pressures exerted by the environment and to proliferate. Cholera prevails in impoverished areas where safe drinking water is not available and where basic sanitary processes are compromised. Many impressive advances have been made in the past several decades, yet recent statistics from the World Health Organization (WHO) clearly show that global trends in the incidence of cholera have increased steadily since the beginning of the millennium. Although the African and Asian continents bear the brunt of endemic and epidemic cholera, all other continents report either endemic or epidemic cholera. Haiti offers the most recent example of how a country can be devastated when cholera strikes. Utilizing a holistic approach to study and understand the biology of *V. cholerae* and of the disease cholera in both endemic (Bangladesh) and epidemic (Haiti) regions, it is evident that *V. cholerae* is autochthonous in the coastal, brackish, and riverine water in with numbers of cells increasing just prior to the annual outbreaks of cholera. This study also demonstrated that the various epidemiologic features of cholera (its ability to spread worldwide and its periodicity in some endemic areas) are primarily the result of environmental factors, complex changes in the microbial community, and the genetic reassortment of *V. cholerae* itself.

A remarkable finding of this study was the close association of salinity, conductivity, water temp and water depth with the number of microbial species

Dhaka University Institutional Repository

detected by 16S metagenomics, which indicated that these predictors might also have predictive value for other water borne diseases, and therefore warrant further attention. From the result of this study, we were able to validate the effect of few previously proposed environmental variables on the occurrence of cholera, as well as propose additional variables, combinations of which might be useful in developing environmental models for the prediction of cholera epidemics.

We also demonstrated the occurrence and distribution of a large number of virulence-associated genes in environmental and clinical strains of *V. cholerae*. The vast majority of environmental *V. cholerae* strains were non-O1/O139 which carry a discrete set of virulence genes; infrequently carried by the toxigenic *V. cholerae* strains, and occasionally carry the genes for the major virulence factors, typically carried by the toxigenic *V. cholerae* strains. This supports the hypothesis that some pathogenic bacteria have evolved from nonpathogenic strains of the same species via horizontal transfer of virulence genes as virulence genes are dispersed among environmental strains of *V. cholerae* and may be ferried about, given the fact that most of the virulence genes are located on mobile elements. Our findings indicate that horizontal transfer of genes has occurred not only between O1 biotypes but also between O1 and non-O1/O139 strains, and stress the need to monitor the emergence and prevalence of non-O1/O139 strains, which not only are a source of morbidity as agents of sporadic disease but also represent a threat, since they might become potentially epidemic.

In this study we successfully implemented metagenomics to characterize the environmental community composition and demonstrated that the aquatic environment of Mathbaria harbor genetically diverse and complex microbial community including primary members of both freshwater and coastal ecosystems, which appears to be the result of frequent coastal water intrusion into the fresh water system in the region. Tremendous diversity was uncovered at every taxonomic level, and many bacterial genera pathogenic to human and animals were detected with possible implications for agriculture and public health. Careful analysis of data for spatial and temporal variation in microbial communities demonstrated that temporal variation in a natural community and their relative abundance was more pronounced than the variation between habitats on a given day, and *Vibrio cholerae* abundance

Dhaka University Institutional Repository

and distribution followed a similar trend. This suggests that the abundance and diversity of the microbial community including the *Vibrio* populations, exhibit a complex seasonal distribution driven by temporal fluctuations of environmental parameters, which helps explain the seasonal occurrence of *V. cholerae* in the environment. Our observation from diversity dissimilarity analysis, as well as species diversity and evenness in distribution, reaffirmed each other and suggests that selection is going on in the circulating community, and a comparison of the community diversity along with observed cholera cases in this region demonstrate that high diversity, alone cannot explain cholera epidemics. Rather, we found a good correlation of cholera cases when the ecosystem experiences high community diversity (alpha and gamma) with least dissimilarity (beta) among the habitats. Taken together, the results of this study suggests that specific consortium among the indigenous population, arising as a result of positive selection, might be crucial in responding on changes in environmental parameters that allow *V. cholerae* to overcome the natural competition and emerge in epidemic proportions. We also observed limitations of species level identity by 16S metagenomics; therefore, similar research could be extended to other endemic areas with similar seasonality as well as geographic proximity to human residence to establish a deep inventory of circulating bacterial species using whole genome shotgun sequencing to enable a genome centric analysis of species identity and to avoid amplification bias of 16 rRNA genes. This would also provide functional characteristics of the community. Such additional studies would establish essential nucleic acid sequence baselines for monitoring of disease-causing bacterial species that may have a significant impact on livestock, agriculture and human health as well as providing data with possible implications in the forecasting of emergence and re-emergence of an infectious disease like cholera.

Results of the genome sequence analysis reported here for *V. cholerae*, *V. mimicus* and *V. antiquarius* provided insight into the biology of these organisms, notably, in the refinement and expansion of our knowledge on metabolism, virulence, evolution, and phylogeny of these organism. These new genome sequences can now serve as reference for studies of communities where *Vibrio* spp. are present. It is now clear that the dynamics of pathogenic *V. cholerae* can be attributed to extensive lateral gene transfer *via* transduction, conjugation and transformation. However, the true evolutionary history of this bacterium remains to be documented. Because

Dhaka University Institutional Repository

phylogenetic analysis can be confounded by the extent of horizontal gene transfer, recombination or movement of mobile elements within and among *V. cholerae* and the environment, the lynchpin for understanding mobile genetic elements (MGE) is DNA sequence analysis. It is evident that many of the genes identified as virulence genes are important for the fitness of the bacterium in its native aquatic environments as evidenced by the presence of these genes in both the clinical and environmental as well as conservation of homologs in other environmental *Vibrio* genomes. Moreover, the presence of these genes in diverse *Vibrio* species (diverse phylogenetically and ecologically) suggests that these affiliations between commensals are likely very old and might help explain the common evolution of *Vibrio* species into pathogens of humans and marine animals. Genomic analysis of *Vibrio cholerae* isolates obtained early in the Haitian cholera epidemic has provided evidence that two distinct *Vibrio* populations, *V cholerae* O1 and *V. cholerae* non-O1/O139, contributed to the cholera epidemic in Haiti. Results of our analysis demonstrated that *V. cholerae* non-O1/O139 populations in Haiti harbor a genomic backbone similar to that of toxigenic *V cholerae* O1 circulating in the Western hemisphere. Core gene and SNP-derived phylogenies suggest, and PCA findings reinforce, that quite quickly, *i.e.*, within a three-week period early in the cholera epidemic, substantial genomic diversity accumulated in the circulating population. Our study speaks to the critical need for an up-to-date, properly curated, and publicly available reference genomic database (including geographically diverse and circulating strains from cholera endemic regions) reflecting quality coverage of global phylogenetic diversity, as well as diversity within this heterogeneous clade and its near neighbors if investigations of future *V. cholerae* epidemics are to be effective and useful for public health applications. Moreover, it may also be necessary to generate the database at the time of a study (Hasan et al., unpublished data) that offers evidence of genetic drift during the short time frame of routine laboratory passage of *V. cholerae*. Therefore, a need to re-sequence reference genome, as the apposite comparator, is also underscored if SNP data are to be used to assign geographic attribution in outbreak investigations.

In conclusion, the study provided evidence that cholera outbreak will not occur simply because *V. cholerae* is present, rather, the cause of the disease will be determined by the interaction of the community, the infectious agent and the environment. The challenge that remains is to develop a cholera prediction model for

Dhaka University Institutional Repository

the endemic and epidemic regions, identifying, optimizing environmental conditions, microbial community interactions and incorporating enabling transmission mechanisms. When either a major natural disaster strikes a coastal town or significant civil disorder seriously damages living conditions of a population center in the tropical belt, a cholera tracking mechanism should be operational that monitors the climate and environment surrounding the vulnerable population by satellite remote sensing. By monitoring water and sanitation infrastructure, displaced populations and convergence with environmental conditions and transmission pathways, a prediction capability with high probability should be developed to provide warning of potential cholera epidemics and mobilization of public health measures.

Moreover, future studies should also focus to investigate the microbial populations in the gastrointestinal tract to explore the etiology of gastrointestinal disease and recovery, providing insight into the dominant microflora populations and how these populations change in response to different disease conditions and therapies to provide a basis for rational development of "pro-/pre-/synbiotic" therapies (i.e., administration of specific bacterial species or nutrients to modify or supplement gut flora), therapies that would be very amenable to production and administration in developing countries. Our ability to design such therapies has in the past been severely limited by our lack of knowledge about the intestinal flora. Utilizing the genomic capabilities that are now available, we can generate data to open up these areas of investigation.

Dhaka University Institutional Repository

LIST OF PUBLICATIONS:

1. Full Text Articles:

A. Published:

1. **Hasan NA**, Choi SY, Eppinger M, Clark PW, Chen A, Alam M, Haley BJ, Taviani E, Hine E, Su Q, Tallon LJ, Prosper JB, Furth K, Hoq MM, Li H, Fraser-Liggett CM, Cravioto A, Huq A, Ravel J, Cebula TA, Colwell RR. Genomic diversity of 2010 Haitian cholera outbreak strains. Proc Natl Acad Sci U S A. 2012 Jul 17; 109(29): E2010-7.
2. Alam M, **Hasan NA**, Sultana M, Sadiq A, Nair GB, Faruque ASG, Endtz HP, Cravioto A, Sack RB, Huq A, Colwell RR, Arakawa E, Ohnishi M, Watanabe H. Diagnostic Limitations to Accurate Diagnosis of Cholera. J Clin Microbiol. 2010 Nov;48(11):3918-22
3. **Hasan NA**, Grim CJ, Haley BJ, Chun J, Alam M, Taviani E, Hoq M, Munk AC, Saunders E, Brettin TS, Bruce DC, Challacombe JF, Detter JC, Han CS, Xie G, Nair GB, Huq A, Colwell RR. 2010. Comparative genomics of clinical and environmental *Vibrio mimicus*. Proc Natl Acad Sci U S A. 107(49): 21134-9.
4. **Hasan NA**, Chowdhury WB, Rahim N, Sultana M, Shabnam SA, Mai V, Ali A, Morris GJ, Sack RB, Huq A, Colwell RR, Endtz HP, Cravioto A, Alam M. Metagenomic 16S rDNA Targeted PCR-DGGE in Determining Bacterial Diversity in Aquatic Ecosystem. Bangladesh J Microbiol. 2010; 27(2): 46-50.
5. Haley B, Grim CJ, **Hasan NA**, Taviani E, Chun J, Brettin T, Bruce D, Challacombe JF, Detter JC, Han CS, Huq A, Nair GB, Colwell RR. The Pre-7th Pandemic *Vibrio cholerae* BX 330286 El Tor Genome; Evidence for the Environment as a Genome Reservoir. Environ Microbiol Reports.2010; 2(1): 208-216.
6. Chun J, Grim CJ, **Hasan NA**, Lee JH, Choi SY, Haley BJ, Taviani E, Jeon YS, Kim DW, Lee JH, Brettin TS, Bruce DC, Challacombe JF, Detter JC, Han CS, Munk AC, Chertkov O, Meincke L, Saunders E, Walters RA, Huq A, Nair GB, Colwell RR. Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. Proc Natl Acad Sci U S A. 2009 Sep 8; 106(36): 15442-7.
7. Grim CJ, **Hasan NA**, Taviani E, Haley B, Chun J, Brettin TS, Bruce DC, Detter JC, Han CS, Chertkov O, Huq A, Nair GB, Colwell RR. Genome Sequence of hybrid *V. cholerae* O1 MJ-1236 and comparative genomics of two El Tor variants, B-33 and CIRS101, with *V. cholerae*. J. Bact. 2010. 192(13): 3524-3533
8. Grim CJ, **Hasan NA**, Ali A, Chowdhury WB, Islam A, Rashid MH, Alam M, Morris JG Jr, Huq A, Colwell RR. RNA colony blot hybridization method for enumeration of culturable *Vibrio cholerae* and *Vibrio mimicus* bacteria. Appl Environ Microbiol. 2009 Sep; 75(17): 5439-44.

B. In Process:

Dhaka University Institutional Repository

9. **Hasan NA**, Grim CJ, Lipp EK, Rivera ING, Chun J, Haley BJ, Taviani E, Hoq MM, Munk AC, Brettin TS, Bruce DC, Challacombe JF, Detter JC, Han CS, Eisen JA, Huq A, Colwell RR. 2012. Deep Sea Bacteria Related to Human Pathogenic *Vibrio* Species. *Proc Natl Acad Sci U S A. (In Revision).*
10. **Hasan NA**, Choi SY, Chun J, Hoq MM, Huq A, Cebula TA, Colwell RR. 2012. Mobile Genetic Elements in the Emergence and Evolution of *Vibrio cholera. (In preparation) PLOS Genetics.*
11. **Hasan NA**, Ceccarelli D, Grim CJ, Taviani E, Choi J, Alam M, Hoq MM, Sack RB, Huq A, Colwell RR. 2012. Virulence gene profile of *V. cholerae* strains isolated from clinical samples and aquatic environment in Bangladesh. *Prepared for Appl Environ Microbiol.*
12. **Hasan NA**, Choi SY, Chen J, Grimes J, Hoq MM, Huq A, Cebula TA, Colwell RR. Incidence of non-toxigenic *V. cholerae* non-O1/O139 causing gastroenteritis in the Gulf Coast. *Under submission to Appl Environ Microbiol.*
13. **Hasan NA**, Constatin deMagny G, Grim CJ, Chun J, Alam M, Hoq MM, Sadique A, Huq A, Colwell RR. 2012. *Vibrio cholerae* seasonality as a function of complex changes within microbial communities of the aquatic environment. (*In preparation*) *PNAS.*

2. Abstract:

1. **Hasan NA**, Fadrosh DW, Abdullah AS, Longmire AG, Choi SY, Chen AJ, Huq A, Cebula TA, Colwell RR. Generation of Genetic Diversity in Vibrio cholerae during Limited Laboratory Passage. 2012. Abstr Ann Meet Amer Soc Microbiol 112.
2. Choi SY, **Hasan NA**, Chun J, Hoq MM, Huq A, Cebula TA, Colwell RR. Comparative Genomic MGE finding process streamlines mobile element search. 2012 ASM Biodefense and Emerging Diseases Research Meeting. Washington, USA. PN: 056 (B)
3. Choi SY. **Hasan NA**, Chun J. Huq A. Colwell RR. 2011. Evolution of the Lipopolysaccharide O antigen gene cluster in *Vibrio cholerae*. B-2051, 111TH ASM General Meeting, May 21-25, 2011, New Orleans, LA.
4. Choi SY, **Hasan NA**, Chun J, Hoq MM, Huq A, Cebula TA, Colwell RR. Comparative Genomic MGE finding process streamlines mobile element search. 2012. ASM Biodefense and Emerging Diseases Research Meeting. Washington, USA. PN: 056 (B)
5. **Hasan NA**, Grim CJ, Lipp EK, Rivera ING, Chun J, Haley BJ, Taviani E, Hoq M, Munk AC, Brettin TS, Bruce DC, Challacombe JF, Detter JC, Han CS, Eisen J, Huq A, and Colwell RR. 2010. Deep Sea Bacteria Related to Human Pathogenic *Vibrio* Species. "Vibrios in the Environment" Conference, Biloxi, MS, November 2010.

   This was selected for contributed oral presentation, and subsequently awarded with ***Outstanding Oral Presentation Award in the conference.

Dhaka University Institutional Repository

6. **Hasan NA**, Grim CJ, Haley BJ, Chun J, Alam M, Taviani E, Hoq M, Munk AC, Saunders E, Brettin TS, Bruce DC, Challacombe JF, Detter JC, Han CS, Xie G, Nair GB, Huq A, and Colwell RR. 2010. Comparative Genomics of Clinical and Environmental *Vibrio mimicus*. P: 128. "Vibrios in the Environment" Conference, Biloxi, MS, November 2010.
   *** This poster was awarded with the ***Best Poster Award*** in the conference.
7. **Hasan NA**, Chen A, Harvey S, White D, McKay R, Tarnowski M, Brohawn K, Johnson CN, Grimes DJ, Huq A, Colwell RR. Environmental Investigations of *Vibrio parahaemolyticus, Vibrio vulnificus* and *V. cholerae* in Oyster, Water and Sediment in the Chesapeake Bay. Control Number 3937. 110TH ASM General Meeting to be held from May 23-27, 2010 at San Diego, CA.
8. **Hasan NA**, Alam M, Islam A, Grim CJ, Huq A, Nair GB, Sack RB, Colwell RR. Seasonality and Distribution of Toxigenic *Vibrio cholerae* in Coastal Villages of Bay of Bengal Monitored by Direct PCR. 108TH ASM General Meeting 2008 held at the Boston Convention and Exhibition Center, from June 1through June 5, 2008 in Boston, Massachusetts. Poster Number: N-301
9. **Hasan NA**, Alam M, Islam A, Siddique AK, Nair GB, Sack RB, Huq A, Colwell RR. Antibiotic Resistance Patterns of Clinical and Environmental *Vibrio cholerae* O1 and O139 Isolated from Coastal Bangladesh. 108TH ASM General Meeting 2008 held at the Boston Convention and Exhibition Center, from June 1through June 5, 2008 in Boston, Massachusetts. Poster Number: C-041
10. Grim CJ, **Hasan NA**, Chowdhury WB, Islam A, Alam M, Huq A, Colwell RR. Enumeration of *Vibrio cholerae* employing a fluorescently labelled 16S rRNA probe. VIBRIO 2007 held in Paris, France, 28 November- 1 December 2007. P97 (Session 3), pp135.
11. Alam M, **Hasan NA**, Bhuiyan NA, Siddique AK, Sack RB, Sack DA, Sultana M, Sadique A, Ahmed KU, Nusrin S, Islam A, Chowdhury WB, Huq A, Colwell RR, Nair GB. Antibiotic resistance profiles of *Vibrio cholerae* O1 and O139 isolated from Coastal Ecosystem of Bangladesh. Abstract: 179 (191), pp; 209 11th ASCON held in Bangladesh, 3-6 March 2007.

Dhaka University Institutional Repository

GRANT ACKNOWLEDGEMENTS

Funding for these studies was provided by the Office of the Chief Scientist (USA), and NIAID Microbial Sequencing Centers (grant no. N01-AI-30001, N01-AI-40001 and HHSN2722009000), the KOSEF National Research Laboratory Program (grant no. R0A-2005-000-10110-0), the Department of Homeland Security (grant No. BAA 10-19), National Institutes of Health (grant nos. 1RO1A139129-01 and g2RO1A1039129-11A2), National Oceanic and Atmospheric Administration, Oceans and Human Health Initiative (grant no. S0660009), and the National Science Foundation Emerging Infection Disease Program (grant no. 0813066, and EF-0813066).

Student support (Nur A. Hasan) was provided by the National Institutes of Health (grant nos. 1RO1A139129-01 and g2RO1A1039129-11A2), the Department of Homeland Security (grant No. BAA 10-19), and the National Science Foundation (grant no. 0813066, and EF-0813066).