Characterization of *Vibrio cholerae* O1 strains that trace the origin of Haitian-like genetic traits

Priyanka Ghosh[1], Dhirendra Kumar[2], Goutam Chowdhury[1], Puneeta Singh[2], Prosenjit Samanta[1], Shanta Dutta[1], T. Ramamurthy[1], N. C. Sharma[2], Preety Sinha[3], Yogendra Prasad[4], Sumio Shinoda[5], Asish K. Mukhopadhyay[1*]

[1]Division of Bacteriology, National Institute of Cholera and Enteric Diseases, Kolkata, India; [2]Maharishi Valmiki Infectious Diseases Hospital, Kingsway Camp, Delhi-110 009, India;[3]Department of Zoology, A.N. College, Patna, Bihar, India; [4]Department of Animal Science, MJP Rohilkhand University, Bareilly, UP, India; [5]Collaborative Research Center of Okayama University for Infectious Diseases at NICED, Kolkata, India.

***Corresponding Author: Dr. Asish K. Mukhopadhyay**
Division of Bacteriology
National Institute of Cholera and Enteric Diseases,
P 33, CIT Road, Scheme XM, Beliaghata
Kolkata 700010
India
E-mail: asish_mukhopadhyay@yahoo.com
FAX: 91-33-2370-5066

Running Title: *V. cholerae* O1 in Delhi traces Haitian genetic traits

Key words: Cholera, *Vibrio cholerae, ctxB, tcpA, rtxA*, *gyrA, rstB*, ctxAB promoter

Abstract

Vibrio cholerae O1 is the etiological agent of the severe diarrheal disease cholera. The bacterium has recently been causing outbreaks in Haiti with catastrophic effects. Numerous mutations have been reported in *V. cholerae* O1 strains associated with the Haitian outbreak. These mutations encompass among other the genes encoding virulence factors such as the pilin subunit of the toxin-co-regulated pilus (*tcpA*), cholera toxin B subunit (*ctxB),* repeat in toxins (*rtxA*), and other genes such as the quinolone resistance-determining region (QRDR) of gyrase A (*gyrA*), *rstB* of RS element along with the alteration in the number of repeat sequences at the promoter region of *ctxAB*. Given the numerous genetic changes in those Haitian isolates, we decided to investigate the possible origins of those variations in the Indian subcontinent. Thus, we determined the genetic traits among *V. cholerae* O1 strains in Delhi, India. A total of 175 strains isolated from cholera patients during 2004 to 2012 were analysed in the present study. Our results showed that all the tested strains carried Haitian type *tcpA* (***tcpA***$^{\mathbf{CIRS}}$) and variant *gyrA* indicating their first appearance before 2004 in Delhi. The Haitian variant *rtxA* and *ctxB7* were first detected in Delhi during 2004 and 2006, respectively. Interestingly, not a single strain with the combination of El Tor *rtxA* and *ctxB*7 was detected in this study. The Delhi strains carried four heptad repeats (TTTTGAT) in the CT promoter region whereas Haitian strains carried 5 such repeats. Delhi strains did not have any deletion mutations in the *rstB* like Haitian strains. Overall, our study demonstrates the sequential accumulation of Haitian-like genetic traits among *V. cholerae* O1 strains in Delhi at different time points prior to the Haitian cholera outbreak.

Word count 280

1. Introduction

Cholera, the life threatening ancient intestinal infection still remains as a serious health burden in the developing countries. Cholera is a major health issue in Asian and African countries where sanitation and supply of clean drinking water are limited. The Gram negative organism *Vibrio cholerae* belonging to serogroups O1 and O139 are responsible for this severe dehydrating diarrheal disease (Kaper et al., 1995; Safa et al., 2010). Classical and El Tor are the two biotypes in the taxonomic classification of *V. cholerae* O1. The unique epidemiologic characteristic of this disease is its propensity to occur as outbreaks that may flare-up into epidemics, if not controlled. Seven pandemics of cholera had been encountered since 1817 in the recorded history. Among them the first six were believed to be caused by classical cholera whereas the ongoing seventh pandemic is caused by the El Tor biotype (Safa et al., 2010). In recent years, the emergence and spread of the new variant *V. cholerae* strains have been reported from different parts of the world. They displayed several cryptic modifications in their genome. The Matlab variants from Bangladesh, the Mozambique variants, the altered El Tor type from different parts of the world, and the newly emerged Haitian variants of the El Tor biotype presented genetic hybrids of classical and El Tor biotype. Among all of these hybrids, Haitian variants are the precarious one, considering its association with number of morbidity and mortality of humans.

In late October 2010, a disastrous cholera outbreak occurred in Haiti, a country that had not experienced cholera for more than a century (Piarroux et al., 2011). This episode of cholera placed the disease at the forefront of the global public health agenda. The outbreak has killed around 8000 Haitians, and infected over 600,000 to date (Mukhopadhyay et al., 2014). Many published reports indicate that the strain that caused the outbreak originates from Asian countries (Reimer et al., 2011; Chin et al., 2011; Hendriksen et al., 2011). Recent study on the global phylogeographic patterns of the *V. cholerae* strains isolated from diverse geographic

areas provide strong molecular evidence pointing to a nonindigenous source of the 2010 Haitian cholera outbreak (Eppinger et al., 2014). Based on the whole-genome-based analysis along with the epidemiological observations, it was commonly believed that the disease was brought to Haiti by battalion of Nepalese soldiers serving as United Nations peace keepers (Piarroux et al., 2011; Eppinger et al., 2014). WHO also reported that the re-emergence of cholera was a significant global public health concern and hence endorsed for the execution of an integrated and comprehensive global approach for the prevention of cholera. (WHO, 2010 & 2011).

Numerous mutations have been described among the Haitian variant strains, among them in the genes encoding toxin-co-regulated pilus **(*tcpA*^{CIRS}),** cholera toxin B subunit **(*ctxB7*),** repeat in toxins **(variant *rtxA*),** quinolone resistance-determining region (QRDR) of gyrase A **(*gyrA*^{Ser83Ile}),** *rstB* of RS element along with the alteration in the number of repeat sequences at the promoter region of *ctxAB* (Chin et al., 2011; Talkington et al., 2011; Son et al., 2011). Interestingly, our earlier studies indicated that, even though these genomic variations became highlighted after the Haitian outbreak, they were detected in Kolkata strains that appeared before the Haitian outbreak (Naha et al., 2012; Ghosh et al., 2014a; Ghosh et al., 2014b). Not only in Kolkata, but variant type of *tcpA* (*tcpA*^{CIRS}), *ctxB* (*ctxB7*) and gyrase A (*gyrA*^{Ser83Ile}) alleles were also reported by Nigeria cholera outbreaks (Marin et al., 2013). Variant *ctxB* allele (*ctxB7)* was also found in South-Western India and Mexico during 2012 and 2013, respectively (Kumar et al., 2014; Diaz-Quinonex et al., 2014). A recent study reported that though *ctxB1* was replaced by *ctxB7* allele in Bangladesh but *ctxB1* outcompetes the *ctxB7* allele since 2013 (Rashid et al., 2015). The current global scenario of the cholera outbreaks and the emergence of this new pathogenic variant motivated us to investigate the genetic characteristics of the *V. cholerae* isolates from Delhi, India. Characterization was based on single nucleotide polymorphism (SNP). This study presented the sequential accumulation of

Haitian genetic traits among *V. cholerae* O1 strains in Delhi at different time frames, which may have a severe threat to the public health and the accumulation events occurred long before the Haitian cholera outbreak. Studies of such changes along with the evolution of new variant *V. cholerae* strains in different parts of the world may have implications to combat cholera.

2. Material and Methods

2.1. Bacterial cultures

V. cholerae O1 strains isolated in Delhi between 2004 and 2012 was characterized in this study. 14 strains from 2004, 13 strains from 2005, 20 from 2006, 39 from 2007, 23 from 2008, 13 from 2009, 11 from 2010, 15 from 2011 and 27 strains from 2012 were considered. *V. cholerae* strains examined for this study were selected from the strain repository of Maharishi Valmiki Infectious Diseases Hospital, Delhi, India and were isolated from the hospitalized cholera patients. After incubation for 18h in Luria Bertani broth (Becton Dickinson, Sparks, MD, USA), all the strains were then streaked on Luria agar (Becton Dickinson) plates. Serology of these strains was confirmed by the slide agglutination with O1 specific polyvalent antiserum and serotype specific antisera (Becton Dickinson). In this study, *V. cholerae* O1 strains EL-1786 (Ogawa), N16961 (Inaba) and O395 (Ogawa) were used as standard strains for the Haitian, El Tor and classical types, respectively.

2.2. Template preparation for PCR

Bacterial cell lysate was used as source for DNA template in all the PCR assays. One loopful of an overnight culture from LA plate was suspended in 200 µl of Tris-EDTA buffer (pH 8.0). Then mixture of phenol: chloroform: isoamyl alcohol (25:24:1) saturated with 10 mM Tris and 1mM EDTA (Sigma-Aldrich, St Louis, MO, USA) was used for lysing the cell by

vigorous vortexing. After centrifugation at 12,000 rpm for 15 min, the supernatant was collected carefully and mixed with 100 µl of mixture of chloroform: isoamyl alcohol (24:1) and centrifuged for 15 min at 12,000 rpm. The supernatant that contained the bacterial DNA was **diluted to 50ng/µl concentration and 3 µl of that diluted DNA was** used as template for PCR analysis.

2.3. Mismatch Amplification Mutation Assay (MAMA)

Utilizing the single base mutation, a simple PCR based method has been used in this study. We used our previously designed allele-specific either forward or reverse primer containing mismatch at the 3' ends ((Table 1). Three individual primers, which include one reverse primer common for both El Tor and variant type of *tcpA* allele (tcpA El -Rev) and two forward primers (tcpAF1 and tcpA F'2) specific for El Tor and variant type, respectively were used (Table 1). Difference between variant and classical *ctxB* alleles was determined with the help of two forward (ctxBF3 and ctxBF4) and one common reverse primer (ctxB Cla-Rev) (Table 1). Further, presence of different alleles of *rtxA* was assayed with one common forward (rtxAF) and two reverse primers (rtxA-R1 and rtxA-R2) specific for El Tor and *rtxA*-null mutant respectively (Table 1). Allele specific reverse primers (gyrA-R1 and gyrA-R2) and a common forward primer (gyrA-F) were used for variant and El Tor type of *gyrA* alleles, respectively (Table 1). Deletion in the *rstB* gene was determined by using rstBF1 and rstBR1 primers (Table 1).

2.4. Sequencing Analysis

PCR amplicons from 16 randomly selected strains of *V. cholerae* O1 were subjected to nucleotide sequence of the *tcpA*, *ctxB*, *rtxA*, *gyrA*, *rstB* and *ctxAB* promoter sequence for confirmation. PCR amplification was carried out in 20 µl reaction mixture. The Qiaquick PCR purification kit (QIAGEN, Hilden, Germany) was used to purify the amplified products.

The purified PCR product obtained was quantitated on 1% agarose gel. For the cycle sequencing reaction, specific amount of DNA present in purified PCR product was added as template depending on the size of the product to be sequenced. The PCR primer pair used to generate the amplicon was generally used for sequencing of both the strands of the amplicon. The sequencing mix was obtained from Big Dye Terminator v3.1 kit (Applied Biosystems, USA) and consisted of Big Dye® Terminator 3.1 Sequencing Buffer (5X) and ready to use reaction mixture (2.5X). The terminators used were dideoxynucleotides labelled with the fluorescent dyes. After purification of cycle sequencing product both the strands of amplicons were sequenced in an automated sequencer. (ABI PRISM 3100 Genetic Analyzer, Applied Biosystems, Foster City, CA, USA). The raw sequencing data was analysed using the ABI PRISM DNA Sequencing Analysis Software (PE Applied Biosystems). The nucleotide sequence data was assembled to generate sequence contig for each gene from respective forward and reverse sequence read using software (DNASTAR, USA). Then searches for identical sequences were performed using the Basic Local Alignment Search Tool (BLAST) programme available on the National Center for Biotechnology Information network server.

2.5. Accession numbers of the gene sequences

The sequences determined in the present study were deposited in GenBank under accession numbers **KP187620-KP187623** for *tcpA*, **KP717075-KP717077** for *ctxB*, **KP717072-KP717074** for *rtxA*, **KP717065-KP717068** for *gyrA*, **KP717069-KP717071** for *rstB* and **KP729599-KP729601** for *ctxAB* promoter repeat.

3. Results

3.1. Variant tcpA allele in V. cholerae isolates from Delhi

In this study, we investigated the emergence of variant *tcpA* allele in Delhi, using previously developed allele specific primers that can broadly discriminate *tcpA* of variants and El Tor strains. *V. cholerae* O1 clinical strains isolated during 2004-2012 in Delhi were screened for understanding the emergence and dissemination of the variant *tcpA* allele **(Fig.1-a1 and a2)**. Sequencing of the 16 representative strains with variant *tcpA* was performed to reconfirm the PCR results. The amino acid sequences of all these strains were found to be identical to the deduced amino acid sequence of the TcpA of the Haitian reference strain EL-1786, containing an asparagine to serine substitution at the 89th position of the sequence containing the signal peptide. Further, our retrospective analysis demonstrated complete replacement of El Tor *tcpA* in Delhi **at least** from the year 2004 and not a single strain with El Tor *tcpA* was detected till 2012.

3.2. ctxB7 allele in El Tor variant strains of Delhi

Existence of SNP at the 58th nucleotide of the *ctxB* gene of the Haitian outbreak strain is known as the 7th allele of *ctxB* (Chin et al., 2011) and this has resulted the emergence of asparagine instead of histidine at the 20th amino acid position of the signal region of CTB. Our allele specific PCR assay has effectively differentiated the classical (*ctxB*1) and variant allele of *ctxB* (*ctxB*7) **(Fig. 1-b1 and b2)**. PCR result was further confirmed by sequencing analysis using different set of primers. We screened **175** *V. cholerae* O1 strains covering different months of each year from 2004 to 2012 using this PCR assay. This analysis revealed the first appearance of the strains harbouring *ctxB7* in Delhi during 2006 with a sudden increase in the isolation rate in 2007 followed by sharp decline in 2008. However, the percentage of the O1 strains with *ctxB7* allele started to surge from 2009 and **around 80%** Delhi strains carried *ctxB7* allele in the year 2012 **(Fig.2)**.

3.3. Prevalence of rtxA-null mutants in Delhi isolates

Dolores and Satchell (2013) described occurrence of a single base mutation in the *rtxA* gene within the recently emerged Haitian outbreak strains which encodes a multifunctional auto-processing RTX toxin (Dolores & Satchell, 2013). This mutation generates a premature stop codon which introduces a non-functional RTX toxin. MAMA based PCR assay was further used to trace the appearance of the *rtxA*-null mutation among the *V. cholerae* Delhi strains. This assay successfully differentiated the El Tor and variant type of *rtxA* **(Fig. 1-c1 and c2)**.

In the sequencing analysis, we found that an *rtxA*-null mutation (variant *rtxA*) in the Delhi strains were isolated during 2004. After the appearance and gradual increase of the *rtxA*-null-mutant up to 2007, El Tor type *rtxA* allele further became dominant in 2008. From **2009** onwards, the El Tor *rtxA* was replaced with a higher percentage of variant *rtxA* **(Fig. 3)**. Our study also demonstrated the appearance of *ctxB7* allele with *rtxA* null background.

3.4. Genetic features of gyrA gene within V. cholerae strains isolated from Delhi

A Ser83→Ile substitution was documented in the *gyrA* of the Haitian strains. This substitution is associated with quinolones resistance in clinical *V. cholerae* (Hasan et al., 2012). Occurrence of the variant type of *gyrA* allele within the *V. cholerae* strains from Delhi, India was determined in this study. The allele specific primers based PCR assay successfully differentiated the two different alleles of *gyrA* **(Fig. 1-d1 and d2).** Another set of primer was used for sequencing the representative Delhi strains and to reconfirm the PCR results. Our study indicated that a variant type of *gyrA* allele has been appeared and completely displaced the El Tor *gyrA* allele **at least from the year** 2004.

3.5. Survey of rstB allele in Delhi isolates

Hasan NA *et al*. (2012) demonstrated a GTA deletion at nucleotide positions 77-79 in the *rstB* of the Haitian outbreak strain. In order to examine these deletions if any, a specific PCR

assay was done. All of the Delhi isolates from the year 2004-2012 gave amplicon at 160 bp regions **(Fig. 1-e), implying no GTA deletion** which was confirmed by sequencing analysis.

3.6. ctxAB promoter repeats in V. cholerae O1 Delhi isolates

Variation in the number of *ctxAB* promoter repeats (TTTTGAT) has been reported earlier. Whole genome sequence (WGS) analysis revealed five copies of the heptads repeat in the ToxR binding region in Haitian outbreak strain El-1786 (Son et al., 2011)**.** Our study showed presence of four heptad repeats in Delhi strains **(Fig. 4).**

3.7. Occurrence of V. cholerae strains with different allelic combination at Delhi

Among six genetic loci studied, there was no variation in *gyrA*, *rstB* and the promoter region of *ctxAB* in *V. cholerae* strains isolated between 2004 and 2012 from Delhi. Based on the changes in *tcpA*, *rtxA* and *ctxB*, we could divide the strains in three different genetic combinations namely, strains with Haitian *tcpA* (***tcpA***$^{\mathbf{CIRS}}$) _ Classical *ctxB* (*ctxB*1) _ El Tor *rtxA;* ***tcpA***$^{\mathbf{CIRS}}$ _ *ctxB*1_ variant *rtxA* and ***tcpA***$^{\mathbf{CIRS}}$_ Haitian *ctxB* (*ctxB*7) _ variant *rtxA*. *V. cholerae* O1 strains with ***tcpA***$^{\mathbf{CIRS}}$ _ *ctxB*1 _ El Tor *rtxA* were isolated mostly in the year 2004. During 2005 the strains with the combination of ***tcpA***$^{\mathbf{CIRS}}$ _ *ctxB*1_ variant *rtxA* became dominant. They prevailed as dominant during 2006 also. Interestingly, among the tested strains, there was not a single isolate with the combination of El Tor *rtxA* and Haitian *ctxB*. The strain with the combination of ***tcpA***$^{\mathbf{CIRS}}$ _ *ctxB*7 _ variant *rtxA* was first isolated in Delhi for the first time in the year 2006 and became dominant during 2007. But the strains with the combination of ***tcpA***$^{\mathbf{CIRS}}$ _ *ctxB*1_ El Tor *rtxA* further became dominant during 2008. Then the strains containing ***tcpA***$^{\mathbf{CIRS}}$_ *ctxB*7_ variant *rtxA* were isolated with a higher percentage from 2010 onwards **(Figure 5)**.

4. Discussion

The Haitian cholera outbreak raised attention in the global incidence of cholera. It also provided researchers with an opportunity to study the genome of several strains of *V. cholerae* from different regions. *V. cholerae* strain responsible for the cholera epidemic in Haiti had similar phenotypic and genetic properties of the seventh pandemic El Tor O1 strains (Hasan et al., 2012; Nair et al., 2002; Siddique et al., 2010). The Genome wide SNP analysis of Haitian cholera outbreak strains presented a different depiction with the canonical El Tor strains. Several studies found variations in major virulence genes, such as *tcpA*, *ctxB*; *rtxA*, *gyrA*, *rstB* of CTX phage and *ctxAB* promoter repeats (Talkington et al., 2011; Son et al., 2011; Dolores & Satchell, 2013). WGS of Haitian strain was compared with two strains from Bangladesh and one isolated in South America including sequence of 23 different strains of *V. cholerae* available in the public domain (Chin et al., 2011). Even though there is a remarkable similarity between El Tor and Haitian strains, large differences were observed with the South American isolates (Chin et al., 2011). Haitian strains were found to be identical and most closely related to the Indian sub continental strains (Mutreja et al., 2011). With the help of single nucleotide polymorphism (SNP) analysis of the *V. cholerae* genome, it was also considered that the seventh cholera pandemic was transmitted worldwide via three major overlapping waves from the Ganges Delta of the Bay of Bengal, the ancient home of Asiatic cholera (Mutreja et al., 2011). Among the 13 recognized *ctxB* genotypes (based on amino acid substitution), *ctxB*3 was represented by wave 1, *ctxB*1 by wave 2 and in early wave 3 strains. Whereas *ctxB*7 was solely found in the wave 3 strains, such as the 2010 Haiti cholera outbreak strains (Kim *et al.*, 2015). The origin of the Haitian outbreak further became multifaceted with WGS analysis that indicated Nepal as the likely origin of this Haitian outbreak. (Hendriksen et al., 2011). In a recent study, comparisons of 108 *V. cholerae* genomes from Thailand, Bangladesh, Nepal, Cameroon, India, Pakistan, and Benin specified

that Haitian strains were not only nearly identical to the isolates from Nepal but also the Nepal-Haiti isolates were clearly distinct from isolates circulating elsewhere in the world (Katz et al., 2013). Whole genome phylogeny and core genome SNPs have shown the genetic relations of Haitian outbreak strain with strains originated from India and Cameroon (Reimer et al., 2011). It was also shown that that Haitian isolates were poorly transformable even in a laboratory, casting doubt on the possibility that non-O1 *V. cholerae* may have contributed to the evolution of this outbreak (Katz et al., 2013). So, horizontal gene transfer probably plays no role in the evolution of this group of *V. cholerae*.

A novel single nucleotide polymorphism (SNP) at the nucleotide position 266 (amino acid 89) of the *tcpA* gene was identified in Haitian cholera outbreak strains (Chin et al., 2011; Talkington et al., 2011; Grim et al., 2010). Our study indicated that this variant *tcpA* allele was also present in Delhi strains since 2004. We were unable to find a single strain with El Tor *tcpA* allele from 2004 onwards. Another interesting observation was the presence of a unique mutation at the 58th nucleotide of *ctxB* gene in Delhi strains from the year 2006. This variant *ctxB* (*ctxB7*) allele was recognized after the Haitian cholera outbreak.

It should be noted that acquisition of Haitian *ctxB* and *tcpA* do not always occur in tandem. This *ctxB7* allele was also reported from other parts of India (Orissa and Kolkata) and West African countries of Nigeria and Cameroon (Naha et al., 2012; Kumar et al., 2014; Chio et al., 2010; Quilici et al., 2010). Furthermore, Haitian strains showed mutation in *rtxA* that encodes a stop codon which ultimately leads to deactivation of RTX (Dolores & Satchell, 2013). Presence of variant *rtxA* (*rtxA*-null mutation) in the altered El Tor has been reported from Bangladesh in 1999, India in 2004 and Haiti in 2010 (Hendriksen et al., 2011; Ghosh et al., 2014; Dolores & Satchell, 2013). The classical strains are also known to have a deletion that removed >7 kb of the *rtx* locus, thereby deactivating the RTX (Lin et al., 1999). We noticed

292 the first appearance of this *rtxA* null mutation in Delhi strains during 2004. Our study also
293 highlighted that all the Delhi strains with Haitian *ctxB* allele has acquired this null mutation.
294 Based on the chronological genetic events, we assume that this *rtx*-null mutant may be the
295 genetic background for the subsequent appearance of the Haitian *ctxB* allele.
296 Epidemiological advantage of this *rtxA* null background is still unclear. It can be explained as
297 the toxin is very large, thus it needs more energy. High energy expenditure may be
298 detrimental to growth and rapid growth is necessary for increased dissemination. A serine to
299 isoleucine substitution at the 83 amino acid position of *gyrA*, which is associated with
300 quinolone resistance in clinical *V. cholerae* strains have been reported in Haitian strains
301 (Hasan et al., 2012). Similar observation was also reported from India, Nigeria and Cameroon
302 (Quilici et al., 2010; Baranwal et al., 2002; Garg et al., 2000; Islam et al., 2009). We found
303 that all the Delhi strains had such mutation in *gyrA*. An additional feature of the Haitian strain
304 was the GTA deletion in the *rstB* (Hasan et al., 2012), which is a unique feature of classical
305 strains of *V. cholerae* (Chio et al., 2010). Such deletion was not detected in any of the Delhi
306 strains. Further, sequence based analysis has demonstrated the presence of four copies ToxR
307 binding repeats (TTTTGAT) in Delhi strains in comparison with the five copies of repeats in
308 Haitian isolates.

309 Finally, our study demonstrated that the circulating *V. cholerae* O1 strains in Delhi presented
310 the sequential accumulation of Haitian genetic traits at different time frames though some
311 dissimilarity are present. It has been believed that the distinct genetic traits of the new variant
312 strains probably increases their relative fitness, perhaps as a consequence of increased
313 pathogenicity (Ghosh-Banerjee et al., 2010). Considering several reports on Haitian cholera
314 epidemic, our results also highlighted existence of a worldwide genomic evolution of recent
315 *V. cholerae* strains which may become a severe threat to the public health. Further studies

316 regarding the cause of such genetic changes in *V. cholerae* are important to detect because it
317 can guide our strategies for the battle against cholera more efficiently.

318

319

320 **Acknowledgements**

321 This study was supported in part by the Indian Council of Medical Research (ICMR),
322 Government of India; Japan Initiative for Global Research Network on Infectious Diseases
323 (J-GRID) of the Japan Agency for Medical Research and Development (AMED) and
324 National Institute of Infectious Diseases (NIID), Japan. P.G. and P.S. acknowledge the
325 Inspire fellowship (IF110706) and CSIR fellowship [09/482(0060)/2014-EMR-I] received
326 from the Department of Science & Technology (DST), India and Council of Scientific and
327 Industrial Research, India, respectively.

328

References:

Baranwal, S., Dey, K., Ramamurthy, T., Nair, G. B. & 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.

Chin, C. S., Sorenson, J., Harris, J. B., Robins, W. P., Charles, R. C., Jean-Charles, R. R., Bullard, J., Webster, D. R., Kasarskis, A., Peluso, P. & other authors (2011). The origin of the Haitian cholera outbreak strain. *N Engl J Med.* **364,** 33-42.

Chio, S. Y., Lee, J. H., Kim, E. J., Lee, H. R., Jeon, Y. S., Von-Seidlein, L., Deen, J., Ansaruzzaman, M., Lucas, G. M., Barreto A. & other authors (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.

Diaz-Quinonez, A., Hernandez-Monroy, I., Montes-Colima, N., Moreno-Perez, A., Galicia-Nicolas, A., Martínez-Rojano, H., Carmona-Ramos, C., Sánchez-Mendoza, M., Rodríguez-Martínez, J. C., Suárez-Idueta, L. & other authors (2014). Outbreak of *Vibrio cholerae* serogroup O1, serotype Ogawa, biotype El Tor strain--La Huasteca Region, Mexico, 2013. *MMWR Morb Mortal Wkly Rep* **63,** 552-553.

Dolores, J. & Satchell, K. J. (2013). Analysis of *Vibrio cholerae* genome sequences reveals unique *rtxA* variants in environmental strains and an *rtxA*-null mutation in recent altered El Tor isolates. *MBio* **4,** e00624.

Eppinger, M., Pearson, T., Koenig, S.S., Pearson, O., Hicks, N., Agrawal, S., Sanjar, F., Galens, K., Daugherty, S., Crabtree, J. & other authors (2014). Genomic epidemiology of the Haitian cholera outbreak: a single introduction followed by rapid, extensive, and continued spread characterized the onset of the epidemic. *MBio* **5(6)**, e01721.

353 **Garg, P., Chakraborty, S., Basu, I., Datta, S., Rajendran, K., Bhattacharya, T.,**
354 **Yamasaki, S., Bhattacharya, S. K., Takeda, Y., Nair, G. B. & other authors (2000)**.
355 Expanding multiple antibiotic resistances among clinical strains of *Vibrio cholerae* isolated
356 from 1992–7 in Calcutta, India. *Epidemiol Infect* **124,** 393–399.
357 **Ghosh, P., Naha, A., Basak, S., Ghosh, S., Ramamurthy, T., Koley, H., Nandy, R. K.,**
358 **Shinoda, S., Watanabe, H., Mukhopadhyay, A. K. (2014).** Haitian variant tcpA in *Vibrio*
359 *cholerae* O1 El Tor strains in Kolkata, India. *J Clin Microbiol* **52**, 1020-1021.
360 **Ghosh, P., Naha, A., Pazhani, G. P., Ramamurthy, T. & Mukhopadhyay, A. K. (2014**).
361 Genetic traits of *Vibrio cholerae* O1 Haitian isolates that are absent in contemporary strains
362 from Kolkata, India. *PLoS One* **9**, e112973.
363 **Ghosh-Banerjee, J., Senoh, M., Takahashi, T., Hamabata, T., Barman, S., Koley, H.,**
364 **Mukhopadhyay, A. K., Ramamurthy, T., Chatterjee, S., Asakura, M. & other authors**
365 **(2010)**. Cholera toxin production by El Tor variant of *Vibrio cholerae* O1 as compared to
366 prototype El Tor and classical biotypes. *J Clin Microbiol* **48**, 4283-4286.
367 **Grim, C. J., Hasan, N. A., Taviani, E., Haley, B., Chun, J., Brettin, T. S., Bruce, D. C.,**
368 **Detter, J. C., Han, C. S., Chertkov, O. & other authors (2010).** Genome sequence of
369 hybrid *Vibrio cholerae* O1 MJ-1236, B33, and CIRS101and comparative genomics with *V.*
370 *cholerae*. *J Bacteriol* **192,** 3524 –3533.
371 **Hasan, N. A., Choi, S.Y., Eppinger, M., Clark, P. W., Chen, A., Alam, M., Haley, B. J.,**
372 **Taviani, E., Hine, E., Su. Q., Tallon, L. J. & other authors (2012).** Genomic diversity of
373 2010 Haitian cholera outbreak strains. *Proc Natl Acad Sci* **109,** E2010-2017.
374 **Hendriksen, R. S., Price, L. B., Schupp, J. M., Gillece, J. D., Kaas, R. S., Engelthaler, D.**
375 **M., Bortolaia, V., Pearson, T., Waters, A.E., Upadhyay, B. P. & other authors (2011).**
376 Population genetics of *Vibrio cholerae* from Nepal in 2010: evidence on the origin of the
377 Haitian outbreak. *MBio* **2,** e00157-11.

Islam, M. S., Midzi, S. M., Charimari, L., Cravioto, A. & Endtz, H. P. (2009). Susceptibility to fluoroquinolones of *Vibrio cholerae* O1 isolated from diarrheal patients in Zimbabwe. *JAMA* **302**, 2321-2332.

Kaper, J. B., Morris, J. G. Jr. & Levine, M. M. (1995). Cholera. *Clin Microbiol Rev* **8,** 48–86.

Katz, L. S., Petkau, A., Beaulaurier, J., Tyler, S., Antonova, E. S. & Turnsek, M. A. (2013). Evolutionary dynamics of *Vibrio cholerae* O1 following a single-source introduction to Haiti. *MBio* **4,** e00398-13.

Kim, E. J., Lee, C. H., Nair, G. B. & Kim, D. W. (2015). Whole-genome sequence comparisons reveal the evolution of *Vibrio cholerae* O1. *Trends Microbiol* **23,** 479–489.

Kumar, P., Mishra, D. K., Deshmukh, D. G., Jain, M., Zade, A. M., Ingole, K. V., Goel, A. K.& Yadava, P. K. (2014). *Vibrio cholerae* O1 Ogawa El Tor strains with the *ctxB7* allele driving cholera outbreaks in south-western India in 2012. *Infect Genet Evol* **25,** 93-96.

Lin, W., Fullner, K. J., Clayton, R., Sexton, J. A., Rogers, M. B., Calia, K. E., Calderwood, S. B., Fraser. C. & Mekalanos, J. J. (1999). Identification of a *Vibrio cholerae* RTX toxin gene cluster that is tightly linked to the cholera toxin prophage. *Proc Natl Acad Sci* **96,** 1071–1076.

Marin, M. A., Thompson, C. C., Freitas, F. S., Fonseca, E. L., Aboderin, A. O., Zailani, S. B., Quartey, N. K., Okeke, I. N. & Vicente, A. C. (2013). Cholera outbreaks in Nigeria are associated with multidrug resistant atypical El Tor and non-O1/non-O139 *Vibrio cholerae*. *PLoS Negl Trop Dis* **7,** e2049.

Mukhopadhyay, A. K., Takeda, Y & Nair, G. B. (2014). Cholera Outbreaks in the El Tor Biotype Era and the Impact of the New El Tor Variants. *Curr Top Microbiol Immunol* **379**, *p17-48. In G.B. Nair and Y Takeda (eds), Cholera Outbreaks.* Springer Newyork.

Mutreja, A., Kim, D. W., Thomson, N. R., Connor, T. R., Lee, J. H., Kariuki, S., Croucher, N. J, Choi, S. Y., Harris, S. R. & Lebens, M. (2011). Evidence for several waves of global transmission in the seventh cholera pandemic. *Nature* **477,** 462-465.

Naha, A., Pazhani, G. P., Ganguly, M., Ghosh, S., Ramamurthy, T., Nandy, R. K, Nair, G. B., Takeda, Y., Mukhopadhyay, A. K. (2012). Development and evaluation of a PCR assay for tracking the emergence and dissemination of Haitian variant *ctxB* in *Vibrio cholerae* O1 strains isolated from Kolkata, India. *J Clin Microbiol* **l50,** 1733-1736.

Nair, G. B., Faruque, S. M., Bhuiyan, N. A., Kamruzzaman, M., Siddique, A. K. & Sack, D. A. (2002). 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* **40,** 3296-3299.

Piarroux, R., Barrais, R., Faucher, B., Haus, R., Piarroux, M., Gaudart, J., Magloire, R. & Raoult, D. (2011). Understanding the cholera epidemic, Haiti. *Emerg Infect Dis* **17,** 1161-1167.

Quilici, M. L., Massenet, D., Gake, B., Bwalki, B.& Olson, D. M. (2010). *Vibrio cholerae* O1 variant with reduced susceptibility to ciprofloxacin, Western Africa. *Emerg Infect Dis* **16,** 1804-1805.

Rashid, M. U., Rashed, S. M., Islam, T., Johura, F. T., Watanabe H., Ohnishi, M. & Alam, M. (2016). CtxB1 outcompetes CtxB7 in *Vibrio cholerae* O1, Bangladesh. *J Med Microbio.* **65,** 101-103.

Reimer, A. R., Van Domselaar, G., Stroika, S., Walker, M., Kent, H., Tarr, C., Talkington, D., Rowe, L., Olsen-Rasmussen, M., Frace, M. & other authors (2011). Comparative genomics of *Vibrio cholerae* from Haiti, Asia, and Africa. *Emerg Infect Dis* **17,** 2113-2121.

Safa, A., Nair, G. B. & Kong, R. Y. C. (2010). Evolution of new variants of *Vibrio cholerae* O1. *Trends Microbiol* **18,** 46–54.

Siddique, A. K., Nair, G. B., Alam, M., Sack, D. A., Huq, A., Nizam, A., Longini, I. M Jr., Qadri, F., Faruque, S. M., Colwell, R. R. & other authors (2010). El Tor cholera with severe disease: a new threat to Asia and beyond. *Epidemiol Infect* **138,** 347-352.

Son, M. S., Megli, C. J., Kovacikova, G., Qadri, F. & Taylor, R.K. (2011). Characterization of *Vibrio cholerae* O1 El Tor biotype variant clinical isolates from Bangladesh and Haiti, including a molecular genetic analysis of virulence genes. J Clin *Microbiol* **49,** 3739-3749.

Talkington, D., Bopp, C., Tarr, C., Parsons, M. B., Dahourou, G., Freeman, M., Joyce, K., Turnsek, M, Garrett, N., Humphrys, M. & other authors (2011). Characterization of toxigenic *Vibrio cholerae* from Haiti, 2010–2011. *Emerg Infect Dis*. **17,** 2122–2129.

World Health Organization. (2010-2011). Cholera .*Wkly Epidemiol Rec* **86,** 325–340.

440

441 **Table 1.** Primer sequences, amplicon size and annealing temperature used in PCR assays

Primer	Sequence (5′-3′)	Amplicon (bp)	Annealing (°C)	Reference
tcpA-F1	ccagctaccgcaaacgcaGA	167	56	Ghosh *et al*., 2014
tcpA-F'2	ccagctaccgcaaacgcaGG			Ghosh *et al*., 2014
tcpA-El-Rev	ccgactgtaattgcgaatgc			Ghosh *et al*., 2014
tcpA-(F)	ggtgactttgtgtggttaaa	895	55	Ghosh *et al*., 2014
tcpA-(R)	atcgcctccaataatccgac			Ghosh *et al*., 2014
ctxB-F3	gttttactatcttcagcatatgcGA	191	58	Naha *et al*., 2012
ctxB-F4	gttttactatcttcagcatatgcGC			Naha *et al*., 2012
(Rv-Cla)	cctggtacttctacttgaaacg			Morita*et al*. 2008
ctxB-F	ggttgcttctcatcatcgaaccac	460	56	Olsvik*et al*. 1993
ctxB-R	gatacacataatagaattaaggat			Olsvik*et al*. 1993
rtxAR1	tgtgaaccacgtctgCC	187	54	Ghosh *et al*., 2014
rtxAR2	tgtgaaccacgtctgCT			Ghosh *et al*., 2014
rtxAF	atcggaatgagtgagaaagacc			Ghosh *et al*., 2014
rtxAF'	tactttaatggtaaccgcgct	422	54	Ghosh *et al*., 2014
rtxAR	cattgtcactgtacttacgtc			Ghosh *et al*., 2014
gyrAR1	gatggtgtcgtaaaccgcTA	177	60	Ghosh *et al*., 2014
gyrAR2	gatggtgtcgtaaaccgcTC			Ghosh *et al*., 2014
gyrAF	tgctcttcctgatgtgcgtgatg	411	61.8	Ghosh *et al*., 2014
gyrAR	ttgatcagcaggttcgggatc			Ghosh *et al*., 2014
rstBF1	attctgaaggggtgagtCgta	160	58	Ghosh *et al*., 2014
rstBR2	ctggtcatcgcgtcactggat			Ghosh *et al*., 2014
cepF	gccaatcacggtaacaatca	820	48	Ghosh *et al*., 2014
rstAR	aggaaattcacgacgattcac			Ghosh *et al*., 2014
ZotF(S)	cgagctaccgctacaaggtgcta	470	55	Naha *et al*. 2013
ctxAR(S)	cgtgcctaacaaatcccgtctgag			Naha *et al*. 2013

442

443

Figure Legends:

Figure 1: PCR based assay for *tcpA, ctxB, rtxA, gyrA* and *rstB* alleles in representative *V. cholerae* O1 Delhi isolates.

MAMA-PCR to detect the type of *tcpA* allele in representative *Vibrio cholerae* O1 strains of Delhi using primers of (tcpAF1/tcpAR) for El Tor **(a1)** and (tcpAF'2/tcpAR) for Variant **(a2).** Lanes 1-9 represent 624 (2004), 463 (2005), 319 (2006), 3621 (2007), 6200 (2008), 15468 (2009), 23855 (2010), 27174 (2011), 31390 (2012), respectively. To detect the type of *ctxB* gene in representative *Vibrio cholerae* O1 strains of Delhi using primers of (ctxBF4/Cla-Rv) for classical **(b1)** and (ctxBF3/Cla-Rv) for Haitian (*ctxB7* allele) **(b2)**. Lanes 1-9 represent 757 (2004), 684 (2005), 1854 (2007), 463 (2005), 7252 (2006), 1851 (2007), 13856 (2009), 23979 (2010), 27122 (2011), respectively. To detect the type of *rtxA* allele in representative *Vibrio cholerae* O1 strains of Delhi using primers (rtxAF/rtxAR1) for El Tor **(c1)** and (rtxAF/rtxAR2) for Variant **(c2)**. Lanes 1-9 represent 831 (2004), 194 (2004), 6479 (2008), 27115 (2011), 22237 (2010), 499 (2005), 26976 (2011), 31267 (2012), 183 (2005), respectively. To detect the type of *gyrase A* allele in representative *Vibrio cholerae* O1 strains of Delhi using primers (gyrAF/gyrAR2) for El Tor **(d1)** and (gyrAF/gyrAR1) for Variant **(d2)**. Lanes 1-9 represent 624 (2004), 463 (2005), 319 (2006), 3621 (2007), 6200 (2008), 15468 (2009), 23855 (2010), 27174 (2011), 31390 (2012) respectively. To detect the type of *rstB* allele in representative *Vibrio cholerae* O1 strains of Delhi using primers rstBF1-rstBR2 **(e)**. Lanes 1-9 represent 624 (2004), 463 (2005), 319 (2006), 3621 (2007), 6200 (2008), 15468 (2009), 23855 (2010), 27174 (2011), 31390 (2012). In all cases expect *ctxB* genotype assessment, N16961 (Lane 10) and El-1786 (Lane 11) were used as control for El Tor and variant strains, respectively. But in case of *ctxB* genotype, O395 (Lane 10) and El-1786 (Lane 11) were used as control for Classical and variant strain. The extreme right lane contains a 100-bp size ladder (New England Biolab Inc., Beverly, MA, USA).

Figure 2: Retrospective analysis of *ctxB* allele in *V. cholerae* Delhi isolates.

Occurrence of *ctxB* allele type in Delhi *Vibrio cholerae* O1 strains from 2004 to 2012. A *V. cholerae* O1 strains with Haitian type of *ctxB* was isolated in Delhi for the first time in the year 2006 and interestingly there was a sudden increase in the isolation profile of *V. cholerae* O1 strains with *ctxB7* allele during the year 2007 encompassing a sudden decrease in the year 2008. After this declining episode of *ctxB7* allele, the percentage of the O1 isolates with *ctxB7* allele started to increase from 2009 and **around 80%** Delhi strains carried *ctxB7* allele in the year 2012.

Figure 3: Isolation profile of variant *rtxA* in Delhi.

Occurrence of *rtxA* allele in Delhi *Vibrio cholerae* O1 strains from 2004 to 2012. *V. cholerae* O1 strains with variant *rtxA* was first time isolated in Delhi during 2004. After the appearance and gradual surge of the *rtxA*-null- mutant up to 2007, El Tor *rtxA* allele further became dominant during 2008. Then the El Tor *rtxA* was slowly replaced by the variant *rtxA* since **2009.**

Figure 4. Sequence based analysis of *ctxAB* promoter repeats.

Schematic representation of the promoter region the *ctxAB* operon of *Vibrio cholerae* O1 El Tor variant strains isolated from Delhi and Haiti and its comparison with classical and El Tor strains. The Delhi strains contained four heptad repeats (TTTTGAT) in their CT promoter region whereas Haitian strain carried 5 heptad repeats.

Figure 5: Isolation profile of *V. cholerae* strains with different genetic makeup in Delhi during 2004 to 2012.

Strains with three different genetic combinations namely, strains with Haitian *tcpA* _ Classical *ctxB* (ctxB1) _ El Tor *rtxA;* Haitian *tcpA* _ ctxB1_ variant *rtxA* and Haitian *tcpA* _

Haitian *ctxB* (ctxB7) _ variant *rtxA,* were isolated during this period. Any strain with the combination of Haitian *tcpA* _ ctxB7 _ El Tor *rtxA* was not detected during the study period.

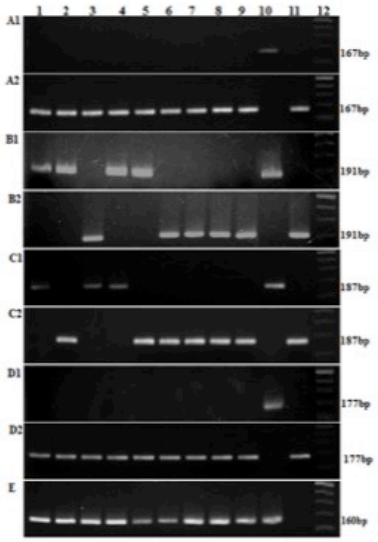
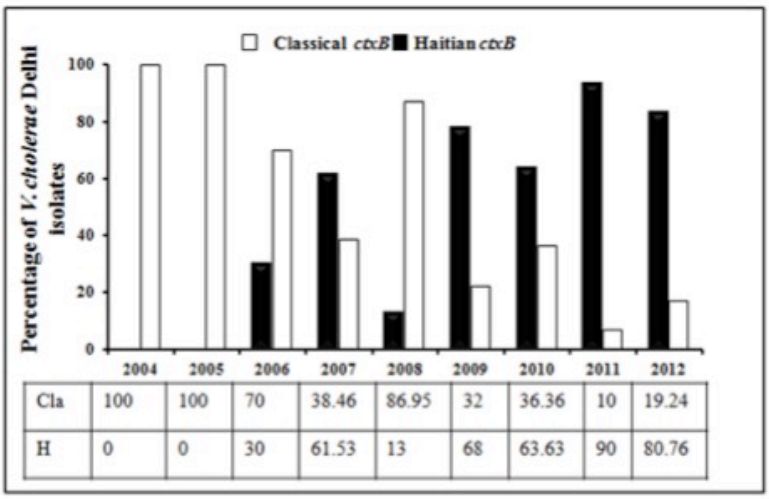
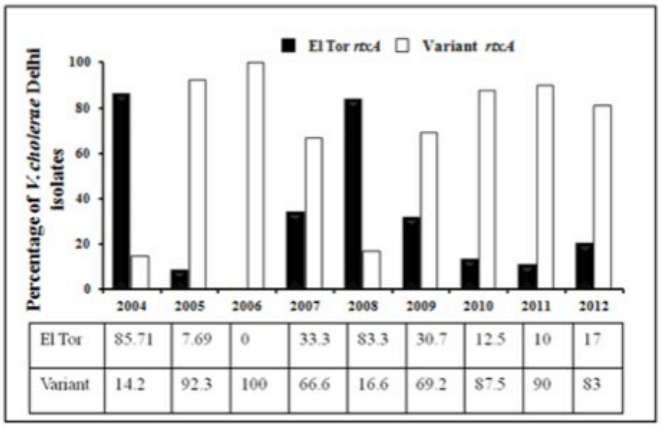


Figure 1



Cla	100	100	70	38.46	86.95	32	36.36	10	19.24
H	0	0	30	61.53	13	68	63.63	90	80.76

Figure 2



El Tor	85.71	7.69	0	33.3	83.3	30.7	12.5	10	17
Variant	14.2	92.3	100	66.6	16.6	69.2	87.5	90	83

Figure 3

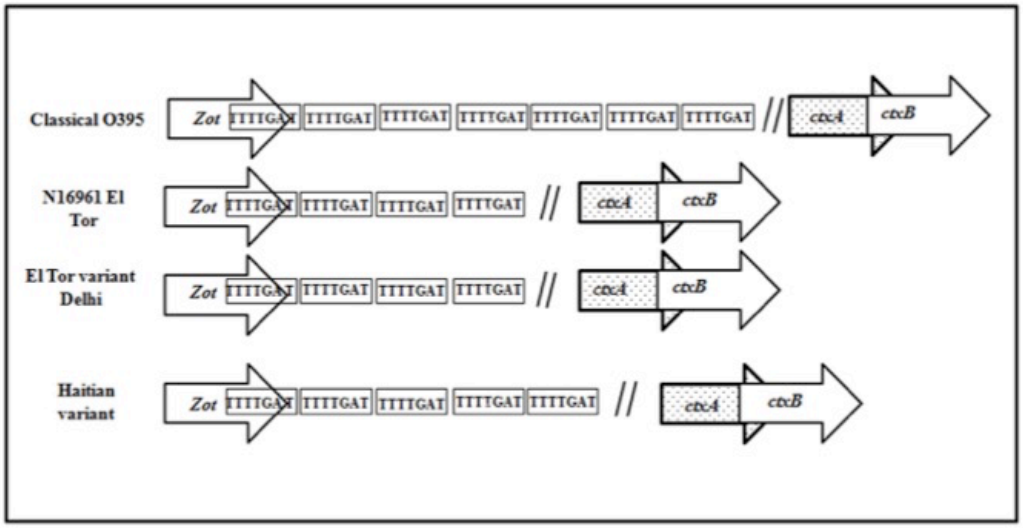


Figure 4

531

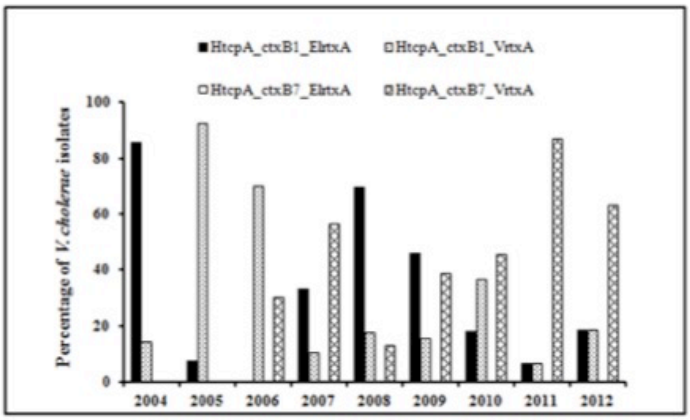


Figure 5

532
533
534