Comparative genomics of Vibrio cholerae O1 isolated from cholera patients in Bangladesh

Hossain, Zenat Zebin; Leekitcharoenphon, Pimlapas; Dalsgaard, Anders; Sultana, Rokaia; Begum, Anowara; Jensen, Peter Kjær Mackie; Hendriksen, Rene S.

Published in:
Letters in Applied Microbiology

Link to article, DOI:
10.1111/lam.13046

Publication date:
2018

Document Version
Peer reviewed version

MISS ZENAT ZEBIN HOSSAIN (Orcid ID : 0000-0002-6238-0323)

Article type : Original Article

Comparative genomics of *Vibrio cholerae* O1 isolated from cholera patients in Bangladesh

Zenat Zebin Hossain[1,2*], Pimlapas Leekitcharoenphon[3], Anders Dalsgaard[4], Rokaia Sultana[1], Anowara Begum[1], Peter Kjær Mackie Jensen[2,5], Rene S. Hendriksen[3]

[1]Department of Microbiology, University of Dhaka, Dhaka 1000, Bangladesh

[2]Institute of Public Health, University of Copenhagen, Copenhagen 1014, Denmark

[3] National Food Institute, Technical University of Denmark, WHO Collaborating Center for Antimicrobial Resistance in Food borne Pathogens & Genomics and European Union Reference Laboratory for Antimicrobial Resistance, Kgs. Lyngby, Denmark

[4]Department of Veterinary and Animal Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, 1870 Frederiksberg C, Copenhagen, Denmark

[5]Copenhagen Centre for Disaster Research, University of Copenhagen, Copenhagen 1014, Denmark

*: Corresponding author: Zenat Zebin Hossain

Department of Microbiology, University of Dhaka

Dhaka 1000, Bangladesh

&

Institute of Public Health, University of Copenhagen

Copenhagen 1014, Denmark

Phone: +8801785636044

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/lam.13046

Fax: +88-02-9667222

E-mail: zenat@sund.ku.dk

Running title: Cholera in Bangladesh

Key words: Whole genome sequencing, *Vibrio cholerae* O1, Bangladesh, SNP, MLST, antimicrobial resistance

Significance and impact of the study:

Vibrio cholerae has frequently experienced genetic changes with rapid evolution of pandemic clones in the Ganges Delta region. Whole Genome Sequencing can reveal genetic information of current pathogenic *V. cholerae* in Bangladesh which includes CTX genotypes, virulence factors, altered antimicrobial resistance pattern as well as mobile genetic element compared to global pandemic strains. This study data could be used in planning future surveillance strategies in Ganges Delta region by informing new epidemiology of current outbreak strains.

Abstract:

Whole Genome Sequencing (WGS) was utilized to investigate the genomic profile of *Vibrio cholerae* O1 strains, isolated from symptomatic patients in a low-income urban area of Dhaka, Bangladesh. Comparative genomics using bioinformatics tools were applied to identify major virulence factors, biotype and antimicrobial resistance genes in three *V. cholerae* O1 strains (VC-1, 2 and 3) isolated from two case patients. A phylogenetic SNP (Single Nucleotide Polymorphism) based analysis was conducted to infer the relatedness to *V. cholerae* O1 strains isolated elsewhere. The *V. cholerae* strains were the El Tor variant carrying *ctxB*1 (standard classical genotype). SNP based global phylogeny revealed that the

three isolates were strictly clonal and the closest neighboring genomes were epidemic clones of *V. cholerae* O1 isolated in 2010 from cholera patients in Pakistan. All strains harbored the integrase gene of the SXT element *(int_{SXT}),* antimicrobial resistance genes for aminoglycosides, phenicol, sulphonamide, and trimethoprim except VC-1 that lacked sulphonamide resistance genes. The Multilocus Sequence Typing (MLST) revealed that the strains belonged to sequence type, ST69. The study provides knowledge on current genetic traits of clinical *V. cholerae* O1 circulating in urban household clusters of Bangladesh which may help in predicting emergence of new pandemic strains in Bangladesh.

Introduction

Cholera has claimed millions of lives globally since the spread of the first pandemic from the Indian subcontinent. Despite ground breaking cholera interventions and diagnostic advancement, cholera cause both endemic and epidemic diseases in Bangladesh (Harris *et al.* 2008).

In urban slum areas of Dhaka, four per 1000 people become affected with cholera annually and the number is increasing due to rapid urbanization (Chowdhury *et al.* 2011). Clinical cholera cases in Bangladesh are entirely attributed to the O1 serogroup of *V. cholerae* with occasional sporadic outbreaks caused by the O139 serogroup (Chowdhury *et al.* 2015b). Although, the Classical biotype of *V. cholerae* O1 (*ctxB*1genotype) was extinct from this region since 1992 (Faruque *et al.* 1993; Chowdhury *et al.* 2015a); new variants of El Tor biotype strains producing classical toxin, *ctxB*1 (altered strains) emerged in Bangladesh in late 1990s (Nguyen *et al.* 2009). The 'altered' biotype strains have replaced the *ctxB*3 genotype of the 7th pandemic El Tor strains since 2001 and continued to be the cause of

severe dehydration and morbidity in hospitalized patients (Nair *et al.* 2006; Chowdhury *et al.* 2015a; Rashid *et al.* 2016).

In this study, whole genome sequencing (WGS) combined with the use of bioinformatics tools were used to determine the genomic and evolutionary characteristics of three *V. cholerae* O1 strains collected from two patients in Dhaka city, Bangladesh. Furthermore, comparative genomic analysis was conducted with spatial and temporal clinical *V. cholerae* O1 strains available from the public domain to determine the genetic relatedness.

Results and Discussion:

The distance between case 1, a girl of 4 years and case 2, a girl of 22 years was 588 m and the households used different water sources.

Two *V. cholerae* strains were isolated from case 1, i.e. VC-1 and VC-3. The patient was reported to have consumed untreated tap water from a closed roof tank supplied with groundwater. In addition, one *V. cholerae* strain was isolated from case 2, i.e. VC-2. The patient had consumed untreated communal tap water at home provided from the Dhaka city corporation (WASA). Subsequently, a number of relevant environmental samples such as drinking water, food and environmental swabs from the respective patients' households were investigated for the presence of *V. cholerae* but were found negative in culturing.

The three *V. cholerae* strains belonged to the 7th pandemic serogroup O1(*rfbV-O1*), serotype Inaba, biotype El Tor variant, ST69, and carried the standard classical *ctxB (ctxB*1) and El Tor *tcpA* genotype ($ctxB^{CC}rstR^{ET}tcpA^{ET}$) (Son *et al.* 2011; Siriphap *et al.* 2017). All three strains harbored 10 major virulence associated genes normally seen in clinical serogroup O1

strains (*ctxA*, *ctxB*, *zot*, *ace*, *tcpA*, *hlyA*, *mshA*, *rtxA*, *ompU*, and *toxR*) (Finkelstein and Mukerjee 1963; Hasan *et al.* 2013; Singh *et al.* 2002) except for the *ace* gene which was not detected in VC-3. The genes for type III secretion system, cholix toxin (*chxA*) and heat-stable enterotoxin (*stn*) were absent in the strains. The three *V. cholerae* strains harbored similar genetic backbone of 7th pandemic El Tor variant trait which has been the highly dominant type in Bangladesh in recent years (Chowdhury *et al.* 2015a). Since the emergence of the *ctxB*1 genotype in early 1990s, a similar temporal change occurred in Bangladesh after 2008 when classical genotype *ctxB*1 shifted to the Haitian genotype *ctxB*7 and *ctx*B1 returned again in 2013-14 (Rashid *et al.* 2016).

The strains showed a unique deletion in the Vibrio pathogenicity island-1 (VPI-1) gene cluster containing three consecutive gene deletions (VC-0819–0821). Variant of Vibrio 7th pandemic island-2 (VSP-2) was found in all the three genome sequences compared to the prototype N16961 with a deletion of 4 ORFs (VC-0495–0498). Similar variants were isolated in India before 2010 (El Ghany *et al.* 2014; Imamura *et al.* 2017). After 2010, another variant with a larger deletion spanning from ORF VC0495 to VC0512 (CIRS101 type VSP-2) became predominant in Kolkata, India and eventually replaced other types (Imamura *et al.* 2017).

The evolution through the larger deletions may occur due to significant genetic rearrangement within the two loci, at the 3′ end of the VC0498 and VC0511 of the genomic backbone of the island and also the event may be more human host specific (Taviani *et al.* 2010). In Bangladesh, the most recent report from 2004-2007 on VSP-2 variant of El Tor strains showed a higher prevalence of the CIRS101 variant (Taviani *et al.* 2010). Our results indicate recent occurrence of genetic shift in VSP-2 which is more similar to prototype 7th pandemic El Tor N16961. Previously, the frequent rearrangements of the variants of VSP-2

islands were also occurring in India and Bangladesh but the explanations for such rearrangements are still unclear (Imamura *et al.* 2017; Taviani *et al.* 2010).

All three strains were phenotypic susceptible to CHL, TET, AZM, NEO, GEN, CTX and resistant to CIP, STP, NAL, TMP, SMX and AMP (VC-3 intermediate). Multiple AMR genes were determined of which VC-2 and VC-3 showed identical profile harboring AMR genes for aminoglycosides (*str*A/*str*B); sulphonamides (*sul*2); chloramphenicol (*cat*B9); and trimethoprim (*dfr*A1). The strains did not show phenotypic resistance to chloramphenicol which is in agreement with a recent report showing that the *cat*B9 gene cassette was associated with very low-level of AMR to chloramphenicol (Kumar *et al.* 2017).VC-1 contained the same resistance genes as VC-2 and VC-3 except for the *sul*2 and *str*A genes. The occurrence of one genetic event may result the deletion as *sul*2 and *str*A genes locate right next to each other in the SXT element. Comparison of the specific genomic region in SXT element of VC-1, 2 and 3 with *V. cholerae* 2010EL-1786 further confirmed the absence of the *sul*2 gene in VC-1 (Figure 1). The locus deletion was found within 102.017 to 102.832 bp in VC-1 as compared to *V. cholerae* 2010 EL-1786 genome. Although, absence of *sul*2 gene in the VC-1 SXT element is unusual, such as *sul*2 gene deletion in SXT variant types has been reported (El Ghany *et al.* 2014; Siriphap *et al.* 2017). In 2009, only one in 38 *V. cholerae* strains had the deletions of the *sul*2 gene in a study of hospitalized patients in Chandigarh, India (El Ghany *et al.* 2014).

All the strains contained the SXT element (*int*SXT gene) and lacked the integrase genes of class 1 integron (*int*I gene). Our results corroborate previous findings in Asia and Africa (Iwanaga *et al.* 2004; Sambe-Ba *et al.* 2017; Siriphap *et al.* 2017). The NAL and CIP resistant strains conferred amino acid substitutions in *gyr*A (Ser83Ile) and *par*C (Ser85Leu) but did

not harbor the fluoroquinolone resistance *qnrVC1* gene. The genomic organization of the integrating conjugative element (ICE) of the *V. cholerae* strains showed an allelic profile similar to the reference ICE*Vch*Hai1 except the 21 common deletions including *flo*R in loci VC1786ICE6-9, VC1786ICE14, VC1786ICE21-29, VC1786ICE49-53, VC1786ICE83, VC1786ICE85 (Supplementary Table S5). VC-3 contains VC1786ICE10 loci, a putative transposase, but it was absent in the other two strains. The strains conferring resistance to trimethoprim harbored the *dfr*A1 gene and lacked *dfr*A18 which reveals the similarity of our strains with ICE*Vch*Ind5 (Spagnoletti *et al.* 2014). The Bangladesh SXT elements have highly structural similarity except for a few additional allelic deletions with the clade 1 isolates from Chandigarh, India isolated in 2009 which clustered with earlier outbreak strains of *V. cholerae* from India, Bangladesh and Nepal (El Ghany *et al.* 2014). Similarity with the Indian O1 strains indicates recent acquisition of the SXT element within strains from the same geographical areas.

The SNP analysis revealed the three *V. cholerae* strains being strictly clonal with zero SNP between them in comparison to the identified 390 SNPs identified among the 38 Bangladeshi strains. SNP accumulation depends on temporal signature and approximately 3.3 SNPs per year can be accumulated in the *V. cholerae* core genome (Mutreja *et al.* 2011). The three strains were isolated only 11 days apart and from local cholera outbreaks with two households located in the same community, which likely explains the lack of SNP differences. In Haiti, *V. cholerae* O1 strains from the outbreak in 2010 showed no polymorphic sites where as strains from Haitian patients in 2012 had 195 SNPs (Azarian *et al.* 2014). Moreover, the identical *V. cholerae* O1 clones found in the two case-households and the separate water sources used by the households indicate a contamination source other than water supported by earlier reports of cholera cases in Arichpur neighborhood (Bi *et al.*

2016). We have found differences in virulence and AMR gene profiles between VC-1 and VC-3 from the same patient; although there was no SNP difference. Moreover, super-integrons of the three strains did not show any diversity. Genetic diversity of *V. cholerae* strains from same patient may occur due to co-infection with multiple environmental strains (Rashed *et al.* 2014). As the three strains were clonal, there would be a possibility of circulation of similar outbreak strains in a short time span and between inter and intra person to person transmission.

The local phylogenetic tree showed that the isolates clustered into two major groups based on temporal patterns: one cluster included strains isolated during the period 1969-2000 and the 2nd cluster included strains isolated during the period 2001-2010 (Figure 2). The three *V. cholerae* O1 strains of this study branched from the group of 2001-2010 strains with their closest ancestral strain being PSC-022 isolated in 2010 which differed with 18 SNPs.

The SNP-based global phylogenetic tree revealed that the 469 *V. cholerae* O1 strains were discriminated by 4,141 high quality SNPs. Our Bangladeshi strains formed a monophyletic branch within a clade consisting of epidemic clones of *V. cholerae* O1 El Tor, isolated from Karachi, Pakistan during a major cholera epidemic in 2010 (Figure 3). The Pakistani clones included in the clade named PSC-1, were of coastal origin and formed a unique clade in the wave three of 7th pandemic lineages (Shah *et al.* 2014). We found 11 SNP differences between strain VC-2 and the closest neighboring genome (S9KCH9) (Supplementary Table S4) which suggests that the strains are closely related. The second nearest genomes of VC-2 with 12 SNPs difference were strains S7KCH20 and PCS-022 isolated from patients in Pakistan and Bangladesh, respectively. Interestingly, S9KCH9 and PCS-022 genomes differ

with only three SNPs suggesting a clonal origin of the *V. cholerae* O1 strains associated with cholera outbreaks in 2010 in Bangladesh and Pakistan.

In conclusion, the current Bangladeshi strains showed a clonal origin from Southeast Asian epidemic progenitor and possess the genetic backbone of current ongoing epidemic *V. cholerae* O1 in Bangladesh with considerably diversity in the pathogenicity islands and SXT elements. Our phylogeny results suggest that current *V. cholerae* O1 strains evolved from recent outbreak strains. Considering the global distribution, our study supports the hypothesis that Bengal Delta region remains the origin ground of cholera and persistent infection occur by either new epidemic clones or reemergence of previously existing clones in this region. This study provides new insights on cholera epidemiology in Bangladesh using WGS technology together with advanced and easily accessible bioinformatics tools. WGS is a more effective, robust and precise method than other molecular approaches for continuous monitoring and possible control of the spread of new epidemic *V. cholerae* O1 clones in Bangladesh, the land which still holds the title as the home of cholera.

Methods and materials

Ethics

The study was approved by the Ethical Review Committee (ERC) of International Centre for Diarrheal Disease Research, Bangladesh (ICDDR, B). An informed consent was obtained from the caretakers of the households, household respondents, and patients for collection of rectal swab, both survived.

Collection of epidemiological data

During 4 months period, 37 rectal swabs were taken by trained field staff in case of watery diarrhea reported by household caretaker and demographic, socioeconomic and clinical data were obtained in a brief interview with patients or respondents (in case of child).

Isolation, identification and antimicrobial susceptibility testing

The rectal swabs were cultured for *V. cholerae*, confirmed by PCR for the *ompW* gene (Nandi *et al.* 2000) and three out of 37 (8.1%) rectal swabs found positive. In October, 2015 three *V. cholerae* O1 strains were isolated from rectal swabs of two diarrhea cases identified during random visits by field staff and a report via choleraphone (mobile phones distributed to households for diarrhea reporting) (Sengupta *et al.* 2013). The distance between the two households was measured by a global positioning system. Suspected *V. cholerae* strains were isolated from rectal swabs by using standard biochemical, serological assays and a species-specific PCR (Huq *et al.* 2006; Nandi *et al.* 2000). According to the standard guideline described by the Clinical and Laboratory Standards Institute (CLSI), antimicrobial susceptibility testing to streptomycin (STP, 10 µg ml^{-1}), ciprofloxacin (CIP, 5 µg ml^{-1}), chloramphenicol (CHL, 30 µg ml^{-1}), trimethoprim (TMP, 5 µg ml^{-1}), sulfamethoxazole (SMX, 25 µg ml^{-1}), nalidixicacid (NAL, 30 µg ml^{-1}), tetracycline (TET, 30 µg ml^{-1}), azithromycin (AZM, 15 µg ml^{-1}), ampicillin (AMP, 10 µg ml^{-1}), neomycin (NEO, 30 µg ml^{-1}), gentamicin (GEN, 10 µg ml^{-1}), cefotaxime(CTX, 30 µg ml^{-1}) (OXOID, UK) was performed by agar disk diffusion method (CLSI 2016). *Escherichia coli* ATCC 25922 was used for quality control.

Whole Genome Sequencing (WGS)

The DNA was extracted from the three isolates for WGS as earlier described (Siriphap *et al.* 2017). Raw sequence data were deposited to Nucleotide Archive (http://www.ebi.ac.uk/ena) under study accession no.: PRJEB18714. The raw reads were de novo assembled using SPAdes 3.7.0 (Bankevich *et al.* 2012). The genomic data are included in the supplementary Table S1.

In silico bioinformatics analysis

The web-server tool MyDbFinder 1.1 (https://cge.cbs.dtu.dk/services/MyDbFinder/) and associated database (supplementary Table S1) was used to analyse the assembled *V. cholerae* genome sequences to identify the species-specific gene, serogroup-specific genes, the biotypes-specific genes, putative virulence genes, *ctxB* genotype specific genes, pathogenic islands (PAI), and the specific gene of the 7^{th} pandemic strain with a selected threshold of 98% identity (Siriphap *et al.* 2017).

The acquired antimicrobial resistance (AMR) genes were detected using ResFinder web server (version 2.1, 80% threshold for %ID/ 60% minimum length) (Zankari *et al.* 2012). Blast atlas tool (https://server.gview.ca/) was used to further compare the strains with reference *V. cholerae* 2010EL-1786, Accession No. NC_016445.1 (contain *sul2* gene). The SXT element, the class 1 integron, and the presence of mutations in the DNA gyrase (*gyr*A gene) and in the DNA topoisomerase IV (*par*C gene) were determined using MyDbFinder and associated database (Supplementary Table S2) (Kaas *et al.* 2016).

The genotypic profile of AMR determinates was compared with the genome of the 2010 Haiti *V. cholerae* O1 integrating conjugative element (ICE), ICE*Vch*Hai1(JN648379) and *dfr*A18 gene of SXTMO10 (AY034138) by using MyDBFinder (threshold, 95% identity). The MLST sequence types (ST) for *V. cholerae* strains were determined using the MLST tool (version 1.8) (https://cge.cbs.dtu.dk/services/MLST/).

Phylogenetic analysis

High quality Single Nucleotide Polymorphisms (SNPs) of the three O1 *V. cholerae* genomes from Bangladesh were identified using the pipeline CSI Phylogeny 1.4 available from the CGE website (https://cge.cbs.dtu.dk/services/CSIPhylogeny/). Raw read data and assembled genomes from 466 *V. cholerae* global pandemic and outbreak strains obtained from the European Nucleotide Archive (ENA) and GenBank were also analysed to determine genome wide SNPs and a global phylodynamic tree was constructed to map the Bangladeshi strains in a global context as previously published (Kaas *et al.* 2016). Moreover, a phylogenetic tree was inferred using SNPs differences in 38 genomes of the 7th pandemic *V. cholerae* O1 El Tor strains from Bangladesh. Raw reads were mapped to the published completed reference strain N16961 (Acc. No.: NC_002505.1) as earlier described (Delcher *et al.* 2002; Kaas *et al.* 2016). The sequence information of 466 global strains and the specific SNP differences are included in the Supplementary Table S3 and Table S4.The phylogenetic trees were created and visualized using online tool iTOL (Version 3) (https://itol.embl.de/).

Acknowledgement

We thank Jacob Dyring Jensen from DTU for his kind help in library preparation, handling and operating sequencer.

Funding

This study was carried out as a part of a Danish Government (DANIDA) funded research project 'Combating Cholera Caused by Climate Change' (C5) (Grant No. 12-040KU) in Bangladesh.

Conflict of interest

None declared.

References

Azarian, T., Ali, A., Johnson, J.A., Mohr, D., Prosperi, M., Veras, N.M., Jubair, M., Strickland, S.L., Rashid, M.H. and Alam, M.T. (2014) Phylodynamic analysis of clinical and environmental *Vibrio cholerae* isolates from Haiti reveals diversification driven by positive selection. *MBio* **5,** e01824-14.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A. and Pevzner, P.A. (2012) SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J Comput Biol* **19,** 455-477.

Bi, Q., Azman, A.S., Satter, S.M., Khan, A.I., Ahmed, D., Riaj, A.A., Gurley, E.S. and Lessler, J. (2016) Micro-scale spatial clustering of cholera risk factors in urban Bangladesh. *PLoS Negl Trop Dis* **10,** e0004400.

Chowdhury, F., Kuchta, A., Khan, A.I., Faruque, A., Calderwood, S.B., Ryan, E.T. and Qadri, F. (2015a) The increased severity in patients presenting to hospital with diarrhea in Dhaka, Bangladesh since the emergence of the hybrid strain of *Vibrio cholerae* O1 is not unique to cholera patients. *Int J Infect Dis* **40,** 9-14.

Chowdhury, F., Mather, A.E., Begum, Y.A., Asaduzzaman, M., Baby, N., Sharmin, S., Biswas, R., Uddin, M.I., LaRocque, R.C. and Harris, J.B. (2015b) *Vibrio cholerae* serogroup O139: isolation from cholera patients and asymptomatic household family members in Bangladesh between 2013 and 2014. *PLoS Negl Trop Dis* **9,** e0004183.

Chowdhury, F., Rahman, M.A., Begum, Y.A., Khan, A.I., Faruque, A.S., Saha, N.C., Baby, N.I., Malek, M., Kumar, A.R. and Svennerholm, A.-M. (2011) Impact of rapid urbanization on the rates of infection by *Vibrio cholerae* O1 and enterotoxigenic *Escherichia coli* in Dhaka, Bangladesh. *PLoS Negl Trop Dis* **5,** e999.

CLSI (2016) Performance Standards for Antimicrobial Susceptibility Testing; Twenty-Sixth Informational Supplement. *Wayne, PA. CLSI Document M100-S26, Clinical and Laboratory Standards Institute*.

Delcher, A.L., Phillippy, A., Carlton, J. and Salzberg, S.L. (2002) Fast algorithms for large-scale genome alignment and comparison. *Nucleic Acids Res* **30,** 2478-2483.

El Ghany, M.A., Chander, J., Mutreja, A., Rashid, M., Hill-Cawthorne, G.A., Ali, S., Naeem, R., Thomson, N.R., Dougan, G. and Pain, A. (2014) The population structure of Vibrio cholerae from the Chandigarh Region of Northern India. *PLoS Negl Trop Dis* **8,** e2981.

Faruque, S.M., Alim, A.A., Rahman, M.M., Siddique, A., Sack, R.B. and Albert, M.J. (1993) Clonal relationships among classical *Vibrio cholerae* O1 strains isolated between 1961 and 1992 in Bangladesh. *J Clin Microbiol* **31,** 2513-2516.

Finkelstein, R.A. and Mukerjee, S. (1963) Hemagglutination: A Rapid Method for Differentiating *Vibrio cholerae* and El Tor Vibrios. *Proc Soc Exp Biol Med* **112,** 355-359.

Harris, A.M., Chowdhury, F., Begum, Y.A., Khan, A.I., Faruque, A.S., Svennerholm, A.-M., Harris, J.B., Ryan, E.T., Cravioto, A. and Calderwood, S.B. (2008) Shifting

prevalence of major diarrheal pathogens in patients seeking hospital care during floods in 1998, 2004, and 2007 in Dhaka, Bangladesh. *Am J Trop Med Hyg* **79,** 708-714.

Hasan, N.A., Ceccarelli, D., Grim, C.J., Taviani, E., Choi, J., Sadique, A., Alam, M., Siddique, A.K., Sack, R.B., Huq, A. and Colwell, R.R. (2013) Distribution of Virulence Genes in Clinical and Environmental *Vibrio cholerae* Strains in Bangladesh. *Appl Environ Microbiol* **79,** 5782-5785.

Huq, A., Haley, B.J., Taviani, E., Chen, A., Hasan, N.A. and Colwell, R.R. (2006) Detection, isolation, and identification of *Vibrio cholerae* from the environment. *Curr Protoc Microbiol* 6A. 5.1-6A. 5.51.

Imamura, D., Morita, M., Sekizuka, T., Mizuno, T., Takemura, T., Yamashiro, T., Chowdhury, G., Pazhani, G.P., Mukhopadhyay, A.K. and Ramamurthy, T. (2017) Comparative genome analysis of VSP-II and SNPs reveals heterogenic variation in contemporary strains of *Vibrio cholerae* O1 isolated from cholera patients in Kolkata, India. *PLoS Negl Trop Dis* **11,** e0005386.

Iwanaga, M., Toma, C., Miyazato, T., Insisiengmay, S., Nakasone, N. and Ehara, M. (2004) Antibiotic Resistance Conferred by a Class I Integron and SXT Constin in *Vibrio cholerae* O1 Strains Isolated in Laos. *Antimicrob Agents Chemother* **48,** 2364-2369.

Kaas, R.S., Ngandjio, A., Nzouankeu, A., Siriphap, A., Fonkoua, M.-C., Aarestrup, F.M. and Hendriksen, R.S. (2016) The Lake Chad Basin, an isolated and persistent reservoir of *Vibrio cholerae* O1: A genomic insight into the outbreak in Cameroon, 2010. *PloS one* **11,** e0155691.

Kumar, P., Yadav, P., Nema, A., Goel, A.K. and Yadava, P.K. (2017) Re-emergence of chloramphenicol resistance and associated genetic background in *Vibrio cholerae* O1. *FASEB J* **31,** 907.3-907.3.

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., Niyogi, S.K., Kim, E.J., Ramamurthy, T., Chun, J., Wood, J.L., Clemens, J.D., Czerkinsky, C., Nair, G.B., Holmgren, J., Parkhill, J. and Dougan, G. (2011) Evidence for several waves of global transmission in the seventh cholera pandemic. *Nature* **477,** 462-5.

Nair, G.B., Qadri, F., Holmgren, J., Svennerholm, A.-M., Safa, A., Bhuiyan, N.A., Ahmad, Q.S., Faruque, S.M., Faruque, A. and Takeda, Y. (2006) Cholera due to altered El Tor strains of *Vibrio cholerae* O1 in Bangladesh. *J Clin Microbiol* **44,** 4211-4213.

Nandi, B., Nandy, R.K., Mukhopadhyay, S., Nair, G.B., Shimada, T. and Ghose, A.C. (2000) Rapid method for species-specific identification of *Vibrio cholerae* using primers targeted to the gene of outer membrane protein OmpW. *J Clin Microbiol* **38,** 4145-4151.

Nguyen, B.M., Lee, J.H., Cuong, N.T., Choi, S.Y., Hien, N.T., Anh, D.D., Lee, H.R., Ansaruzzaman, M., Endtz, H.P. and Chun, J. (2009) Cholera outbreaks caused by an altered *Vibrio cholerae* O1 El Tor biotype strain producing classical cholera toxin B in Vietnam in 2007 to 2008. *J Clin Microbiol* **47,** 1568-1571.

Rashed, S.M., Azman, A.S., Alam, M., Li, S., Sack, D.A., Morris Jr, J.G., Longini, I., Siddique, A.K., Iqbal, A. and Huq, A. (2014) Genetic variation of *Vibrio cholerae* during outbreaks, Bangladesh, 2010–2011. *Emerg Infect Dis* **20,** 54.

Rashid, M.-u., Rashed, S.M., Islam, T., Johura, F.-T., Watanabe, H., Ohnishi, M. and Alam, M. (2016) *ctxB*1 outcompetes *ctxB*7 in *Vibrio cholerae* O1, Bangladesh. *J Med Microbiol* **65,** 101-103.

Sambe-Ba, B., Diallo, M.H., Seck, A., Wane, A.A., Constantin de Magny, G., Boye, C.S.B., Sow, A.I. and Gassama-Sow, A. (2017) Identification of Atypical El Tor *V. cholerae* O1 Ogawa Hosting SXT Element in Senegal, Africa. *Front Microbiol* **8,** 748.

Sengupta, L.C., Tamason, C.C. and Jensen, P.K. (2013) The Choleraphone. *Trop Med Int Health* **18(Supplement s1),** 212-212.

Shah, M.A., Mutreja, A., Thomson, N., Baker, S., Parkhill, J., Dougan, G., Bokhari, H. and Wren, B.W. (2014) Genomic epidemiology of *Vibrio cholerae* O1 associated with floods, Pakistan, 2010. *Emerg Infect Dis* **20,** 13.

Singh, D., Isac, S.R. and Colwell, R. (2002) Development of a hexaplex PCR assay for rapid detection of virulence and regulatory genes in *Vibrio cholerae* and *Vibrio mimicus*. *J Clin Microbiol* **40,** 4321-4324.

Siriphap, A., Leekitcharoenphon, P., Kaas, R.S., Theethakaew, C., Aarestrup, F.M., Sutheinkul, O. and Hendriksen, R.S. (2017) Characterization and Genetic Variation of *Vibrio cholerae* Isolated from Clinical and Environmental Sources in Thailand. *PloS one* **12,** e0169324.

Son, M.S., Megli, C.J., Kovacikova, G., Qadri, F. and 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.

Spagnoletti, M., Ceccarelli, D., Rieux, A., Fondi, M., Taviani, E., Fani, R., Colombo, M.M., Colwell, R.R. and Balloux, F. (2014) Acquisition and Evolution of SXT-R391 Integrative Conjugative Elements in the Seventh-Pandemic *Vibrio cholerae* Lineage. *mBio* **5,** e01356-14.

Taviani, E., Grim, C.J., Choi, J., Chun, J., Haley, B., Hasan, N.A., Huq, A. and Colwell, R.R. (2010) Discovery of novel *Vibrio cholerae* VSP-II genomic islands using comparative genomic analysis. *FEMS Microbiol Lett* **308,** 130-137.

Zankari, E., Hasman, H., Kaas, R.S., Seyfarth, A.M., Agersø, Y., Lund, O., Larsen, M.V. and Aarestrup, F.M. (2012) Genotyping using whole-genome sequencing is a realistic

alternative to surveillance based on phenotypic antimicrobial susceptibility testing. *J Antimicrob Chemother* **68,** 771-777.

List of Supplementary files:

Supplementary Table 1 (S1). Genomic sequence data, virulence profile and occurrence of antimicrobial resistance genes in the three *V. cholerae* strains

Supplementary Table 2 (S2). The accession numbers of the genes in this study

Supplementary Table 3 (S3).Sequence information of 466 global *V. cholerae* strains

Supplementary Table 4 (S4). SNP differences in the global *V. cholerae* strains

Supplementary Table 5 (S5). Comparison of ICEVchHai1 gene loci with the three *V. cholerae* strains

Figure Legends:

Figure 1. Genomic variation of the three *V. cholerae* strains indicating *sul*2 gene using *V. cholerae* 2010EL-1786 chromosome one as reference (VC-3 VC-2 VC-1 *V. cholerae* 2010EL-1786)

Figure 2. SNP-based local phylogenetic tree constructed including 38 database strains (Branches are highlighted in colors according to years: Bangladesh 2015 Bangladesh 2001-2010

Bangladesh 1969-2000)

Figure 3. Global phylogenetic tree constructed with 466 database strains based on SNPs (Regions are highlighted in colors). The box in the lower left corner is focusing the position

of the 3 study strains in the global tree. 1. Country (Philippines China Thailand Pakistan Bangladesh India Nepal Vietnam Indonesia Afghanistan Bahrain Malaysia Haiti USA Mexico Dominican Republic Kenya Djibouti Mozambique South Africa Zambia Mauritania Tanzania Zimbabwe Peru Argentina Colombia Bolivia Chile Russia Cameroon) 2. Region (Southeast Asia/China/Bangladesh South America Africa Pakistan/ Afghanistan North America/Russia Bangladesh 2015 Middle East) 3. Time range (Pre-7th pandemic Predecessors 1957-1962 1963-1970 1971-1980 1981-1990 1991-2000 2001-2010 2011-2015)