Virulence-Related Genes Identified from the Genome Sequence of the Non-O1/Non-O139 *Vibrio cholerae* Strain VcN1, Isolated from Dhaka, Bangladesh

Maqsud Hossain,[a,b] Munirul Alam,[c] Abdul Khaleque,[a] Sohidul Islam,[a] Abdus Sadique,[b] Nayeim Khan,[a,b] Zahra Halim,[a,b] Mrinmoy Sarker,[a] Najib M. El-Sayed,[d,e] Anwar Huq,[d] Gias Uddin Ahsan,[b,f] Rita R. Colwell[d,g,h]

[a]Department of Biochemistry and Microbiology, North South University, Dhaka, Bangladesh

[b]NSU Genome Research Institute, North South University, Dhaka, Bangladesh

[c]International Centre for Diarrheal Disease Research, Bangladesh (ICDDR,B), Dhaka, Bangladesh

[d]Maryland Pathogen Research Institute, University of Maryland, College Park, Maryland, USA

[e]Center for Bioinformatics and Computational Biology, University of Maryland, College Park, Maryland, USA

[f]Department of Public Health, North South University, Dhaka, Bangladesh

[g]University of Maryland Institute of Advanced Computer Studies, University of Maryland, College Park, Maryland, USA

[h]Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

ABSTRACT   We report here the first draft genome sequence of the non-O1/non-O139 *Vibrio cholerae* strain VcN1, isolated from Dhaka, Bangladesh. The data submitted to GenBank for this strain will contribute to advancing our understanding of this environmentally disseminated bacterium, including its virulence and its evolution as an important pathogen.

*V*ibrio cholerae, the causative agent of cholera, is a Gram-negative bacterium that is adapted to the aquatic environment and is also a human pathogen. Of more than 200 known serogroups of *V. cholerae*, O1 and O139 are associated with a major virulence factor cholera toxin (CTX) and with epidemic cholera worldwide (1). The noncholera, or non-O1/non-O139, serogroups rarely carry the CTX gene cassette, but they can serve as a reservoir for virulence and related genes routinely found in cholera serogroup strains (2) and can cause diarrhea and extraintestinal infections (3).

Here, we report the draft genome sequence of the non-O1/non-O139 *V. cholerae* strain VcN1, isolated from a river in Dhaka, Bangladesh. A genomic library was constructed and employed for 300-bp paired-end whole-genome sequencing using an Illumina MiSeq platform (Illumina, San Diego, CA, USA) at the Genome Research Institute of North South University, Bangladesh. A total number of 293,954 raw reads were generated ($\sim$15$\times$ coverage) and assembled using SPAdes version 3.11 (4). The scaffold generated was mapped and ordered using ABACAS (5) and included the reference genome of *V. cholerae* O1 El Tor strain MS6 (GenBank accession no. AP014524 for chromosome 1 and AP014525 for chromosome 2). Structural gene prediction and functional annotation were performed using the Rapid Annotations using Subsystems Technology (RAST) server (6).

The total size of the draft assembly was 4,146,313 bp, arranged into 178 contigs with an N_{50} of 89 kb. The GC content was determined to be 47.57%. After scaffolding using the reference genome, 40 contigs totaling 2,997,625 bp in size mapped to chromosome 1 and 33 contigs totaling 1,113,511 bp in size mapped to chromosome 2. Another 105 contigs, with a cumulative size of 906,989 bp (N_{50} = 80 kb), did not map to the reference genome.

Several genes, gene clusters, and operons responsible for virulence, disease, and defense mechanisms were detected. Although the cholera enterotoxin *ctx* was absent, other known virulence factors, such as *zot* (zonula occludens toxin) and *toxr* and *toxs* (regulators of the expression of cholera pathogenicity) (7), were detected in the genome. Genes coding for resistance to antibiotics and toxic substances were also detected, including those for fluoroquinolone and tetracycline resistance, multidrug resistance efflux pumps, and the multidrug resistance tripartite system found in Gram-negative bacteria.

In future studies, the whole-genome sequence of *V. cholerae* VcN1 and additional strains of other cholera serogroups of *V. cholerae* will be analyzed using comparative genomics to understand differences in virulence and related factors, the incidence of antibiotic resistance genes, genome plasticity, and the evolutionary dynamics of this important human pathogen.

Accession number(s). This whole-genome shotgun project has been deposited in GenBank under the accession no. PDNJ00000000.

REFERENCES

1. Singh D, Matte M, Matte G, Jiang S, Sabeena F, Shukla B, Sanyal S, Huq A, Colwell R. 2001. Molecular analysis of *Vibrio cholerae* O1, O139, non-O1, and non-O139 strains: clonal relationships between clinical and environmental isolates. Appl Environ Microbiol 67:910–921. https://doi.org/10.1128/AEM.67.2.910-921.2001.

2. Ceccarelli D, Chen A, Hasan N, Rashed S, Huq A, Colwell R. 2015. Non-O1/non-O139 *Vibrio cholerae* carrying multiple virulence factors and *V. cholerae* O1 in the Chesapeake Bay, Maryland. Appl Environ Microbiol 81:1909–1918. https://doi.org/10.1128/AEM.03540-14.

3. Ramamurthy T, Bag P, Pal A, Bhattacharya S, Bhattarcharya M, Shimada T, Takeda T, Karasawa T, Kurazono H, Takeda Y, Nair G. 1993. Virulence patterns of *Vibrio cholerae* non-01 strains isolated from hospitalised patients with acute diarrhoea in Calcutta, India. J Med Microbiol 39:310–317. https://doi.org/10.1099/00222615-39-4-310.

4. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.

5. Assefa S, Keane T, Otto T, Newbold C, Berriman M. 2009. ABACAS: algorithm-based automatic contiguation of assembled sequences. Bioinformatics 25:1968–1969. https://doi.org/10.1093/bioinformatics/btp347.

6. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.

7. DiRita V. 1992. Co-ordinate expression of virulence genes by ToxR in *Vibrio cholerae*. Mol Microbiol 6:451–458. https://doi.org/10.1111/j.1365-2958.1992.tb01489.x.