

Draft Genome Sequencing of *Vibrio cholerae* O1 El Tor Isolates Collected in the Russian Federation from Imported Cholera Cases

Konstantin V. Kuleshov,^a Sergey O. Vodop'ianov,^b Vladimir G. Dedkov,^a Mikhail L. Markelov,^c Anton V. Kermanov,^b Vladimir D. Kruglikov,^b Alexey S. Vodop'ianov,^b Ruslan V. Pisanov,^b Olga S. Chemisova,^b Alexey B. Mazrukho,^b Svetlana V. Titova,^b German A. Shipulin^a

Federal Budget Institute of Science Central Research Institute for Epidemiology, Moscow, Russia^a; Federal Government Health Institution Rostov-on-Don Plague Control Research Institute, Rostov-on-Don, Russia^b; Institute of Occupational Health RAMS, Moscow, Russia^c

We report the draft genome sequencing of five *Vibrio cholerae* O1 El Tor clinical isolates collected in the Russian Federation from imported cholera cases in 2006, 2010, and 2012. In the initial phylogenetic analysis, one isolate clustered with the Haiti/Nepal-4 group.

Received 23 June 2014 Accepted 26 June 2014 Published 17 July 2014

Citation Kuleshov KV, Vodop'ianov SO, Dedkov VG, Markelov ML, Kermanov AV, Kruglikov VD, Vodop'ianov AS, Pisanov RV, Chemisova OS, Mazrukho AB, Titova SV, Shipulin GA. 2014. Draft genome sequencing of *Vibrio cholerae* O1 El Tor isolates collected in the Russian Federation from imported cholera cases. *Genome Announc.* 2(4):e00624-14. doi:10.1128/genomeA.00624-14.

Copyright © 2014 Kuleshov et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Konstantin V. Kuleshov, konstantinkul@gmail.com.

The geographical location of the Russia Federation and its integration with different countries are basic prerequisites for the importation of new strains of *Vibrio cholerae* from regions where the pathogen is endemic. The analysis of such strains might provide useful information about the circulation and a source of epidemiologically significant clones of *V. cholerae* on a global scale. In our study, we present the initial steps performed to obtain and process the whole-genome sequencing data of five *V. cholerae* O1 El Tor clinical isolates that were collected in the Russian Federation from imported cholera cases.

The clinical isolates were identified as *V. cholerae* O1 El Tor on the basis of standard biochemical and serologic testing. All studied isolates, RND6878, RND19191, RND1889, RND19188, and RND19187, represent cases of cholera imported from India. Two of them, RND19188 and RND19187, were isolated from a woman and her 10-month-old daughter.

Genomic DNA was extracted using the standard phenol-chloroform extraction protocol. Libraries for paired-end sequencing were prepared from 1 to 5 µg of genomic DNA by nebulization. Two steps of size selection were applied to get the libraries, with an approximately 500- to 700-bp insert size. The libraries were sequenced on the MiSeq platform using reagent kit version 2 (Illumina; San Diego, CA, USA). Raw sequencing data were processed by CASAVA1.8 software (Illumina; San Diego, CA, USA). The resulting raw reads were submitted to the NCBI Short Read Archive. To exclude low-quality data and adapter sequences for further analysis, reads were filtered out by Trimmomatic (1). Reads containing at least one base with a quality score of <20 were removed. The average insert size distribution estimated by PICARD tools (<http://picard.sourceforge.net>) varied in a narrow range, with a median of 500 bp. *De novo* assembly was performed using default options with CLC Genomics Workbench version 5.1 (CLC bio; Aarhus, Denmark) and SPAdes 3.0.0 (2). As a result, 164 and 234 contigs were generated, with estimated sizes of the genomes ranging between 3,991,231 and 4,005,480 bp.

BLASTn (3) searches against the nonredundant nucleotide database were performed, and taxonomy assignments were manually validated to exclude potential contaminating sequences from the resulting set of contigs. The average coverage for each resulting draft genome varied between 88× and 136×.

It was found that isolate RND19191 is closely related to a group of Haitian isolates on the basis of identified core single nucleotide polymorphisms (SNPs) using a read mapping approach (4) of globally representative *V. cholerae* O1 El Tor strains (4–8).

A detailed report of the travel histories from cholera case patients as well as the results of comparative and phylogenetic analyses of these genomes and other available *V. cholerae* genomes will be published elsewhere.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [AYNL00000000](https://www.ncbi.nlm.nih.gov/nuclink/AYNL00000000) (RND6878), [JNGT00000000](https://www.ncbi.nlm.nih.gov/nuclink/JNGT00000000) (RND19191), [JNGU00000000](https://www.ncbi.nlm.nih.gov/nuclink/JNGU00000000) (RND19188), [AYNM00000000](https://www.ncbi.nlm.nih.gov/nuclink/AYNM00000000) (RND19187), and [AYNN00000000](https://www.ncbi.nlm.nih.gov/nuclink/AYNN00000000) (RND18899).

ACKNOWLEDGMENT

This work was supported by the Russian Federal Target Program “National system of chemical and biological safety of the Russian Federation (2009–2014)” (ID 11-D 08042014).

REFERENCES

1. Bolger AM, Lohse M, Usadel B. 1 April 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. <http://dx.doi.org/10.1093/bioinformatics/btu170>.
2. 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. <http://dx.doi.org/10.1089/cmb.2012.0021>.
3. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403–410. [http://dx.doi.org/10.1016/S0022-2836\(05\)80360-2](http://dx.doi.org/10.1016/S0022-2836(05)80360-2).
4. Katz LS, Petkau A, Beaulaurier J, Tyler S, Antonova ES, Turnsek MA,

- Guo Y, Wang S, Paxinos EE, Orata F, Gladney LM, Stroika S, Folster JP, Rowe L, Freeman MM, Knox N, Frace M, Boncy J, Graham M, Hammer BK, Boucher Y, Bashir A, Hanage WP, Van Domselaar G, Tarr CL. 2013. Evolutionary dynamics of *Vibrio cholerae* O1 following a single-source introduction to Haiti. *mBio* 4(4):e00398-13. <http://dx.doi.org/10.1128/mBio.00398-13>.
5. Hendriksen RS, Price LB, Schupp JM, Gillette JD, Kaas RS, Engelthaler DM, Bortolaia V, Pearson T, Waters AE, Upadhyay BP, Shrestha SD, Adhikari S, Shakya G, Keim PS, Aarestrup FM. 2011. Population genetics of *Vibrio cholerae* from Nepal in 2010: evidence on the origin of the Haitian outbreak. *mBio* 2(4):e00157-11. <http://dx.doi.org/10.1128/mBio.00157-11>.
6. Reimer AR, Van Domselaar G, Stroika S, Walker M, Kent H, Tarr C, Talkington D, Rowe L, Olsen-Rasmussen M, Frace M, Sammons S, Dahourou GA, Boncy J, Smith AM, Mabon P, Petkau A, Graham M, Gilmour MW, Gerner-Smidt P, V. cholerae Outbreak Genomics Task Force. 2011. Comparative genomics of *Vibrio cholerae* from Haiti, Asia, and Africa. *Emerg. Infect. Dis.* 17:2113–2121.
7. 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. 2012. Genomic diversity of 2010 Haitian cholera outbreak strains. *Proc. Natl. Acad. Sci. U. S. A.* 109:E2010–E2017. <http://dx.doi.org/10.1073/pnas.1207359109>.
8. Shah MA, Mutreja A, Thomson N, Baker S, Parkhill J, Dougan G, Bokhari H, Wren BW. 2014. Genomic epidemiology of *Vibrio cholerae* O1 associated with floods, Pakistan, 2010. *Emerg. Infect. Dis.* 20:13–20. <http://dx.doi.org/10.3201/eid2001.130428>.