

# Genome Sequence of Non-O1 *Vibrio cholerae* PS15

Sanath Kumar,<sup>a</sup> Ingrid E. Lindquist,<sup>b</sup> Anitha Sundararajan,<sup>b</sup> Chythanya Rajanna,<sup>c</sup> Jared T. Floyd,<sup>a</sup> Kenneth P. Smith,<sup>a</sup> Jody L. Andersen,<sup>a</sup> Guixin He,<sup>d</sup> Ryan M. Ayers,<sup>a</sup> Judith A. Johnson,<sup>c</sup> James J. Werdann,<sup>a</sup> Ava A. Sandoval,<sup>a</sup> Nadia M. Mojica,<sup>a</sup> Faye D. Schilkey,<sup>b</sup> Joann Mudge,<sup>b</sup> Manuel F. Varela<sup>a</sup>

Eastern New Mexico University, Department of Biology, Portales, New Mexico, USA<sup>a</sup>; National Center for Genomic Resources, Santa Fe, New Mexico, USA<sup>b</sup>; Emerging Pathogens Institute, University of Florida, Gainesville, Florida, USA<sup>c</sup>; University of Massachusetts—Lowell, Department of Clinical Laboratory and Nutritional Sciences, Lowell, Massachusetts, USA<sup>d</sup>

**The draft genome sequence of a non-O1 *Vibrio cholerae* strain, PS15, organized into 3,512 open reading frames within a 3.9-Mb genome, was determined. The PS15 genome sequence will allow for the study of the evolution of virulence and environmental adaptation in *V. cholerae*.**

Received 18 December 2012 Accepted 21 December 2012 Published 14 February 2013

**Citation** Kumar S, Lindquist IE, Sundararajan A, Rajanna C, Floyd JT, Smith KP, Andersen JL, He G, Ayers RM, Johnson JA, Werdann JJ, Sandoval AA, Mojica NM, Schilkey FD, Mudge J, Varela MF. 2013. Genome sequence of non-O1 *Vibrio cholerae* PS15. *Genome Announc*. 1(1):e00227-12. doi:10.1128/genomeA.00227-12.

**Copyright** © 2013 Kumar et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Manuel F. Varela, [Manuel.Varela@enmu.edu](mailto:Manuel.Varela@enmu.edu).

*Vibrio cholerae* is an estuarine bacterium that is free living or lives in association with crustaceans, algae, and fish (1). *V. cholerae* serogroups O1 and O139, which produce cholera toxin (CTX), cause the serious gastrointestinal disorder cholera (2). Non-O1 and non-O139 serogroups, of which there are roughly 200, generally do not produce CTX and are apparently not involved in epidemic-scale infections, though their involvement in diarrhea and extraintestinal infections is documented (3).

Here, we report the whole-genome sequence of a non-O1 *V. cholerae* strain, PS15, isolated from sediment in Puget Sound, WA (kind gift of Charles Kaysner). Purified genomic DNA was constructed into a sequencing library and was sequenced on the Illumina Genome Analyzer II platform. A total of 25,451,494 singleton 36-bp passing reads, approximately 234× coverage of the expected genome size, were assembled with a Kmer-sweep in Assembly By Short Sequences (ABYSS) (4), followed by dederundification with CD-HIT (5) and merging of the assembly with Phrap (6). The final assembly includes 131 contigs that are >1 kb and captures most of the *V. cholerae* genome without redundancy (95% of reads realigned uniquely to the assembly). Structural gene prediction, functional annotation, and a comparative gene-based analysis with O1 *V. cholerae* N16961 were performed using the Rapid Annotations using Subsystems Technology (RAST) server (7).

*V. cholerae* PS15 has a 3,910,387-bp genome with a G+C content of 47.55%. We aligned our 131 contigs to the well-established *V. cholerae* O395 reference genome and uniquely placed 33 of our contigs to chromosome I and 65 of our contigs to chromosome II. The remaining 27 contigs that aligned to O395 had portions mapping to chromosomes, suggesting possible chromosomal rearrangements between the PS15 and O395 genomes. Six contigs did not align to *V. cholerae* O395 and may represent novel genomic regions or plasmid sequences. No complete plasmids were assembled, but similarities with plasmid sequences were found sporadically throughout the genome, suggesting prior plasmid integration.

A comparison of the PS15 genome with that of O1 *V. cholerae* N16961 (8) revealed the absence of about 619 genes. The majority

(432 open reading frames [ORFs]) of these encode hypothetical proteins. Others include the *V. cholerae* virulence cassette containing *ctxAB* (cholera toxin), *zot* (zonula occludens toxin), *ace* (accessory cholera enterotoxin), *cep* (core-encoded pilin), and *orfU* genes, and virulence-associated genes, such as the *tcp* (toxin-coregulated protein) gene cluster, the *acfA* (accessory colonization factor) gene, and the *hig* toxin-antitoxin system. However, the *par* toxin-antitoxin system (*parA*, *parB*, and *parC*) was identified. About 245 ORFs were unique to the *V. cholerae* PS15 genome. A 57-kb *Vibrio* pathogenic island-2 (VPI-2), found in toxigenic *V. cholerae* strains (9), was not found in the PS15 genome, though two genes, sialic acid-related *nanM* and neuraminidase-encoding *nanH*, were identified. Interestingly, the non-O1 *V. cholerae* NRT36S genome contains an altered VPI-2 with an intact core *nan-nag* region (10). Further, PS15 harbors genes of *Vibrio* seventh pandemic island-II (VC0494 to VC0497).

The whole-genome sequence of *V. cholerae* PS15 will help our future studies to determine the differences in virulence gene composition and the molecular bases for differences in the physiology of substrate transport and utilization, environmental fitness, biofilm formation, and genome plasticity between O1 and non-O1 serogroups of *V. cholerae*.

**Nucleotide sequence accession numbers.** This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [AIJR000000000](https://www.ncbi.nlm.nih.gov/nuccore/AIJR000000000). The version described in this article is the first version, [AIJR010000000](https://www.ncbi.nlm.nih.gov/nuccore/AIJR010000000).

## ACKNOWLEDGMENTS

This work was supported by grants from the National Institute of General Medical Sciences (8P20 GM103451-12) and the National Institutes of Health (5P20 RR016480-12 and 8P20 GM103451-12) and an Internal Research Grant from Eastern New Mexico University.

## REFERENCES

1. Colwell RR. 1996. Global climate and infectious disease: the cholera paradigm. *Science* 274:2025–2031.
2. Faruque SM, Nair GB. 2002. Molecular ecology of toxigenic *Vibrio cholerae*. *Microbiol. Immunol.* 46:59–66.

3. Ramamurthy T, Bag PK, Pal A, Bhattacharya SK, Bhattacharya MK, Shimada T, Takeda T, Karasawa T, Kurazono H, Takeda Y, Nair GB. 1993. Virulence patterns of *Vibrio cholerae* non-O1 strains isolated from hospitalised patients with acute diarrhoea in Calcutta, India. *J. Med. Microbiol.* 39:310–317.
4. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I. 2009. ABySS: A parallel assembler for short read sequence data. *Genome Res.* 19:1117–1123.
5. Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22: 1658–1659.
6. Ewing B, Green P. 1998. Base-calling of automated sequencer traces using Phred. II. Error probabilities. *Genome Res.* 8:186–194.
7. 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.
8. Heidelberg JF, Eisen JA, Nelson WC, Clayton RA, Gwinn ML, Dodson RJ, Haft DH, Hickey EK, Peterson JD, Umayam L, Gill SR, Nelson KE, Read TD, Tettelin H, Richardson D, Ermolaeva MD, Vamathevan J, Bass S, Qin H, Dragoi I, Sellers P, McDonald L, Utterback T, Fleischmann RD, Nierman WC, White O, Salzberg SL, Smith HO, Colwell RR, Mekalanos JJ, Venter JC, Fraser CM. 2000. DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* 406: 477–483.
9. Jermyn WS, Boyd EF. 2002. Characterization of a novel *Vibrio* pathogenicity island (VPI-2) encoding neuraminidase (*nanH*) among toxigenic *Vibrio cholerae* isolates. *Microbiology* 148:3681–3693.
10. Chen Y, Johnson JA, Pusch GD, Morris JG Jr, Stine OC. 2007. The genome of non-O1 *Vibrio cholerae* NRT36S demonstrates the presence of pathogenic mechanisms that are distinct from those of O1 *Vibrio cholerae*. *Infect. Immun.* 75:2645–2647.