Genome Sequence of the Pathogenic Bacterium *Vibrio vulnificus* Biotype 3

Yael Danin-Poleg,[a] Sharona Elgavish,[b] Nili Raz,[a] Vera Efimov,[a] Yechezkel Kashi[a]

Faculty of Biotechnology and Food Engineering, Technion-Israel Institute of Technology, Haifa, Israel[a]; Bioinformatics Knowledge Unit, Lorry I. Lokey Interdisciplinary Center for Life Sciences and Engineering, Technion-Israel Institute of Technology, Haifa, Israel[b]

Y.D.-P. and S.E. contributed equally to this work.

We report the first genome sequence of the pathogenic *Vibrio vulnificus* biotype 3. This draft genome sequence of the environmental strain VVyb1(BT3), isolated in Israel, provides a representation of this newly emerged clonal group, which reveals higher similarity to the clinical strains of biotype 1 than to the environmental ones.

Address correspondence to Yechezkel Kashi, kashi@tx.technion.ac.il.

*V*ibrio vulnificus is an aquatic bacterium and an important human pathogen (1–4). Strains of *V. vulnificus* are biochemically divided into three biotypes. The highly virulent biotype 3, initially isolated in 1996 (5), appears to be geographically restricted to Israel and found to be rather clonal (6–8), although discrimination was achieved among strains, enabling epidemiology studies (9).

Here we describe the draft genome sequence of the environmental *V. vulnificus* biotype 3 strain VVyb1(BT3), isolated from *Tilapia* fish in 2004 (7, 9) and subjected to whole-genome shotgun sequencing.

The genome complexity of VVyb1(BT3), derived from the high rates of horizontal gene transfer in the *Vibrio* species (10–12) and the presence of multiple repetitive regions, required implementation of a multistep strategy in sequencing and in the bioinformatics analysis. Three different libraries, with average insert sizes of 200 bp, 4,500 bp, and 455 bp, were prepared. The first two libraries were sequenced using Illumina Genome Analyzer IIx, and the third was sequenced using Illumina HiSeq 2000, generating 37,426,728 33-bp paired-end reads, 18,820,750 36-bp mate-pair reads, and 164,399,024 100-bp paired-end reads, with coverages of 215×, 118×, and 2,864×, respectively. All reads of the two first libraries were *de novo* assembled simultaneously with different insert lengths, providing 306 contigs of >500 bp using Velvet 0.7.54 (13). The third library was *de novo* assembled with CLCbio Genomics workbench 5.0 (CLC Bio, Denmark) to generate 470 contigs of >500 bp, using reads with a minimum quality of 30 for each base. The contigs of the third library's assembly were contiguated (aligned, ordered, and oriented) using ABACAS (14), with *V. vulnificus* CMCP6 (10) serving as a reference genome for chromosomes 1 and 2 and YJ016 for the plasmid (15), to produce 3 scaffolds that split into 95 contigs. This set of contigs was then merged with the assembly of the first two libraries using Minimus2 (AMOS project), providing 69 contigs and numerous singletons that have undergone ABACAS analysis, generating 3 scaffolds plus 93 contigs. Cleaning out sequence redundancy left 84 contigs and 3 scaffolds, resulting in 140 contigs after splitting, with an N_{50} of 230,903 bp and the longest contig of 644,564 bp.

The genome of VVyb1(BT3) consists of 2 chromosomes and a plasmid (5.74 Mbp; 46.7% G+C content). A total of 5,361 coding sequences (CDS), 6 rRNAs, and 76 tRNAs were predicted and annotated by the RAST annotation server (16). Of the 5,361 CDS, 217 are unique to VVyb1(BT3) and have low/no similarity (BLASTn, filtering below 85% identity and 80% query coverage) to sequences in the known *V. vulnificus* genomes, and the majority (65%) were annotated as encoding hypothetical proteins. These genes were probably acquired horizontally from other bacterial species sharing the same ecological niche, such as other *Vibrio* species and *Pseudomonas* species. Of the unique genes, 116 show no sequence similarity to any of the known sequences in the NCBI nr database (BLASTn: query coverage, <30%, and E value, >10^{-10}). Since biotype 3 is highly clonal, the VVyb1(BT3) genome provides representation of this group, contributing to the understanding of the evolution of this human pathogen.

Nucleotide sequence accession numbers. The draft genome sequence for this project has been deposited at DDBJ/EMBL/GenBank under the accession number AOCM00000000. The version described here is the first version, AOCM01000000.

REFERENCES

1. **Jones MK, Oliver JD.** 2009. Vibrio vulnificus: disease and pathogenesis. Infect. Immun. 77:1723–1733.

2. **Oliver JD.** 2006. Vibrio vulnificus, p 349–366. *In* Thompson FL, Austin B, Swings J (ed), The biology of vibrios. ASM Press, Washington, DC.

3. **Oliver JD, Warner RA, Cleland DR.** 1982. Distribution and ecology of *Vibrio vulnificus* and other lactose-fermenting marine vibrios in coastal waters of the southeastern United States. Appl. Environ. Microbiol. **44**: 1404–1414.

4. **Strom MS, Paranjpye RN.** 2000. Epidemiology and pathogenesis of *Vibrio vulnificus*. Microbes Infect. **2**:177–188.

5. **Bisharat N, Agmon V, Finkelstein R, Raz R, Ben-Dror G, Lerner L, Soboh S, Colodner R, Cameron DN, Wykstra DL, Swerdlow DL, Farmer JJ, and the Israel Vibrio Study Group.** 1999. Clinical, epidemiological, and microbiological features of *Vibrio vulnificus* biogroup 3 causing outbreaks of wound infection and bacteraemia in Israel. Lancet **354**:1421–1424.

6. **Bisharat N, Amaro C, Fouz B, Llorens A, Cohen DI.** 2007. Serological and molecular characteristics of *Vibrio vulnificus* biotype 3: evidence for high clonality. Microbiology **153**:847–856.

7. **Broza YY, Danin-Poleg Y, Lerner L, Broza M, Kashi Y.** 2007. Vibrio vulnificus typing based on simple sequence repeats: insights into the biotype 3 group. J. Clin. Microbiol. **45**:2951–2959.

8. **Broza YY, Raz N, Lerner L, Danin-Poleg Y, Kashi Y.** 2012. Genetic diversity of the human pathogen *Vibrio vulnificus*: a new phylogroup. Int. J. Food Microbiol. **153**:436–443.

9. **Broza YY, Danin-Poleg Y, Lerner L, Valinsky L, Broza M, Kashi Y.** 2009. Epidemiologic study of *Vibrio vulnificus* infections by using variable number tandem repeats. Emerg. Infect. Dis. **15**:1282–1285.

10. **Kim HU, Kim SY, Jeong H, Kim TY, Kim JJ, Choy HE, Yi KY, Rhee JH, Lee SY.** 2011. Integrative genome-scale metabolic analysis of *Vibrio vulnificus* for drug targeting and discovery. Mol. Syst. Biol. **7**:460.

11. **Quirke AM, Reen FJ, Claesson MJ, Boyd EF.** 2006. Genomic island identification in *Vibrio vulnificus* reveals significant genome plasticity in this human pathogen. Bioinformatics **22**:905–910.

12. **Urbanczyk H, Ast JC, Kaeding AJ, Oliver JD, Dunlap PV.** 2008. Phylogenetic analysis of the incidence of lux gene horizontal transfer in Vibrionaceae. J. Bacteriol. **190**:3494–3504.

13. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18**:821–829.

14. **Assefa S, Keane TM, Otto TD, Newbold C, Berriman M.** 2009. ABACAS: algorithm-based automatic contiguation of assembled sequences. Bioinformatics **25**:1968–1969.

15. **Chen CY, Wu KM, Chang YC, Chang CH, Tsai HC, Liao TL, Liu YM, Chen HJ, Shen AB, Li JC, Su TL, Shao CP, Lee CT, Hor LI, Tsai SF.** 2003. Comparative genome analysis of *Vibrio vulnificus*, a marine pathogen. Genome Res. **13**:2577–2587.

16. **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.