

GENOME REPORT

Open Access



Complete genome sequence of the *Vibrio vulnificus* strain VV2014DJH, a human-pathogenic bacterium isolated from a death case in China

Junhang Pan^{1,2}, Yi Sun², Wenwu Yao², Haiyan Mao², Yanjun Zhang^{2*} and Muyuan Zhu^{1*}

Abstract

Background: *Vibrio vulnificus*, an opportunistic pathogen, is the causative agent of life-threatening septicemia and severe wound infections. However, the pathogenicity and virulence factors of *V. vulnificus* are not fully understood. Here we report the complete genome sequence of *V. vulnificus* VV2014DJH, which was isolated from a death case.

Results: The genome of the *V. vulnificus* VV2014DJH contains two circular chromosomes with a mean G+C content of 46.8%, but does not consist of any plasmids. The chromosome I and chromosome II consist of 3,303,590 and 1,770,972 bp, respectively. In addition, the genome consists of 4617 protein coding genes, 172 RNA genes and type I, II and III secretion systems were predicted.

Conclusions: In this study, the genomic information of the *V. vulnificus* VV2014DJH has been described. The information would contribute to the increasing scope and depth of *Vibrio* genome database, and provide insights into the pathogenicity and virulence factors of *V. vulnificus*.

Introduction

Vibrio vulnificus is an opportunistic Gram-negative pathogen, which is widely distributed in marine environments around the world and has been isolated from water, sediments and seafood [1]. *V. vulnificus* is the leading cause of deaths reported from seafood in the United States with approximately 40 cases reported per year, and the fatality rate exceeds 50% [2]. However, the mortality rate ranges from 18 to 56% in China [3, 4]. In the recent years, more cases of amputation and death caused by *V. vulnificus* infections were reported in China, and most of these cases occurred in the coastal area of Zhejiang province [5]. In this paper, we report the complete genome sequence of *V. vulnificus* strain VV2014DJH, which was isolated from the blood culture of the death case of a shellfish aquaculture worker in Taizhou, Zhejiang

province [6]. Determination of the genomic information of *V. vulnificus* strain VV2014DJH is necessary to understand the pathogenesis of the strain. Our study was approved by the ethics committee of the Zhejiang Provincial Center for Disease Control and Prevention (ZJCDC), China.

Methods

The genomic DNA of VV2014DJH was extracted using the DNeasy Blood and Tissue Kit (QIAGEN, Germany) for Gram-negative bacteria, according to the manufacturer's instructions. Total DNA was subjected to quality control by 2% agarose gel electrophoresis and quantified by a NanoDropTM spectrophotometer. The sequencing of VV2014DJH was performed on Pacbio RS with Single Molecule, Real-Time (SMRT) technology. SMRT Analysis 2.3.0 was used to filter low quality reads, whose quality scores were under 40, and the filtered reads were assembled by SOAPdenovo2 to generate one contig without gaps [7]. According to Clusters of Orthologous

*Correspondence: yjzhang@cdc.zj.cn; myzhu@zju.edu.cn

¹ College of Life Sciences, Zhejiang University, Hangzhou 310058, China

² Zhejiang Provincial Center for Disease Control and Prevention, Hangzhou 310051, China

Group (COG) category, circular representations with the annotated genes were obtained using Circos software and displayed the diagram of the VV2014DJH genome.

Transfer RNA (tRNA) genes were predicted by tRNAscan-SE [8]. Ribosome RNA (rRNA) genes were predicted with rRNAmmer [9] and sRNAs were obtained by BLAST against Rfam database [10]. RepeatMasker [11] (<http://www.repeatmasker.org/>) was used to detect repetitive sequences, and Tandem Repeat Finder [12] (<http://tandem.bu.edu/trf/trf.html>) was used to find Tandem Repeats. Gene prediction was performed on the VV2014DJH genome assembly by GeneMarkS [13] (<http://topaz.gatech.edu/>) with integrated model which combine the GeneMarkS generated (native) and Heuristic model parameters.

A whole genome BLAST [14] search (E-value less than $1E-5$, minimal alignment length percentage larger than 40%) was performed against five databases. They are KEGG [15–17] (Kyoto Encyclopedia of Genes and Genomes), COG [18, 19], NR (Non-Redundant Protein Database), Swiss-Prot [20], and GO [21] (Gene Ontology) [22]. Secretory proteins were detected in the genome assembly by SignalP [23]. Type I–VII secretion system related proteins were extracted from all the annotation results. Type III secretion system effector proteins were detected by EffectiveT3 also [24]. The MUMmer system (version 3.0) was used applied for colinearity analysis of to compare the genome sequences of VV2014DJH, with CMCP6 and YJ016 [25]. The virulence factors annotation in VV2014DJH was performed by blasting the genome sequence to the Virulence factors Pathogenic Bacteria database (VFDB, <http://www.mgc.ac.cn/cgi-bin/VFs/jsif/main.cgi>) [26].

Quality assurance

The genomic DNA used for sequencing was isolated from a single colony of the VV2014DJH. The 16S rRNA gene was sequenced and BLAST was performed against the NCBI database. In addition, the raw read sequences were selected and assembled only when their quality scores were more than 40 as cutoffs.

Results

A PacBio RS sequencing run resulted in 56,579 reads with a mean read length of 8107 bp and a N50 read length of 11,532 bp. PacBio RS reads were assembled into one polished contigs with a N50 contig length of 3,307,273. The genome of VV2014DJH has a size of 5,074,562 bp with a mean G+C content of 46.8% and is composed of two circular chromosomes (chromosome I: 3,303,590 bp, 46.6% G+C content; chromosome II: 1,770,972 bp, 47.3% G+C content). It does not contain any plasmids.

A total of 4617 protein coding genes and 172 RNA genes were predicted, and the general features of strain VV2014DJH are summarized in Tables 1, 2 and 3. The putative functions of the majority of the protein coding genes were prognosticated. The distribution of these genes into COGs functional categories was showed in Fig. 1, and the schematic circular representation of the VV2014DJH genome was created by Circos [27] in Fig. 2 to show the annotation information.

Three hundred and fifty-seven secretory proteins were detected in the 4617 proteins, and 45 proteins related to type I, II, III, IV and VI secretion systems were predicted, no protein related to type V and VII secretion systems was found. Type III secretion systems are complex bacterial structures that provide Gram-negative pathogens with unique virulence mechanisms. In the strain VV2014DJH, 184 effector proteins were detected by EffectiveT3 software.

The results of colinearity analysis for these three genome sequences of VV2014DJH, CMCP6 and YJ016 showed that the majority of the genomes were conserved

Table 1 General features of strain VV2014DJH genome

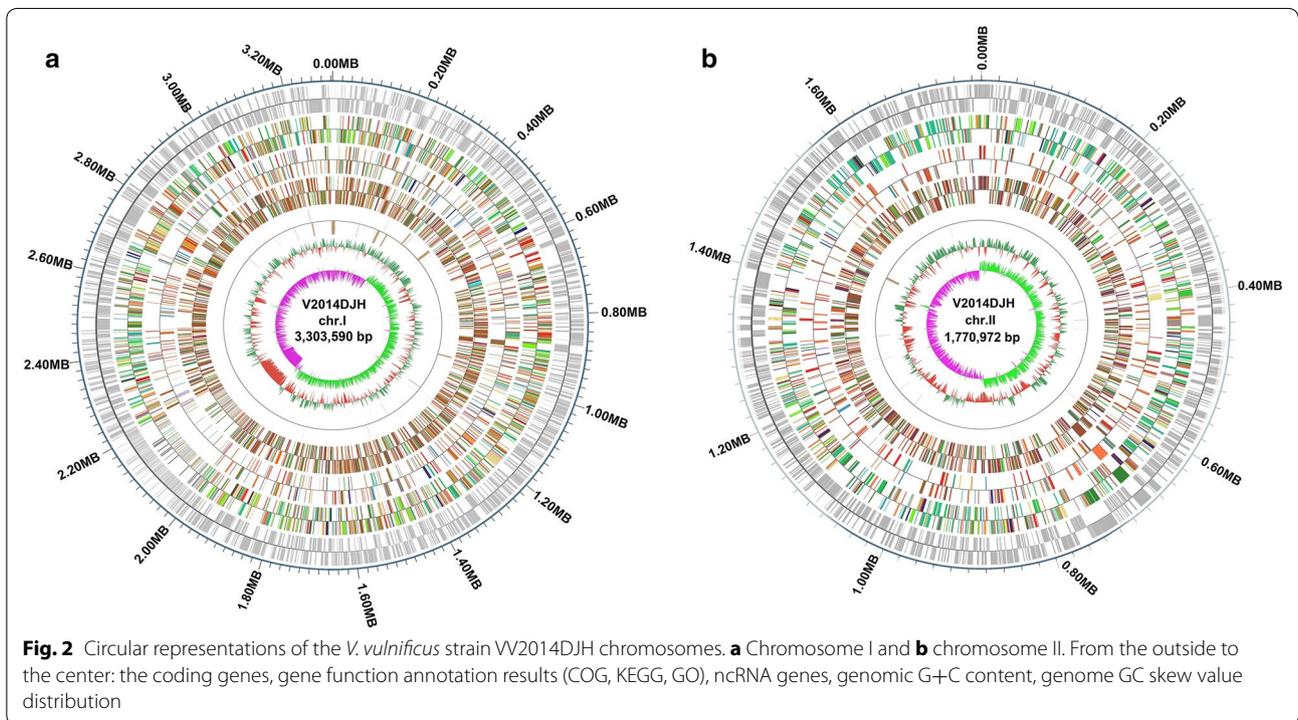
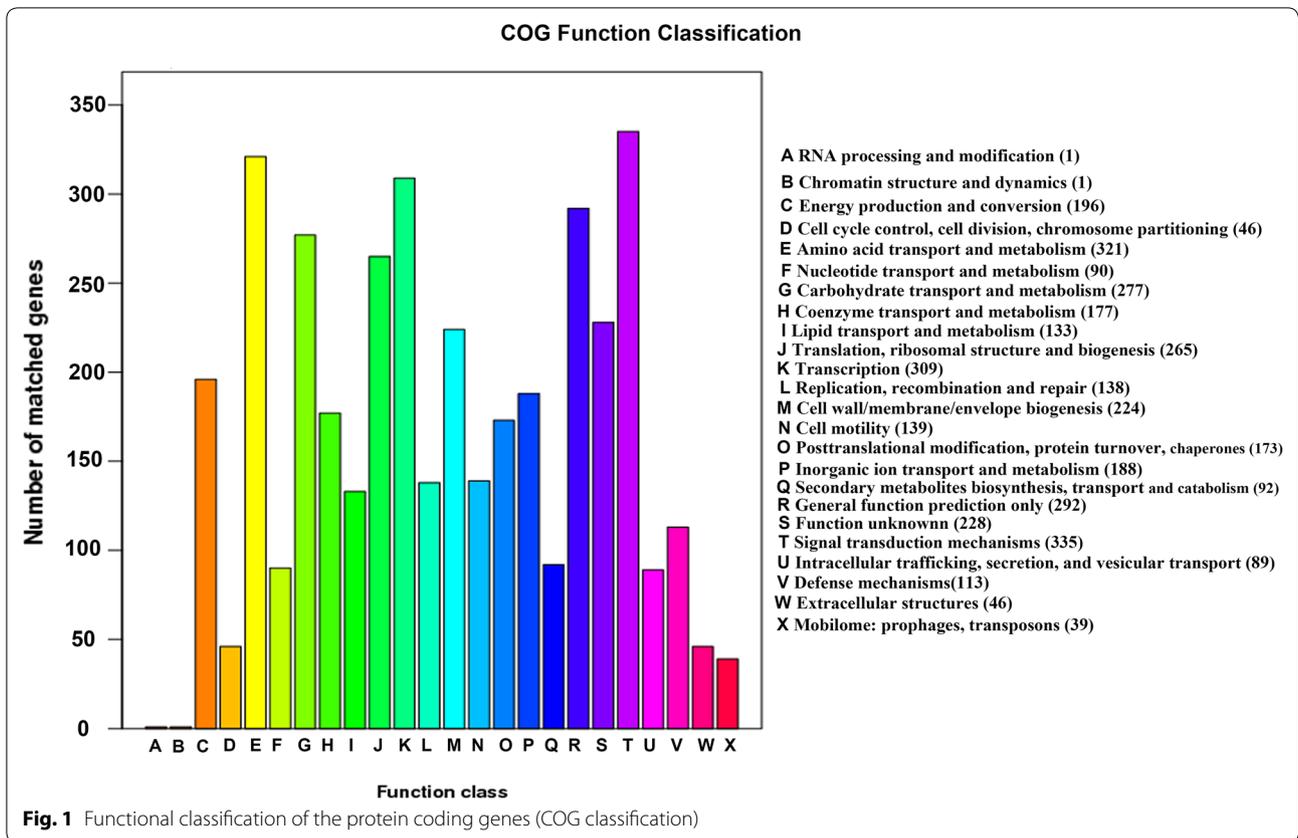
| Type | Number |
|----------------------|--------|
| tRNA | 120 |
| 5S rRNA | 13 |
| 16 S rRNA | 12 |
| 23 S rRNA | 12 |
| sRNA | 15 |
| Protein coding genes | 4617 |

Table 2 Interspersed repetitive sequences

| Type | Number | Total length (bp) | In genome (%) | Average length (bp) |
|---------|--------|-------------------|---------------|---------------------|
| LTR | 121 | 10,795 | 0.213 | 89 |
| DNA | 38 | 2554 | 0.050 | 68 |
| LINE | 40 | 2442 | 0.048 | 64 |
| SINE | 26 | 2291 | 0.045 | 90 |
| RC | 3 | 164 | 0.003 | 55 |
| Unknown | 1 | 98 | 0.002 | 98 |

Table 3 Tandem repetitive sequences

| Type | Number | Repeat size (bp) | Total length (bp) | In genome (%) |
|----------------|--------|------------------|-------------------|---------------|
| TR | 157 | 6–801 | 41,505 | 0.818 |
| Minisatellite | 72 | 11–51 | 3113 | 0.061 |
| Microsatellite | 1 | 6 | 124 | 0.002 |



in all three strains, although these were some transpositions between VV2014DJH and CMCP6, or VV2014DJH and YJ016. The Blast results showed that the pathogenic factors of VV2014DJH including the aforementioned type II and type VI secretion systems, flagellin C/D/E, and other pathogenic factors were also found in *V. vulnificus* VV2014DJH (Additional file 1: Fig S1, Table S1).

Significance

Our report provides an extended understanding on *Vibrio vulnificus* at genomic level and may provide useful information for epidemiological investigation, pathogenicity and virulence factors. In addition, the genome information is probably helpful for rapid detection and prevention of food poisoning by the strain.

Additional file

Additional file 1: Figure S1. Colinearity analysis of VV2014DJH, CMCP6 and YJ016. **Table S1.** The result of *Vibrio vulnificus* VV2014DJH blast against virulence factors of pathogenic bacteria database.

Authors' contributions

JHP was responsible for the most of the experiments; YS, WWY and HYM collected and analyzed the sequencing data; YJZ and MYZ reviewed the article. All authors read and approved the final manuscript.

Acknowledgements

This work was supported by the Major Science and Technology Special Projects of Zhejiang, China (No. 2013C03045-1) and the Medical Scientific Research Foundation of Zhejiang Province, China (No. 2012RCA013).

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

This whole genome sequences of *V. vulnificus* VV2014DJH has been deposited in the DDBJ/EMBL/GenBank database under the accession numbers CP019320 (ChI) and CP019321 (ChII).

Consent for publication

Not applicable.

Ethics approval and consent to participate

Our study was approved by the ethics committee of the Zhejiang Provincial Center for Disease Control and Prevention (ZJCDC), China.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 24 July 2017 Accepted: 8 November 2017

Published online: 21 November 2017

References

- Gulig PA, Bourdage KL, Starks AM. Molecular pathogenesis of *Vibrio vulnificus*. *J Microbiol*. 2005;43 Spec No:118–31.
- Jones MK, Oliver JD. *Vibrio vulnificus*: disease and pathogenesis. *Infect Immun*. 2009;77:1723–33. <https://doi.org/10.1128/iai.01046-08>.
- Hong G, et al. Emergency treatment of 16 patients with necrotizing fasciitis caused by *Vibrio vulnificus* infection complicated with septic shock. *Chin Med J*. 2014;127:1984–6.
- Chao WN, et al. Impact of timing of surgery on outcome of *Vibrio vulnificus*-related necrotizing fasciitis. *Am J Surg*. 2013;206:32–9. <https://doi.org/10.1016/j.amjsurg.2012.08.008>.
- Zhao H, et al. Correlations between clinical features and mortality in patients with *Vibrio vulnificus* infection. *PLoS ONE*. 2015;10:e0136019. <https://doi.org/10.1371/journal.pone.0136019>.
- Pan JH, et al. Drug resistance phenotype and molecular characteristics of a virulent *Vibrio vulnificus* strain isolated from a fatal case. *Dis Surveill*. 2016;31:229–32.
- Luo R, et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience*. 2012;1:18. <https://doi.org/10.1186/2047-217x-1-18>.
- Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*. 1997;25:955–64.
- Lagesen K, et al. RNAMmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res*. 2007;35:3100–8. <https://doi.org/10.1093/nar/gkm160>.
- Gardner PP, et al. Rfam: updates to the RNA families database. *Nucleic Acids Res*. 2009;37:D136–40. <https://doi.org/10.1093/nar/gkn766>.
- Saha S, Bridges S, Magbanua ZV, Peterson DG. Empirical comparison of ab initio repeat finding programs. *Nucleic Acids Res*. 2008;36:2284–94. <https://doi.org/10.1093/nar/gkn064>.
- Benson G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res*. 1999;27:573–80.
- Besemer J, Lomsadze A, Borodovsky M. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res*. 2001;29:2607–18.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215:403–10. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
- Kanehisa M. A database for post-genome analysis. *Trends Genet*. 1997;13:375–6.
- Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The KEGG resource for deciphering the genome. *Nucleic Acids Res*. 2004;32:D277–80. <https://doi.org/10.1093/nar/gkh063>.
- Kanehisa M, et al. From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res*. 2006;34:D354–7. <https://doi.org/10.1093/nar/gkj102>.
- Tatusov RL, Koonin EV, Lipman DJ. A genomic perspective on protein families. *Science*. 1997;278:631–7.
- Tatusov RL, et al. The COG database: an updated version includes eukaryotes. *BMC Bioinform*. 2003;4:41. <https://doi.org/10.1186/1471-2105-4-41>.
- Magrane M. UniProt Knowledgebase: a hub of integrated protein data. *Database*. 2011;2011:bar009. <https://doi.org/10.1093/database/bar009>.
- Ashburner M, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000;25:25–9. <https://doi.org/10.1038/75556>.
- Yu K, Zhang T. Construction of customized sub-databases from NCBI-nr database for rapid annotation of huge metagenomic datasets using a combined BLAST and MEGAN approach. *PLoS ONE*. 2013;8:e59831. <https://doi.org/10.1371/journal.pone.0059831>.
- Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods*. 2011;8:785–6. <https://doi.org/10.1038/nmeth.1701>.
- Jehl MA, Arnold R, Rattei T. Effective—a database of predicted secreted bacterial proteins. *Nucleic Acids Res*. 2011;39:D591–5. <https://doi.org/10.1093/nar/gkq1154>.
- Kurtz S, et al. Versatile and open software for comparing large genomes. *Genome Biol*. 2004;5:R12. <https://doi.org/10.1186/gb-2004-5-2-r12>.
- Chen L, Zheng D, Liu B, Yang J, Jin Q. VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. *Nucleic Acids Res*. 2016;44:D694–7. <https://doi.org/10.1093/nar/gkv1239>.
- Krzywinski M, et al. Circos: an information aesthetic for comparative genomics. *Genome Res*. 2009;19:1639–45. <https://doi.org/10.1101/gr.092759.109>.