Data Article

Draft genome sequence data of *Vibrio harveyi* VH1 isolated from a diseased tiger grouper, *Epinephelus fuscoguttatus*, cultured in Malaysia

Md. Ali Amatul-Samahah\textsuperscript{a,b}, Aslah Mohamad\textsuperscript{c}, Nurhidayu Al-saari\textsuperscript{d}, Mohd Zamri-Saad\textsuperscript{e}, Mohamad Noor Amal Azmai\textsuperscript{f}, Mohd Termizi Yusof\textsuperscript{g}, Md. Yasin Ina-Salwany\textsuperscript{a,c,*}, Mami Tanaka\textsuperscript{h}, Sayaka Mino\textsuperscript{h}, Tomoo Sawabe\textsuperscript{h}

\textsuperscript{a} Department of Aquaculture, Faculty of Agriculture, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
\textsuperscript{b} Freshwater Fisheries Research Division, Fisheries Research Institute Glami Lemt, Jelebu 71650, Malaysia
\textsuperscript{c} Aquatic Animal Health and Therapeutics Laboratory, Institute of Bioscience, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
\textsuperscript{d} International Institute for Halal Research And Training, International Islamic University Malaysia (IIUM), 53100 Jalan Gombak, Selangor, Malaysia
\textsuperscript{e} Department of Veterinary Laboratory Diagnosis, Faculty of Veterinary Medicine, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
\textsuperscript{f} Department of Biology, Faculty of Science, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
\textsuperscript{g} Department of Microbiology, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
\textsuperscript{h} Laboratory of Microbiology, Faculty of Fisheries Sciences, Hokkaido University, 3-1-1 Minato-cho, Hakodate 041-8611, Japan

\section*{ARTICLE INFO}

\textbf{Article history:}
Received 22 May 2022
Revised 28 July 2022
Accepted 3 August 2022
Available online 8 August 2022

Dataset link: Draft genome sequence data of \textit{Vibrio harveyi} VH1 (Original data)

\section*{ABSTRACT}

Vibriosis accounts for 66.7% of diseases reported in groupers’ cultures and affects almost all stages of growth. The disease could lead up to mortality up to 50% mortality, and it was reported that high stocking density and poor fish handling were among the factors that contributed to the disease dissemination. \textit{V. harveyi} has been reported to be among the causative agent and has caused acute mortality in cage groupers. In this study, we report the genome of \textit{V. harveyi} VH1 isolated from a
diseased tiger grouper *Epinephelus fuscoguttatus*, reared in a cage farm located in the coastal area of Langkawi.

© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

---

**Keywords:**
Aquaculture  
Vibriosis  
Grouper  
Genome  
*Vibrio* sp.

---

**Specification Table**

| Subject | Biology |
|---------|---------|
| Specific subject area | Microbiology, Genomics, Biotechnology |
| Type of data | Table, figures |
| How the data were acquired | The draft genome sequence was processed using MinION instrument, Oxford Nanopore Technology, UK. |
| Data format | Raw, analyzed and deposited |
| Description of data collection | *Vibrio harveyi* VH1 was isolated from the skin lesion samples originated from a cultured tiger grouper (*Epinephelus fuscoguttatus*) in Malaysia. Genomic DNA extraction and sequencing were performed. |
| Data source location | *Vibrio harveyi* VH1 was isolated from the skin lesion samples originated from a cultured tiger grouper (*Epinephelus fuscoguttatus*) in a cage farm located in the coastal area of Langkawi, Kedah, Malaysia (latitude and longitude: 6°13'16.1"N 99°46'07.5"E). *Vibrio harveyi* VH1 genome was analyzed at the Aquatic Animal Health and Therapeutics Laboratory, Institute of Bioscience, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia (latitude and longitude: 2° 59' 59.1684" N, 101° 43' 21.918" E). |
| Data accessibility | Data are publicly available at NCBI GenBank  
https://www.ncbi.nlm.nih.gov/nuccore/JAAIXX000000000  
https://www.ncbi.nlm.nih.gov/biosample/SAMN14091020  
https://www.ncbi.nlm.nih.gov/bioproject/PRJNA606422 |
| Related research article | Y.K. Chin, M.Y. Ina-Salwany, M. Zamri-Saad, M.N.A. Amal, A. Mohamad, J.Y. Lee, S. Annas & N. Al-saari, Efficacy of bath vaccination with a live attenuated *Vibrio harveyi* against vibriosis in Asian seabass fingerling, *Lates calcarifer*, Disease of Aquatic Organism 137: 167-173 (2020). https://doi.org/10.3354/dao03435 |

---

**Value of the Data**

- The draft genome of *Vibrio harveyi* VH1, isolated from cultured marine fish (*E. fuscoguttatus*) will be useful for further research on the virulence gene transfer of *V. harveyi* and *Vibrio* spp. in general.
- Data on the genome sequence of *Vibrio harveyi* VH1 can be used for comparative genomic studies with other *Vibrio* spp. disease isolates from other places.
- Data on the genome sequence of *Vibrio harveyi* VH1 could be used to identify and characterize important virulence factors that contribute to the pathogenesis.
- Data is useful for the bioinformatician and bacteriologist to better understand the genetic features of *Vibrio harveyi* VH1 and novel insights about its key virulence determinants.

---

**1. Data Description**

The draft genome of *V. harveyi* strain VH1 was reported in this finding. The draft genome assembly of *V. harveyi* strain VH1 has a length of 6,094,415 bp and a GC content of 44.8% (Fig. 1). The paired-end reads were assembled *de novo* into three huge contigs using Canu 1.6 with error and mismatch correction (N50, 3,675,737 bp). Annotation of the draft genome with RAST (Rapid Annotation Using Subsystem Technology) 2.0 identified 416 subsystems, 8763 coding sequences (CDS), and 161 total RNAs in the genome (Table 2). 108 coding sequences involves in virulence, diseases, and defense which inclusive of bacteriocins productions, ribosomally synthesized an-
Table 1
General information and genome sequencing project information of V. harveyi VH1.

| Items                                      | Description                       |
|--------------------------------------------|------------------------------------|
| Classification                             | Domain: Bacteria                  |
|                                            | Phylum: Proteobacteria             |
|                                            | Class: Gammaproteobacteria         |
|                                            | Order: Vibrionales                 |
|                                            | Family: Vibrionaceae               |
|                                            | Genus: Vibrio                      |
|                                            | Species: Vibrio harveyi            |
| Gram stain                                 | Negative                           |
| Cell shape                                 | Rod-shaped                         |
| Pigmentation                               | Non-pigmented                      |
| Sporulation                                | Non-sporulating                    |
| Optimum temperature                        | 25 °C                              |
| Salinity                                   | 30 ppt                             |
| Oxygen                                     | Aerobic                            |
| MICS Data                                  | GenBank                            |
| Investigation type                         | Bacteria                           |
| Project Name                               | Genome of Vibrio harveyi VH1       |
| Collection date                            | May 2017                           |
| Longitude and Latitude                     | 6°13’16.1"N 99°46’07.5"E           |
| Geographic location name                   | Langkawi, Kedah, Malaysia          |
| Environment biome                          | Coastal area                       |
| Environment feature                        | Cage farm                          |
| Environment material                       | Water                              |
| Depth                                      | 3.0-5.0 m                          |
| Biotic relationship                        | Free living                        |

Table 2
Genome features of V. harveyi VH1.

| Attribute                      | Description         |
|--------------------------------|---------------------|
| Genome size (bp)               | 6,094,415 bp        |
| G+C content (%)                | 44.8%               |
| CDS (coding sequences)         | 8763                |
| rRNA number                    | 37                  |
| tRNA number                    | 124                 |
| Genbank accession              | JAAIKJ0000000000    |
| BioSample accession            | SAMN14091020        |
| BioProject accession           | PRJNA606422         |

Tibacterial peptides, and resistance to antibiotics (Fig. 2). Four coding sequences were found to have identity with phages, prophages, transposable elements and plasmids [1–4].

Besides that, 219 coding sequences involved in motility and chemotaxis, 120 coding sequences involved in regulation and cell signaling, 139 coding sequences involved in stress response, and 158 coding sequences involved in cell respiration. The genome sequence of V. harveyi VH1 serves as an additional genomic resource for comparative genomic studies of other V. harveyi strains that infected marine fish (Fig. 2).
Fig. 1. Circular map of the *V. harveyi* VH1 genome. From the outermost circle to the center: CDSs on forward strand (including tRNA, rRNA and mRNA), open reading frame (ORF), Contigs, GC skew+ and GC skew−, GC content, BLAST results, and the marker of genome size is the innermost circle.
2. Experimental Design, Materials and Methods

V. harveyi strain VH1 was isolated from the skin lesion samples originated from a male diseased tiger grouper Epinephelus fuscoguttatus, reared in a cage farm located in the coastal area of Langkawi, Malaysia (Table 1). The V. harveyi isolate was cultured and maintained in thiosulfate-citrate-bile salts (TCBS) (Oxoid) agar and tryptone soy broth (TSB) (Oxoid), supplemented with NaCl (1.5% w/v) at 30 °C. A TCBS agar is a selective medium for enteropathogenic Vibrio spp. When cultured on TCBS agar, the colonies of pathogenic V. harveyi strain VH1 appeared as yellow colonies. Genomic DNA of V. harveyi VH1 was extracted from the culture using the DNA kit (Thermo Fisher Scientific). Sequencing library was prepared using the Rapid Barcoding Kit (SQK-RBK001) (Oxford Nanopore Technologies, Oxford, UK) as per instruction in the manual provided by the manufacturer. The library was then loaded to a MinION R9 flow cell (FLO-MIN106) (Oxford Nanopore Technologies, Oxford, UK), and the sequencing analysis was performed using MinKNOW software version 1.7.14. Fast5s from Nanopore sequencing were basecalled with ONT Albacore Sequencing Pipeline software version 2.0.2 and reads passing the internal test were used for subsequent analysis [5]. Porechop 0.2.2 (https://github.com/rrwick/Porechop) was used for debarcoding and adaptor trimming. Nanopore reads were assembled using Canu 1.6 [6]. For Nanopore-only assembly, output contigs were polished using Nanopolish software version 0.8.1 (https://github.com/jts/nanopolish). Contigs from Canu 1.6 were manually closed based on the assembly graph with Bandage software version 0.8.1 [7].

Ethics Statements

The study was conducted according to the guidelines by the Animal Care and Use Committee Universiti Putra Malaysia (UPM/IACUC/AUP-R078/2019). All animal experiments are reported in compliance with the ARRIVE guidelines and carried out in accordance with the U.K. Animals
(Scientific Procedures) Act, 1986 and associated guidelines, EU Directive 2010/63/EU for animal experiments, or the National Institutes of Health guide for the care and use of Laboratory animals (NIH Publications No. 8023, revised 1978).

CRediT Author Statement

The authors here declare their individual contributions:

Md. Ali Amatul-Samahah: Writing – original draft and editing; Aslah Mohamad: Investigation and writing; Nurhidayu Al-saari: Investigation and writing; Mohd Zamri-Saad: Validation and reviewing; Mohamad Noor Amal Azmai: Validation and reviewing; Mohd Termizi Yusof: Validation and reviewing; Ina-Salwany Md.Yasin: Validation, reviewing, editing & supervision; Mami Tanaka: Investigation, data curation, software & validation; Sayaka Mino: Data curation, software, validation & investigation; Tomoo Sawabe: Software, validation, investigation & supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Draft genome sequence data of Vibrio harveyi VH1 (Original data) (NCBI).

Acknowledgments

This research was supported by a research grant from Transdisciplinary Research Grant Scheme (TRGS/1/2019/UPM/02/5) and Higher Institution Centre of Excellence (HiCoE) funded by the Ministry of Education (MOE), Malaysia. We are thankful to all the members of the Laboratory of Microbiology, Faculty of Fisheries Sciences, Hokkaido University, for providing materials and nanopore sequencing facilities needed to carry out this project.

References

[1] R. Overbeek, R. Olson, G.D. Pusch, G.J. Olsen, J.J. Davis, T. Disz, R.A. Edwards, S. Gerdes, B. Parello, M. Shukla, V. Vonstein, A.R. Wattam, F. Xia, R. Stevens, The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST), Nucleic Acids Res. 5 (2014) 8365, doi:10.1093/nar/gkt08365.
[2] T. Brettin, J.J. Davis, T. Disz, R.A. Edwards, S. Gerdes, G.J. Olsen, R. Olson, R. Overbeek, B. Parello, G.D. Pusch, M. Shukla, J.A. Thomason, R. Stevens, V. Vonstein, A.R. Wattam, F. Xia, RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes, Sci. Rep. 42 (2015) 206–214, doi:10.1093/nar/gkt1226.
[3] M.A. Amatul-Samahah, S. Muthukrishnan, N. Al-saari, N.F.M. Ikhsan, M. Zamri-Saad, M.N.A. Azmai, M.T. Yusof, I.S.M. Yasin, M. Tanaka, S. Mino, T. Sawabe, Genome sequence of Vibrio parahaemolyticus CSA causing acute hepatopancreatic necrosis disease in shrimp isolated from a Malaysian shrimp culture pond, Gene Rep. 27 (2022) 101601, doi:10.1016/j.genrep.2022.101601.
[4] M.A. Amatul-Samahah, S. Muthukrishnan, W.H.H.W. Omar, N.F.M. Ikhsan, M.Y. Ina-Salwany, Vibrio spp. associated with acute hepatopancreatic necrosis disease (AHPND) found in penaeid shrimp pond from east coast of peninsular Malaysia, J. Environ. Biol. 41 (2020) 1160–1170, doi:10.22438/jeb/41/5(SI)/MS_07.
[5] M. Tanaka, S. Mino, Y. Ogura, T. Hayashi, T. Sawabe, Availability of Nanopore sequences in the genome taxonomy for Vibrionaceae systematics: rumoienosis clade species as a test case, PeerJ 6 (2018) e5018, doi:10.7717/peerj.5018.
[6] S. Koren, B.P. Walenz, K. Berlin, J.R. Miller, N.H. Bergman, A.M. Philippey, Canu: A scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation, Genome Res. 27 (2017) 722–736, doi:10.1101/gr.215087.116.
[7] R. Wick, L. Judd, C. Gorrie, K. Holt, Completing bacterial genome assemblies with multiplex MinION sequencing, Microb. Genom. 3 (2017) 10, doi:10.1099/mgen.0.000132.