AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

PROKARYOTES

New Sequence Types of *Vibrio parahaemolyticus* Isolated from a Malaysian Aquaculture Pond, as Revealed by Whole-Genome Sequencing

Soon Man Foo,[a] Wilhelm Wei Han Eng,[a,b] Yin Peng Lee,[a,b] Kimberly Gui,[a] Han Ming Gan[a,b]

School of Science, Monash University Malaysia, Petaling Jaya, Malaysia[a]; Genomics Facility, Tropical Medicine and Biology Platform, Monash University Malaysia, Petaling Jaya, Malaysia[b]

ABSTRACT The acquisition of *Photorhabdus* insect-related (Pir) toxin-like genes in *Vibrio parahaemolyticus* has been linked to hepatopancreatic necrosis disease in shrimp. We report the whole-genome sequences of genetically virulent and avirulent *V. parahaemolyticus* isolated from a Malaysian aquaculture pond and show that they represent previously unreported sequence types of *V. parahaemolyticus*.

Received 14 March 2017 **Accepted** 16 March 2017 **Published** 11 May 2017

Citation Foo SM, Eng WWH, Lee YP, Gui K, Gan HM. 2017. New sequence types of *Vibrio parahaemolyticus* isolated from a Malaysian aquaculture pond, as revealed by whole-genome sequencing. Genome Announc 5: e00302-17. https://doi.org/10.1128/genomeA.00302-17.

Address correspondence to Han Ming Gan, gan.han.ming@monash.edu.

Vibrio parahaemolyticus is a marine Gram-negative bacterium (1) that has been occasionally associated with acute hepatopancreatic necrosis disease (AHPND) in the white leg shrimp, *Litopenaeus vannamei*, resulting in severe economic losses in shrimp production in Southeast Asian countries (2). *Photorhabdus* insect-related (Pir) toxin-like genes have been recently identified in various AHPND-causing *V. parahaemolyticus* strains and these genes (*pirA*- and *pirB*-like) were shown to be the primary virulence factor in these strains (3).

Five *V. parahaemolyticus* strains (MVP1, MVP2, MVP4, MVP6, and MVP9) were isolated from a shrimp pond that was tested positive for *V. parahaemolyticus* harboring the Pir genes. Genomic DNA was extracted from a 2-day-old marine nutrient agar culture (ATCC Medium 8) using the Solokov method (4). Library preparation was performed using the NexteraXT DNA library preparation kit (Illumina, San Diego, CA) according to the manufacturer's instructions and sequenced on a MiSeq desktop sequencer (2 × 75-bp and 2 × 250-bp configurations) located at the Monash University Malaysia Genomics Facility.

Nextera adapter trimming was performed using Trimmomatic version 0.32 (5) and the filtered paired-end reads were assembled using SPAdes version 3.8.1 (6). After the removal of short (<300 bp) and/or low-coverage (<2×) contigs, *in silico* scaffolding and gap closing were performed using SSPACE version 2.1 (7) and Gapfiller version 1.10 (8), respectively. To confirm the identity of the isolated strains as *V. parahaemolyticus*, Jspecies version 1.2 (9) was used to calculate the average nucleotide identity of strains MVP1, -2, -4, -6, and -9 in comparison to the whole genome of *V. parahaemolyticus* DSM 10027^T. Subsequently, gene prediction was performed using Prodigal version 2.6 (10) and searched against the multilocus sequence typing (MLST) locus database (http://www.mlst.net/) to infer the sequence type of each sequenced strain based on their genetic similarity to seven housekeeping genes, namely, *pyrC*, *gyrB*, *recA*, *dnaE*, *tnaA*, *pntA*, and *dtdS*. The identification of the Pir genes was performed via a local BlastN search against *pirA* (GenBank accession no. AIL49948.1) and *pirB* (GenBank accession no. AIL49949.1).

A summary of the assembly statistics for the genomes of all isolates is available in Table 1. All five strains exhibited more than 95% average nucleotide identity (ANI) to

TABLE 1 Accession numbers and genetic analyses of *V. parahaemolyticus* strains reported in this study

Strain	Accession no.	Genome size (bp)	N_{50} (bp)	No. of contigs	*pyrC*[a]	*gyrB*[a]	*recA*[a]	*dnaE*[a]	*tnaA*[a]	*pntA*[a]	*dtdS*[a]	Pir[b]
MVP1	MQMQ01000000	5,230,330	60,033	172	303	143[c] (591/592)	218	308[c] (556/557)	26	30	355[c] (457/458)	+
MVP2	MSBY01000000	5,275,177	129,821	89	303	143[c] (591/592)	218	308[c] (556/557)	26	30	355[c] (457/458)	+
MVP4	MSBZ01000000	5,270,749	94,277	122	27	141	31[c] (728/729)	110[c] (556/557)	26	18	232	−
MVP6	MSCA01000000	5,195,990	47,800	190	303	143[c] (591/592)	218	308[c] (556/557)	26	30	355[c] (457/458)	+
MVP9	MSCB01000000	4,967,664	87,871	115	54[c] (492/493)	144	116	28	61	26	252	−

[a]Numerical values indicate the MLST allele for the respective genes.

[b]+, presence of both *pirA* and *pirB* genes; −, absence of both *pirA* and *pirB* genes.

[c]Closest allele hit, with values in parentheses indicating the number of positions over the length of the gene fragment where all of the bases at that position are identical.

the type strain of *V. parahaemolyticus*. Based on the lack of 100% sequence identity to seven housekeeping genes, new MLST sequence types of *V. parahaemolyticus* were identified in this study (Table 1). Subsequent blastN searches showed that strains MVP1, MVP2, and MVP6 contain the identical nucleotide sequence for all seven housekeeping genes (Table 1, data not shown), thus classifying them as the same sequence type. In addition to sharing the identical sequence type, these 3 strains also harbor the Pir toxin genes. On the contrary, strains MVP4 and MVP9 belong to two different sequence types and do not harbor the Pir toxin genes, suggesting a potential association between the sequence types and the presence of Pir toxin genes in *V. parahaemolyticus*.

Accession number(s). This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the accession numbers listed in Table 1. The versions described in this paper are the first versions.

ACKNOWLEDGMENTS

This research was supported by the Malaysian Ministry of Education (grant code FRGS/1/2016/STG05/MUSM/03/1) and by the Monash University Malaysia Tropical and Medicine Biology Multidisciplinary Platform (grant code 5140754-313). We are grateful to the staff at Lab-Ind Resource SB for technical assistant in the laboratory.

REFERENCES

1. Tanabe T, Miyamoto K, Tsujibo H, Yamamoto S, Funahashi T. 2015. The small RNA Spot 42 regulates the expression of the type III secretion system 1 (T3SS1) chaperone protein VP1682 in *Vibrio parahaemolyticus*. FEMS Microbiol Lett 362:fnv173. https://doi.org/10.1093/femsle/fnv173.
2. Sirikharin R, Taengchaiyaphum S, Sanguanrut P, Chi TD, Mavichak R, Proespraiwong P, Nuangsaeng B, Thitamadee S, Flegel TW, Sritunyalucksana K. 2015. Characterization and PCR detection of binary, pir-like toxins from *Vibrio parahaemolyticus* isolates that cause acute hepatopancreatic necrosis disease (AHPND) in shrimp. PLoS One 10:e0126987. https://doi.org/10.1371/journal.pone.0126987.
3. Han JE, Tang KFJ, Tran LH, Lightner DV. 2015. Photorhabdus insect-related (Pir) toxin-like genes in a plasmid of *Vibrio parahaemolyticus*, the causative agent of acute hepatopancreatic necrosis disease (AHPND) of shrimp. Dis Aquat Organ 113:33–40. https://doi.org/10.3354/dao02830.
4. Sokolov EP. 2000. An improved method for DNA isolation from mucopolysaccharide-rich molluscan tissues. J Molluscan Stud 66:573–575. https://doi.org/10.1093/mollus/66.4.573.
5. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
6. 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. https://doi.org/10.1089/cmb.2012.0021.
7. Boetzer M, Pirovano W. 2014. SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information. BMC Bioinformatics 15:211. https://doi.org/10.1186/1471-2105-15-211.
8. Boetzer M, Pirovano W. 2012. Toward almost closed genomes with GapFiller. Genome Biol 13:R56. https://doi.org/10.1186/gb-2012-13-6-r56.
9. Richter M, Rosselló-Móra R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A 106:19126–19131. https://doi.org/10.1073/pnas.0906412106.
10. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119. https://doi.org/10.1186/1471-2105-11-119.