Genome Sequence of *Bacillus* sp. Strain UMTAT18 Isolated from the Dinoflagellate *Alexandrium tamiyavanichii* Found in the Straits of Malacca

Muhd Danish-Daniel,[a,b] Gan Han Ming,[c,d] Mohd Ezhar Mohd Noor,[a] Yeong Yik Sung,[b] Gires Usup[e]

School of Fisheries and Aquaculture Sciences, Universiti Malaysia Terengganu, Kuala Terengganu, Terengganu, Malaysia[a]; Institute of Marine Biotechnology, Universiti Malaysia Terengganu, Kuala Terengganu, Terengganu, Malaysia[b]; School of Science, Monash University Malaysia, Jalan Lagoon Selatan, Bandar Sunway, Petaling Jaya, Selangor, Malaysia[c]; Monash University Malaysia Genomics Facility, Jalan Lagoon Selatan, Bandar Sunway, Petaling Jaya, Selangor, Malaysia[d]; Faculty of Science and Technology, School of Environment and Natural Resource Sciences, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia[e]

***Bacillus* sp. strain UMTAT18 was isolated from the harmful dinoflagellate *Alexandrium tamiyavanichii*. Its genome consists of 5,479,367 bp with 5,546 open reading frames, 102 tRNAs, and 29 rRNAs. Gene clusters for biosynthesis of nonribosomal peptides, bacteriocin, and lantipeptide were identified. It also contains siderophore and genes related to stress tolerance.**

Address correspondence to Muhd Danish-Daniel, mdda@umt.edu.my.

Many *Bacillus* spp. have been reported to have biotechnological importance (1), while many others of this genus are also well-known pathogens (2). *Bacillus* spp. with the capacity to produce antimicrobial compounds and enzymes important to industry have been reported (3, 4). Some *Bacillus* spp. are used as probiotics in aquaculture to improve water quality, to enhance the hosts' digestive systems, or to prevent the spread of diseases (5). *Bacillus* sp. strain UMTAT18 was isolated from the paralytic shellfish toxin-producing dinoflagellate *Alexandrium tamiyavanichii* found in the Straits of Malacca, Malaysia. We sequenced the genome of this strain to gain insights into its potential role as a probiotic for aquaculture applications.

Bacillus sp. strain UMTAT18 was cultured in marine broth 2216 (MB; Difco). Genomic DNA was then extracted using the GF-1 nucleic acid extraction kit (Vivantis, Malaysia). Sequencing was performed on the Illumina HiSeq2000 platform, generating 38,220,522 raw FASTQ paired-end reads. Two million reads were subsampled for error correction and *de novo* assembly using SPAdes version 3.1.0 (6). The resulting contigs were used for scaffolding, followed by gap-closing using SSPACE version 2.0 and GapFiller version 1.11 (7, 8). Sixty-nine gap-filled contigs with an N_{50} of 5 sequences longer than 317,317 bp were produced, and the total sequence length was 5,479,367 bp with a $73\times$ coverage.

The Prokka version 1.8 annotation pipeline, comprised of Prodigal version 2.60, RNAmmer version 1.2, and Aragorn version 1.2.36, was used to annotate the genome, predicting 5,546 open reading frames, 29 rRNAs, and 102 tRNAs (9–12). The predicted 16S rRNA was queried with BLASTn (13) against the nucleotide collection database, identifying our sample as belonging to the genus *Bacillus*. Further validation of the species was performed using GGDC2.1 *in silico* genome-to-genome comparison of the UMTAT18 strain to other closely related *Bacillus* spp., showing that strain UMTAT18 is a novel species within the *Bacillus cereus* group (14). InterProScan5 was used to provide additional annotation to the predicted protein sequences (15). Furthermore, antiSMASH was used to identify the presence of secondary metabolite biosynthesis gene clusters in the genome (16).

Bacillus sp. UMTAT18 carries genes involved in the biosynthesis of bacillibactin and siderophore, which suggests that this strain could be applied in aquaculture to deprive pathogens that require iron for growth and pathogenesis (17). The genome of this strain also contains genes involved in the synthesis of antimicrobial properties like bacteriocin, lantipeptide, and nonribosomal peptide synthase, which allow it to compete with other bacteria in the aquatic environment. This strain possesses other environmental stress-adaptation-related genes, including those for osmotic and oxidative stress. The genome of strain UMTAT18 provides valuable information at the molecular level as a candidate probiotic for aquaculture applications.

Accession number(s). This whole-genome shotgun project has been deposited in DDBJ/EMBL/GenBank under the accession number JSFD00000000.

REFERENCES

1. **Hernández-González IL, Olmedo-Álvarez G**. 2016. Draft whole-genome sequence of the type strain *Bacillus horikoshii* DSM 8719. Genome Announc 4(4):e00641-16. http://dx.doi.org/10.1128/genomeA.00641-16.
2. **Vilas-Bôas GT, Peruca AP, Arantes OM**. 2007. Biology and taxonomy of

Bacillus cereus, *Bacillus anthracis*, and *Bacillus thuringiensis*. Can J Microbiol **53**:673–687. http://dx.doi.org/10.1139/W07-029.

3. **Thenmozhi R, Nithyanand P, Rathna J, Pandian SK**. 2009. Antibiofilm activity of coral-associated bacteria against different clinical M serotypes of *Streptococcus pyogenes*. FEMS Immunol Med Microbiol **57**:284–294. http://dx.doi.org/10.1111/j.1574-695X.2009.00613.x.

4. **Joo HS, Choi JW**. 2012. Purification and characterization of a novel alkaline protease from *Bacillus horikoshii*. J Microbiol Biotechnol **22**:58–68. http://dx.doi.org/10.4014/jmb.1109.09006.

5. **Cruz PM, Ibáñez AL, Hermosillo OAM, Saad HCR**. 2012. Use of probiotics in aquaculture. ISRN Microbiol **2012**:916845. http://dx.doi.org/10.5402/2012/916845.

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. http://dx.doi.org/10.1089/cmb.2012.0021.

7. **Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W**. 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics **27**:578–579. http://dx.doi.org/10.1093/bioinformatics/btq683.

8. **Boetzer M, Pirovano W**. 2012. Toward almost closed genomes with Gap-Filler. Genome Biol **13**:R56. http://dx.doi.org/10.1186/gb-2012-13-6-r56.

9. **Seemann T**. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics **30**:2068–2069. http://dx.doi.org/10.1093/bioinformatics/btu153.

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. http://dx.doi.org/10.1186/1471-2105-11-119.

11. **Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW**. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res **35**:3100–3108. http://dx.doi.org/10.1093/nar/gkm160.

12. **Laslett D, Canback B**. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res **32**:11–16. http://dx.doi.org/10.1093/nar/gkh152.

13. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ**. 1990. Basic local alignment search tool. J Mol Biol **215**:403–410. http://dx.doi.org/10.1016/S0022-2836(05)80360-2.

14. **Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M**. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics **14**:60. http://dx.doi.org/10.1186/1471-2105-14-60.

15. **Jones P, Binns D, Chang HY, Fraser M, Li W, Mcanulla C, Mcwilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S**. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics **30**:1236–1240. http://dx.doi.org/10.1093/bioinformatics/btu031.

16. **Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, Weber T, Takano E, Breitling R**. 2011. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res **39**(suppl 2):W339–W346. http://dx.doi.org/10.1093/nar/gkr466.

17. **Jagadeesh KS, Kulkarni JH, Krishnaraj PU**. 2001. Evaluation of the role of fluorescent siderophore in the biological control of bacterial wilt in tomato using Tn5 mutants of fluorescent *Pseudomonas* sp. Curr Sci **81**:882.