Complete Genome Sequence and Methylome Analysis of *Aeromonas hydrophila* Strain YL17, Isolated from a Compost Pile

Yan-Lue Lim,[a] Richard J. Roberts,[b] Robson Ee,[a] Wai-Fong Yin,[a] Kok-Gan Chan[a]

Division of Genetics and Molecular Biology, Institute of Biological Sciences, Faculty of Science, University of Malaya, Kuala Lumpur, Malaysia[a]; New England Biolabs, Ipswich, Massachusetts, USA[b]

In this report, we announce the complete genome sequence of *Aeromonas hydrophila* strain YL17. Single-molecule real-time (SMRT) DNA sequencing was used to generate the complete genome sequence and the genome-wide DNA methylation profile of this environmental isolate. A total of five unique DNA methyltransferase recognition motifs were reported here.

Studies of prokaryotic DNA methylation are gaining major attention because of its significant implication in bacterial physiology and virulence. However, the methylomes of a vast number of bacterial organisms still remain unexplored, including *Aeromonas hydrophila*. Due to their ubiquity in the aquatic environment and their enterotoxigenic properties, *A. hydrophila* strains are extensively studied as they are increasingly regarded as emerging human pathogens (1, 2). In this study, we report the complete genome sequence of *Aeromonas hydrophila* strain YL17, a compost pile isolate, and its whole-genome and methylome analyses.

Genomic DNA extraction was performed using a MasterPure DNA purification kit (Epicentre, USA) followed by gel electrophoresis and NanoDrop spectrophotometer absorbance measurement (Thermo Scientific, USA) to assess the integrity and purity of the extracted genomic DNA. A Qubit fluorometer was used in combination with a Qubit dsDNA broad range (BR) assay kit (Invitrogen, USA) for DNA quantitation. The genomic DNA was sheared to an average size of 10-Kb using G-tubes (Covaris, USA) and was subjected to purification and concentration using Ampure PB beads (Pacific Biosciences, USA). Briefly, 5 µg of sheared genomic DNA was then proceeded with SMRTbell library construction following the "Procedure & Checklist-10-kb Template Preparation and Sequencing" protocol (http://www.pacb.com/wp-content/uploads/2015/09/Procedure-Checklist-10-kb-Template-Preparation-and-Sequencing.pdf). A Pacific Biosciences RSII sequencer (Pacific Biosciences, USA) was used to sequence the 10-Kb library by using C2 chemistry and four

single-molecule real-time (SMRT) cells. 173.05-fold coverage was achieved and the reads were assembled using the Hierarchical genome assembly process (HGAP) version 2 (PacBio DevNet; Pacific Biosciences) (3). The genome sequences were successfully assembled to closure by using a seed read length of 8,000 bp and yielded a single contig (4,808,605 bp). The assembled genome was then circularized prior to annotation using the NCBI Prokaryotic Genomes Annotation Pipeline (PGAP) and Rapid Annotations using Subsystems Technology (RAST) (4).

The circularized chromosome of *A. hydrophila* strain YL12 was imported into the SMRT portal and base modification and methyltransferase motif detection and analysis were performed using the RS_Modification_and_Motif Analysis workflow (5). Six DNA methyltransferase specificities were detected as listed in Table 1. Interestingly, five motifs detected in this genome (besides G^{m6}ATC) are novel and have not been reported previously. Restriction modification (RM) system annotation was carried out using the SEQWARE computer resource (6) and enabled the reliable assignment of candidate methyltransferase genes to each specificity based on their RM types and their similarity to previously characterized methyltransferase homologs (Table 1). The methylome data have been deposited in REBASE (7).

Nucleotide sequence accession numbers. The complete genome sequence of *Aeromonas hydrophila* YL17 is available in GenBank under the accession number CP007518. The version described in this paper is the second version, CP007518.2.

TABLE 1 Summary of methylation motif identified in *Aeromonas hydrophila* strain YL17 and the most likely genes responsible for each motif

Motif[a]	Percentage	Unique	Probable gene	Methylation type	RM type
G^{m6}ATC	99.97%	No	M.AhyYL17Dam	m6A	II
CC^{m6}AGNNNNNN**T**GAY	100%	Yes	S.AhyYL17ORF17275P	m6A	I
CC^{m6}AYNNNNNN**T**RTC	100%	Yes	S.AhyYL17ORF670P	m6A	I
CT^{m6}ANNNNNNNG**T**TC	100%	Yes	S.AhyYL17ORF11910P	m6A	I
S**C**TCGAKG	41.07%	Yes	M.AhyYL17ORF13945P	m4C or m5C	II
YAAMG^{m6}A**G**	99.86%	Yes	AhyYL17I	m6A	II

[a] Modified bases are highlighted in bold.

REFERENCES

1. **Seshadri R, Joseph SW, Chopra AK, Sha J, Shaw J, Graf J, Haft D, Wu M, Ren Q, Rosovitz MJ, Madupu R, Tallon L, Kim M, Jin S, Vuong H, Stine OC, Ali A, Horneman AJ, Heidelberg JF**. 2006. Genome sequence of *Aeromonas hydrophila* ATCC 7966^T: Jack of all trades. J Bacteriol **188**:8272–8282. http://dx.doi.org/10.1128/JB.00621-06.

2. **Jones BL, Wilcox MH**. 1995. *Aeromonas* infections and their treatment. J Antimicrob Chemother **35**:453–461. http://dx.doi.org/10.1093/jac/35.4.453.

3. **Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J**. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods **10**:563–569. http://dx.doi.org/10.1038/nmeth.2474.

4. **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**. 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics **9**:75. http://dx.doi.org/10.1186/1471-2164-9-75.

5. **Clark TA, Murray IA, Morgan RD, Kislyuk AO, Spittle KE, Boitano M, Fomenkov A, Roberts RJ, Korlach J**. 2012. Characterization of DNA methyltransferase specificities using single-molecule, real-time DNA sequencing. Nucleic Acids Res **40**:e29. http://dx.doi.org/10.1093/nar/gkr1146.

6. **Murray IA, Clark TA, Morgan RD, Boitano M, Anton BP, Luong K, Fomenkov A, Turner SW, Korlach J, Roberts RJ**. 2012. The methylomes of six bacteria. Nucleic Acids Res **40**:11450–11462. http://dx.doi.org/10.1093/nar/gks891.

7. **Roberts RJ, Vincze T, Posfai J, Macelis D**. 2015. Rebase—a database for DNA restriction and modification: enzymes, genes and genomes. Nucleic Acids Res **43**:D298–D299. http://dx.doi.org/10.1093/nar/gku1046.