Draft Genome Sequences of Five *Legionella pneumophila* Strains Isolated from Environmental Water Samples

Kenta Watanabe,[a,b] Haruo Suzuki,[c] Ryo Nakao,[d] Takashi Shimizu,[a,b] Masahisa Watarai[a,b]

The United Graduate School of Veterinary Science, Yamaguchi University, Yamaguchi, Japan[a]; Joint Faculty of Veterinary Medicine, Laboratory of Veterinary Public Health, Yamaguchi University, Yamaguchi, Japan[b]; Department of Environmental Science and Engineering, Graduate School of Science and Engineering, Yamaguchi University, Yamaguchi, Japan[c]; Unit of Risk Analysis and Management, Hokkaido University Research Center for Zoonosis Control, Sapporo, Hokkaido, Japan[d]

Legionella pneumophila is the causative agent of legionellosis. Here, we report the draft genome sequences of five *L. pneumophila* strains, Bnt314, Ofk308, Twr292, Ymg289, and Ymt294, isolated from environmental water samples. Comparative analyses of these genomes may reveal the survival mechanisms and virulence of *L. pneumophila* in the natural environment.

*L*egionella pneumophila is the most frequent cause of legionellosis in humans. *L. pneumophila* can withstand temperatures of 0 to 68°C and a pH range of 5.0 to 8.5 and can survive in most environments for long periods (1). Furthermore, *L. pneumophila* is a facultative intracellular bacterium and can survive within free-living protozoa, such as amoebae and ciliates, in environmental waters (2–4). These environmental habitats of *L. pneumophila* have been suggested as important evolutionary sites to develop their virulence traits in humans, although the detailed mechanisms remain largely unclear.

Previously, we isolated five *L. pneumophila* strains from environmental water samples (5). *L. pneumophila* strains Ofk308, Twr292, and Ymt294 were isolated from an Ashiyu foot spa strain Ymg289 was isolated from a water fountain, and strain Bnt314 was isolated from a pond. These strains were classified as *L. pneumophila* serotype I or IV. Here, we report the draft genome sequences of the five *L. pneumophila* strains (Table 1).

Chromosomal DNA was extracted from an overnight culture of the five *L. pneumophila* strains using the DNeasy blood and tissue kit (Qiagen). Whole-genome sequencing of the five environmental strains was performed using paired-end sequencing on the Illumina MiSeq kit version 3. The sequencer produced 300-bp paired-end reads that were obtained from 550-bp inserts. The quality of the reads was checked using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), and the raw sequences were trimmed to 250 bp using FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). *De novo* genome assembly was performed using SPAdes version 3.5.0 (6). After the removal of low-coverage contigs, the resulting contigs were ordered against the complete genome of *L. pneumophila* subsp. *pneumophila* strain Philadelphia 1 chromosome (7) using the Mauve aligner (8). Genome annotation was performed using Prokka version 1.11 (9). The genome sizes and G+C contents were estimated for all contigs of each strain using G-language Genome Analysis Environment version 1.9.0 (http://www.g-language.org) (10). Among the five strains, the genome size and G+C content varied from 3.39 to 3.69 Mb and 38.2 to 38.4%, respectively (Table 1) and were close to those of the reference genome of strain Philadelphia 1. The genome statistics for the five environmental strains are summarized in Table 1.

Nucleotide sequence accession numbers. The sequences have been deposited as whole-genome shotgun projects at DDBJ/EMBL/GenBank under the accession numbers listed in Table 1. The versions described in this paper are the first versions.

ACKNOWLEDGMENT

TABLE 1 Information for draft genome sequences of five *L. pneumophila* strains isolated from environmental water samples

Strain	Source	Serotype	Accession no.	No. of contigs (>200 bp)	G+C content (%)	Genome size (bp)	No. of protein-coding genes
Bnt314	Pond	IV	BBUG00000000	80	38.2	3,471,799	3,105
Ofk308	Ashiyu foot spa	IV	BBUH00000000	86	38.2	3,473,188	3,104
Twr292	Ashiyu foot spa	I	BBUI00000000	66	38.2	3,394,434	3,007
Ymg289	Water fountain	I	BBUJ00000000	54	38.3	3,689,833	3,291
Ymt294	Ashiyu foot spa	I	BBUK00000000	141	38.4	3,401,814	3,076

REFERENCES

1. **Diederen BM**. 2008. *Legionella* spp. and Legionnaires' disease. J Infect **56:**1–12. http://dx.doi.org/10.1016/j.jinf.2007.09.010.
2. **Berk SG, Faulkner G, Garduño E, Joy MC, Ortiz-Jimenez MA, Garduño RA**. 2008. Packaging of live *Legionella pneumophila* into pellets expelled by *Tetrahymena* spp. does not require bacterial replication and depends on a Dot/Icm-mediated survival mechanism. Appl Environ Microbiol **74:** 2187–2199. http://dx.doi.org/10.1128/AEM.01214-07.
3. **Berk SG, Ting RS, Turner GW, Ashburn RJ**. 1998. Production of respirable vesicles containing live *Legionella pneumophila* cells by two *Acanthamoeba* spp. Appl Environ Microbiol **64:**279–286.
4. **Hoffmann C, Harrison CF, Hilbi H**. 2014. The natural alternative: protozoa as cellular models for *Legionella* infection. Cell Microbiol **16:**15–26. http://dx.doi.org/10.1111/cmi.12235.
5. **Tachibana M, Nakamoto M, Kimura Y, Shimizu T, Watarai M**. 2013. Characterization of *Legionella pneumophila* isolated from environmental water and Ashiyu foot spa. BioMed Res Int **2013:**514395. http://dx.doi.org/10.1155/2013/514395.
6. **Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Prjibelski AD, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanaus-kas R, Clingenpeel SR, Woyke T, McLean JS, Lasken R, Tesler G, Alekseyev MA, Pevzner PA**. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol **20:** 714–737. http://dx.doi.org/10.1089/cmb.2013.0084.
7. **Chien M, Morozova I, Shi S, Sheng H, Chen J, Gomez SM, Asamani G, Hill K, Nuara J, Feder M, Rineer J, Greenberg JJ, Steshenko V, Park SH, Zhao B, Teplitskaya E, Edwards JR, Pampou S, Georghiou A, Chou IC, Iannuccilli W, Ulz ME, Kim DH, Geringer-Sameth A, Goldsberry C, Morozov P, Fischer SG, Segal G, Qu X, Rzhetsky A, Zhang P, Cayanis E, De Jong PJ, Ju J, Kalachikov S, Shuman HA, Russo JJ**. 2004. The genomic sequence of the accidental pathogen *Legionella pneumophila*. Science **305:**1966–1968. http://dx.doi.org/10.1126/science.1099776.
8. **Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT**. 2009. Reordering contigs of draft genomes using the Mauve aligner. Bioinformatics **25:**2071–2073. http://dx.doi.org/10.1093/bioinformatics/btp356.
9. **Seemann T**. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics **30:**2068–2069. http://dx.doi.org/10.1093/bioinformatics/btu153.
10. **Arakawa K, Mori K, Ikeda K, Matsuzaki T, Kobayashi Y, Tomita M**. 2003. G-language genome analysis environment: a workbench for nucleotide sequence data mining. Bioinformatics **19:**305–306. http://dx.doi.org/10.1093/bioinformatics/19.2.305.