

Genome Sequence of an Environmental Isolate of the Bacterial Pathogen *Legionella pneumophila*

Jian Ma,^a Yongqun He,^{b,c} Bijie Hu,^a Zhao-Qing Luo^d

Zhongshan Hospital, Fudan University, Shanghai, China^a; Unit for Laboratory Animal Medicine, Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, Michigan, USA^b; Center for Computational Medicine and Bioinformatics, University of Michigan Medical School, Ann Arbor, Michigan, USA^c; Department of Biological Sciences, Purdue University, West Lafayette, Indiana, USA^d

We report here the genomic sequence of *Legionella pneumophila* strain LPE509 from the water distribution system of a hospital in Shanghai, China. This is the first complete genome sequence of an environmental *L. pneumophila* isolate. Genomic analyses identified approximately 600 genes unique to LPE509 compared to those of the 7 available *L. pneumophila* genomes.

Received 22 April 2013 Accepted 26 April 2013 Published 6 June 2013

Citation Ma J, He Y, Hu B, Luo Z-Q. 2013. Genome sequence of an environmental isolate of the bacterial pathogen *Legionella pneumophila*. *Genome Announc.* 1(3):e00320-13. doi:10.1128/genomeA.00320-13.

Copyright © 2013 Ma et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Bijie Hu, hujie@vip.sina.com, or Zhao-Qing Luo, luoz@purdue.edu.

Species of the Gram-negative genus *Legionella* are the causative agent of Legionnaires' disease. These bacteria are parasites of fresh amoebae in aquatic environments, which are considered the primary source of infection in Legionnaires' disease outbreaks. Currently, 7 *Legionella pneumophila* genomes are available, all from clinical isolates, i.e., the six serogroup 1 strains 130b (1), Alcoy (2), Corby (3), Lens (4), Paris (4), and Philadelphia 1 (5) and the serogroup 12 strain 570-CO-H (6). These genomic data have provided the information necessary for large-scale comparative genomic analysis, which is instrumental to the study of the evolution of virulence of this important pathogen (7, 8).

To gain insights into the potential differences between clinical *L. pneumophila* isolates and environmental strains, we determined the entire genome sequence of a strain (*L. pneumophila* LPE509) that was isolated from the water distribution system of a hospital in Shanghai, China (9). Genomic DNA was used to prepare 500-bp shotgun libraries to obtain a total of 1,944,445 (94-fold coverage) reads in each library by an Illumina DNA analyzer, which was *de novo* assembled into 36 contigs >500 bp with SOAPdenovo 1.05. Full assembly was achieved with Geneious by sequencing PCR products that spanned the gaps.

We used the software Glimmer 3 (10) to predict the open reading frames, which were then annotated by BLAST searches against GenBank. tRNA genes and RNA genes were identified with the tRNAscan and RNAmmer programs, respectively, followed by BLAST searches. Results from BLAST searches against known *L. pneumophila* genomes were used to assign gene names. Genome comparison between LPE509 and the other known *L. pneumophila* genomes (strains 130b, 570-CO-H, Alcoy, Corby, Lens, Paris, and Philadelphia 1) was performed with Vaxign (11) and incorporated the ortholog searching program OrthoMCL (12).

The genome of strain LPE509 consists of a single circular chromosome of 3,434,224 bp and a plasmid of 73,490 bp. The average G+C content is 38.32%. It contains 9 rRNA operons, 43 tRNA genes, and 4,285 putative open reading frames, of which 3,246 are predicted genes (75.8%). The genomic organization of LPE509 is

highly syntenic with those of the seven already-sequenced *L. pneumophila* strains. Gene clusters encoding all the structural components of the type I secretion system (T1SS), the T2SS, and the Dot/Icm type IVb secretion system (T4BSS) were found in syntenic positions. Components of the type IVb secretion system Lvh were detected in the genome. Interestingly, only five of the proteins LvhA, LvhB11, LvhD4, LvhB4, LvrB, LvrE, LvrC, and LvhB10 are highly similar (e value = $1.0E - 10$ to $4.0E - 73$) to those found in strain Philadelphia 1. LPE509 harbors 274 of the 277 experimentally verified protein substrates in the Dot/Icm transporter (7, 13). The plasmid unique to strain LPE509 carries genes that are putatively involved in resistance to multiple drugs. Compared to the other 7 sequenced strains (130b, 570-CO-H, Alcoy, Corby, Lens, Paris, and Philadelphia 1), LPE509 genes code for 627, 519, 563, 570, 720, 605, and 496 unique proteins, respectively, highlighting the high levels of genomic plasticity among different *L. pneumophila* isolates.

Nucleotide sequence accession numbers. The NCBI accession numbers of strain LPE509 are [CP003885](https://www.ncbi.nlm.nih.gov/nuccore/CP003885) and [CP003886](https://www.ncbi.nlm.nih.gov/nuccore/CP003886).

ACKNOWLEDGMENTS

We thank Lili Tao for the bacterial strain LPE509.

This work was supported by NIAID grant K02AI085403 (to Z.-Q.L.), by the Special Fund for Health-Scientific Research in the Public Interest research and application for the prevention and control of nosocomial infections caused by multidrug-resistant bacteria (201002021) (to B.H.), and by NIAID grant 1R01AI081062 (to Y.H.).

REFERENCES

- Schroeder GN, Petty NK, Mousnier A, Harding CR, Vogrin AJ, Wee B, Fry NK, Harrison TG, Newton HJ, Thomson NR, Beatson SA, Dougan G, Hartland EL, Frankel G. 2010. *Legionella pneumophila* strain 130b possesses a unique combination of type IV secretion systems and novel Dot/Icm secretion system effector proteins. *J. Bacteriol.* 192:6001–6016.
- D'Auria G, Jiménez-Hernández N, Peris-Bondía F, Moya A, Latorre A. 2010. *Legionella pneumophila* pangenome reveals strain-specific virulence factors. *BMC Genomics* 11:181.
- Glöckner G, Albert-Weissenberger C, Weinmann E, Jacobi S, Schunder

- E, Steinert M, Hacker J, Heuner K. 2008. Identification and characterization of a new conjugation/type IVA secretion system (trb/tra) of *Legionella pneumophila* Corby localized on two mobile genomic islands. *Int. J. Med. Microbiol.* 298:411–428.
4. Cazalet C, Rusniok C, Brüggemann H, Zidane N, Magnier A, Ma L, Tichit M, Jarraud S, Bouchier C, Vandenesch F, Kunst F, Etienne J, Glaser P, Buchrieser C. 2004. Evidence in the *Legionella pneumophila* genome for exploitation of host cell functions and high genome plasticity. *Nat. Genet.* 36:1165–1173.
 5. 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.
 6. Amaro F, Gilbert JA, Owens S, Trimble W, Shuman HA. 2012. Whole-genome sequence of the human pathogen *Legionella pneumophila* serogroup 12 strain 570-CO-H. *J. Bacteriol.* 194:1613–1614.
 7. Gomez-Valero L, Rusniok C, Cazalet C, Buchrieser C. 2011. Comparative and functional genomics of *Legionella* identified eukaryotic-like proteins as key players in host-pathogen interactions. *Front. Microbiol.* 2:208.
 8. Xu L, Luo ZQ. 2013. Cell biology of infection by *Legionella pneumophila*. *Microbes Infect.* 15:157–167.
 9. Tao LL, Hu BJ, Yu LL, Zhou ZY, Xie HM, Zhou CM. 2011. Genetic diversity of *Legionella pneumophila* serogroup 1 in hospital water distribution systems in Shanghai. *Zhonghua Jie He He Hu Xi Za Zhi* 34: 100–103.
 10. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 23:673–679.
 11. He Y, Xiang Z, Mobley HL. 2010. Vaxign: the first Web-based vaccine design program for reverse vaccinology and applications for vaccine development. *J. Biomed. Biotechnol.* 2010:297505. doi:10.1155/2010/297505.
 12. Li L, Stoeckert CJ, Jr, Roos DS. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* 13:2178–2189.
 13. Zhu W, Banga S, Tan Y, Zheng C, Stephenson R, Gately J, Luo ZQ. 2011. Comprehensive identification of protein substrates of the Dot/Icm type IV transporter of *Legionella pneumophila*. *PLoS One* 6:e17638.