AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

Check for updates

| Bacteriology | Announcement

Draft genome sequence of *Legionella pneumophila* strains isolated from engineered water systems in Brazil

Dândrea Driely de Melo Ferrari,[1] Edson Machado,[1] Emilyn Costa Conceição,[2,3,4] Abhinav Sharma,[3] Solange Lima,[5] Karla Valéria Batista Lima,[6] Raquel Lima de Figueiredo Teixeira,[1] Harrison Magdinier Gomes,[1] Philip Suffys[1]

AUTHOR AFFILIATIONS See affiliation list on p. 3.

ABSTRACT We report the draft genome sequence of 12 isolates of *Legionella pneumophila* subsp. *pneumophila* from Brazil. The genomes have a GC content of approximately 38%, genome sizes ranging from 3.20 to 4.09 Mb, and 2,845–3,615 predicted coding sequences.

KEYWORDS *Legionella pneumophila*

Legionella pneumophila is a ubiquitous environmental bacterium belonging to the Legionellaceae family. It is a small, gram-negative, strictly aerobic bacillus that does not produce spores or possess a capsule (1). Human infection occurs through the inhalation of aerosolized particles containing the pathogen, typically associated with engineered water systems. However, person-to-person transmission of Legionnaires' disease has also been reported (2). The clinical spectrum of disease comprises two well-defined syndromes: Pontiac fever, an acute, self-limiting influenza-like illness, and Legionnaires' disease, a severe pneumonia associated with significant morbidity and mortality (3). We report here the draft genome sequence of 12 strains of *L. pneumophila*, isolated as part of a surveillance study in Brazil, from nine hotels, a laboratory, and a shopping mall in eight Brazilian states (Table 1).

The bacteria were isolated in accordance with ISO 11731:2017 by Conforlab and sent to our laboratory, where a fresh culture prepared by inoculation in a plate with BCYE medium (Sigma-Aldrich, São Paulo, Brazil) at 37°C and 5% of CO_2 up to 14 days. A loop of wet bacterial mass was then submitted to nucleic acid extraction in parallel with adding a similar amount to BYE broth with glycerol for long-term storage in liquid nitrogen.

Genomic DNA was extracted using the method described by Jahan et al. (4), including cell lysis using proteinase K followed by DNA purification using cetyltrimethylammonium bromide, isoamyl-chloroform, and isopropanol precipitation. We determined DNA quantity and quality using the Qubit Double-Strand DNA Broad Range Assay Kit (Thermo Fisher Scientific, Waltham, USA). For whole-genome sequencing, a library was prepared using the Nextera XT DNA Library Preparation Kit (Illumina, San Diego, USA) and after quality control using the 2100 Bioanalyzer (Agilent Technologies, Santa Clara, United States) with libraries requiring a minimum of 30 ng of total DNA, sequenced on a NextSeq 500 platform (Illumina), using the NextSeq 500/550 High Output 300 Cycles Kit, generating 2 × 150 paired-end fragments.

Quality control and trimming of the reads was done using Trimmomatic v0.39 (5), with the following parameters: LEADING:3, TRAILING:3, SLIDINGWINDOW:4:20, and MINLEN:40. Read quality was verified using FastQC v0.12.1 (6), and genomes were *de novo* assembled using SPAdes v3.15.4 (7). We checked the quality of the genomes with CheckM v1.2.2 (8), and the annotation was performed by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v6.10 (9). Default parameters were used for all software unless otherwise specified. The characteristics of the assembled genomes are

Editor Kenneth M. Stedman, Portland State University, Portland, Oregon, United States

Address correspondence to Dândrea Driely de Melo Ferrari, dan.mferrari@hotmail.com, or Philip Suffys, psuffys@gmail.com.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 13 January 2026
Accepted 16 March 2026
Published 20 April 2026

Copyright © 2026 Ferrari et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

TABLE 1 Information on sequenced *Legionella* sp. isolates[b]

Sample	Year	Source	Local	City	State	GenBank accession no.	SRA	Organism	Contigs	Bases	Genes	% GC	N50	L50	Completeness	Contamination
ISHP06	2015	Drinking fountain	Hotel	Belo Horizonte	MG	JBPZJB000000000	SRR34759054	*L. pneumophila* subsp. *pneumophila*	64	3,551,892	3,214	38	116,016	10	100	0.19
ISHP07	2015	Faucet	Hotel	Porto Alegre	RS	JAJCCN000000000	SRR16377044	*L. pneumophila* subsp. *pneumophila*	52	3,350,592	3,038	38	145,133	8	100	0.38
ISHP03	2015	Faucet	Hotel	Vera Cruz	BA	JAJCCM000000000	SRR16377045	*L. pneumophila* subsp. *pneumophila*	106	3,219,274	2,897	38	64,772	15	100	0.38
18HP16	2018	Faucet	Hotel	Belo Horizonte	MG	JBPZJC000000000	SRR34759053	*L. pneumophila* subsp. *pneumophila*	58	3,554,197	3,212	38	137,947	9	100	0.19
18HP14	2018	Shower	Hotel	Rio de Janeiro	RJ	JBRKJP000000000	SRR16377043	*L. pneumophila* subsp. *pneumophila*	68	3,299,477	2,974	38	132,763	12	100	0.19
18SI20	2018	Cooling tower	Mall	São Paulo	SP	JBPZJD000000000	SRR34759052	*L. pneumophila* subsp. *pneumophila*	107	3,783,046	3,459	38	108,431	10	100	0.38
21121	2019	Faucet	Hotel[a]	Paranaguá	PR	JBPZJE000000000	SRR34759051	*L. pneumophila* subsp. *pneumophila*	30	3,276,353	2,962	38	346,681	4	100	0.19
21123	2019	Faucet	Hotel[a]	Paranaguá	PR	JBPZJF000000000	SRR34759050	*L. pneumophila* subsp. *pneumophila*	31	3,256,285	2,940	38	225,298	4	100	0.19
36803	2019	Faucet	Hotel	Aracaju	SE	JBPZJG000000000	SRR34759049	*L. pneumophila* subsp. *pneumophila*	61	3,442,241	3,111	38	133,048	9	100	0.19
37065	2019	Faucet	Hotel	Recife	PE	JBPZJI000000000	SRR34759047	*L. pneumophila* subsp. *pneumophila*	55	3,482,255	3,145	38	190,718	9	100	0.19
20935	2019	Shower	Hotel	Foz do Iguaçu	PR	JAJCCL000000000	SRR16377047	*L. pneumophila* subsp. *pneumophila*	35	3,361,576	3,021	38	174,743	6	100	0.19

[a]Samples belong to the same hotel. CheckM estimates genome completeness and contamination based on the presence or absence of marker genes.
[b]BA, Bahia; % GC, GC content; L50, the smallest number of contigs accounting for 50% of the total assembly length; MG, Minas Gerais; N50, the length of the shortest contig within this subset; PE, Pernambuco; PR, Paraná; RJ, Rio de Janeiro; RS, Rio Grande do Suk; SE, Sergipe; SP, São Paulo.

summarized in Table 1. Briefly, genome size ranged from 3.20 to 4.09 Mb; GC content was 38%; the number of contigs ranged from 31 to 107; and annotation predicted between 2,845 and 3,615 coding sequences.

To our knowledge, these are the first genomes from *L. pneumophila* isolated in Brazil. These data will allow better analysis of genetic variability of environmental *Legionella* strains in the country, which was previously only available through sequence typing (10), and will be the basis of more detailed genome analysis.

ACKNOWLEDGMENTS

We are grateful to Progenix Lab (Ribeirão Preto, SP, Brazil) and Evandro Chagas Institute (Ananindeua, PA, Brazil) for carrying out the genomic sequencing of the isolates.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior– Brasil–Finance Code 001.

AUTHOR AFFILIATIONS

[1]Laboratory of Molecular Biology Applied to Mycobacteria, Oswaldo Cruz Institute, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil
[2]Program in Clinical Research and Infectious Diseases, Evandro Chagas National Institute of Infectious Diseases, Oswaldo Cruz Foundation, Rio de Janeiro, Manguinhos, Brazil
[3]South African Medical Research Council Centre for Tuberculosis Research, Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa
[4]Centre for Epidemic Response and Innovation (CERI), School of Data Science and Computational Thinking, Stellenbosch University, Cape Town, Western Cape, South Africa
[5]Conforlab Engenharia Ambiental, São Paulo, Brazil
[6]Bacteriology and Mycology Section, Evandro Chagas Institute, Ananindeua, Pará, Brazil

AUTHOR ORCIDs

Dândrea Driely de Melo Ferrari http://orcid.org/0000-0001-5846-3647
Edson Machado http://orcid.org/0000-0002-7585-0005
Emilyn Costa Conceição http://orcid.org/0000-0002-7445-6620
Harrison Magdinier Gomes http://orcid.org/0000-0002-6489-8003
Philip Suffys http://orcid.org/0000-0001-8527-9300

FUNDING

Funder	Grant(s)	Author(s)
Coordenação de Aperfeiçoamento de Pessoal de Nível Superior		Dândrea Ferrari
Conselho Nacional de Desenvolvimento Científico e Tecnológico		Philip Suffys
Instituto Oswaldo Cruz (IOC)		Harrison Magdinier Gomes

AUTHOR CONTRIBUTIONS

Dândrea Driely de Melo Ferrari, Conceptualization, Data curation, Formal analysis, Methodology, Project administration, Writing – original draft | Edson Machado, Investigation, Methodology, Software, Writing – review and editing | Emilyn Costa Conceição, Investigation, Methodology, Writing – review and editing | Abhinav Sharma, Investigation, Methodology | Solange Lima, Data curation, Methodology | Raquel Lima de Figueiredo Teixeira, Resources, Supervision | Harrison Magdinier Gomes, Funding acquisition, Project administration, Supervision, Writing – review and editing | Philip

Suffys, Funding acquisition, Project administration, Supervision, Writing – review and editing.

DATA AVAILABILITY

The draft genome sequence of *L. pneumophila* strains has been deposited at GenBank. The raw reads are available in the Sequence Read Archive under the studies SRP604359 and SRP341755, associated with BioProject numbers PRJNA1297308 and PRJNA772000, respectively.

REFERENCES

1. Iliadi V, Staykova J, Iliadis S, Konstantinidou I, Sivykh P, Romanidou G, Vardikov DF, Cassimos D, Konstantinidis TG. 2022. *Legionella pneumophila*: the journey from the environment to the blood. J Clin Med 11:6126. https://doi.org/10.3390/jcm11206126
2. Borges V, Nunes A, Sampaio DA, Vieira L, Machado J, Simões MJ, Gonçalves P, Gomes JP. 2016. *Legionella pneumophila* strain associated with the first evidence of person-to-person transmission of Legionnaires' disease: a unique mosaic genetic backbone. Sci Rep 6:26261. https://doi.org/10.1038/srep26261
3. Mondino S, Schmidt S, Rolando M, Escoll P, Gomez-Valero L, Buchrieser C. 2020. Legionnaires' disease: state of the art knowledge of pathogenesis mechanisms of *Legionella*. Annu Rev Pathol 15:439–466. https://doi.org/10.1146/annurev-pathmechdis-012419-032742
4. Jahan SJ, Chowdhury SF, Mitu SA, Shahriar MS, Bhuiyan MAB. 2015. Genomic DNA extraction methods: a comparative case study with Gram-negative organisms. bjbabe VI:61–68. https://doi.org/10.7904/2068-4738-VI(11)-61
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. Andrews S. 2015. FastQC: A quality control tool for high throughput sequence data. Babraham Inst. http://www.bioinformatics.babraham.ac.uk/projects.
7. 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
8. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114
9. Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretdin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F. 2021. RefSeq: expanding the prokaryotic genome annotation pipeline reach with protein family model curation. Nucleic Acids Res 49:D1020–D1028. https://doi.org/10.1093/nar/gkaa1105
10. Ferrari D de M, Lima SC, Teixeira RLF, Lopes MQP, Vaconcellos SEG, Machado ES Jr, Suffys PN, Gomes HM. 2024. Genetic diversity of Legionella pneumophila isolates from artificial water sources in Brazil. Curr Microbiol 81:165. https://doi.org/10.1007/s00284-024-03645-5