INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY

TAXONOMIC DESCRIPTION

Girolamini *et al.*, *Int. J. Syst. Evol. Microbiol.* 2022;72:005512

DOI 10.1099/ijsem.0.005512

MICROBIOLOGY SOCIETY

OPEN ACCESS

Legionella bononiensis sp. nov., isolated from a hotel water distribution system in northern Italy

Luna Girolamini[1,2]†, Maria Rosaria Pascale[1]†, Silvano Salaris[1], Marta Mazzotta[1], Massimiliano Orsini[3], Antonella Grottola[2,4], Nicoletta Zini[5,6] and Sandra Cristino[1,2,*]

Abstract

Legionella-like isolates, strains 27fs60, 30fs61 and 30cs62^T, were isolated from a hotel water distribution system in the Emilia-Romagna region, Italy. Isolates were Gram- and Ziehl Neelsen-stain-negative, rod-shaped, with transitory flagella presence and able to grow at 32–37 °C (with an optimum at 32 °C) on buffered charcoal–yeast extract agar with L-cysteine, glycine–vancomycin–polymyxin B–cycloheximide agar and Wadowsky–Yee medium agar. The strains showed positive reactions for oxidase, hippurate and gelatinase and a weakly positive reaction for catalase. Based on the EUCAST cut-off, strain 30cs62^T was resistant to ciprofloxacin (5 mg l^{-1}). The *mip* and *rpoB* gene sequences of the three strains showed close matches to those of *Legionella quateirensis* ATCC 49507^T with similarity values of 98.2 and 94.5%, respectively. Whole genome sequencing of the three strains was performed, resulting in G+C contents of 39.0, 39.1 and 39.0 mol%, respectively. The identity percentage measured by average nucleotide identity between the three strains and their respective closest strains were: 91.32% *L. quateirensis* NCTC 12376^T, 91.45% *L. quateirensis* ATCC 49507^T and 91.45% *L. quateirensis* ATCC 49507^T, respectively. The digital DNA–DNA hybridization analysis demonstrated how the isolates were separated from the most related phylogenetic *Legionella* species (*L. quateirensis* ATCC 49507^T, ≤40.10% DNA–DNA relatedness). The concatenated phylogenetic tree based on 16S *rRNA*, *mip*, *rpoB* and *rnpB* genes, shows a close relationship with *L. quateirensis* ATCC 49507^T. The results obtained confirm the status of an independent species. The name proposed for this species is *Legionella bononiensis* sp. nov. with 30cs62^T (=ATCC TSD-262^T=DSM 112526^T) as the type strain.

ISOLATION AND ECOLOGY

Legionella is the only genus of the family *Legionellaceae*. *Legionella* are Gram-negative aerobic bacilli that live in fresh and man-made water environments. To date, a total of 63 species of *Legionella* have been characterized, 24 of these are associated with human infections [1, 2] and involved in Legionnaire's disease and Pontiac fever [3, 4]. The infection is acquired via inhalation of contaminated aerosol or, less commonly, by aspiration of drinking water [5]. Only two cases of person-to-person infection have been reported [6, 7].

The study presents the taxonomy characterization of three strains (27fs60, 30fs61, 30cs62^T) belonging to the genus *Legionella* as a novel species, under the proposed name of *Legionella bononiensis* sp. nov.

Author affiliations: [1]Department of Biological, Geological and Environmental Sciences, University of Bologna, via San Giacomo 12, 40126 Bologna (BO), Italy; [2]European Society of Clinical Microbiology and Infectious Diseases (ESCMID) Study Group for *Legionella* Infections (ESGLI), Basel, Switzerland; [3]Laboratory of Microbial Ecology and Genomics of Microorganisms, Istituto Zooprofilattico Sperimentale delle Venezie, Legnaro, Italy; [4]Regional Reference Laboratory for Clinical Diagnosis of Legionellosis, Unit of Molecular Virology and Microbiology, Modena University Hospital, Modena, Italy; [5]CNR-National Research Council of Italy, Institute of Molecular Genetics "Luigi Luca Cavalli-Sforza", Unit of Bologna, via di Barbiano 1/10, 40136, Bologna, Italy; [6]IRCCS Istituto Ortopedico Rizzoli, via di Barbiano 1/10, Bologna, Italy.

***Correspondence:** Sandra Cristino sandra.cristino@unibo.it

Keywords: aquatic environment; *Legionella bononiensis* sp. nov.; new species; polyphasic taxonomy; whole genome sequencing (WGS).

Abbreviations: BCYE, buffered charcoal yeast extract agar; Cys+, with L-cysteine; Cys-, without L-cysteine; dDDH, digital DNA-DNA hybridization; GLM, generalized linear model; GVPC, glycine, vancomycin, polymyxin B, cycloheximide agar medium; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; MWY, Wadowsky Yee medium.

The GenBank accession numbers for the 16S rRNA genes of strains 27fs60, 30fs61 and 30cs62^T are OK495707, OK495709 and OK495708, respectively. The GenBank accession numbers for the whole genome sequences of strains 27fs60, 30fs61 and 30cs62^T are JADOBG000000000, JADWVM000000000 and JADWVN000000000, respectively.

† These authors contributed equally to this work.

Eight supplementary figures and four supplementary tables are available with the online version of this article.

Table 1. *rpoB* and *mip* mismatches and amino acid variations of the three strains compared to *Legionella quateirensis* reference strain (ATCC 49507^T)

Gene size (bp)	Percentage of nucleotide identity (%)	No. of mismatches in DNA (DNA)	No. of mismatches in amino acid (AA)
Entire *rpoB* gene (4107)	94.8	214	11
Partial *rpoB* (329)	94.5	18	0
Partial *mip* (611)	98.2	11	2

Water samples were collected from a hotel in Emilia-Romagna region, Italy, during July 2019. The hot and cold water samples (2l) were collected in post-flushing modality following the UNI EN ISO 19458:2006 [8]. *Legionella* isolation was performed by a standard culture technique on glycine–vancomycin–polymyxin B–cycloheximide (GVPC) agar medium (Thermo Fisher Scientific Diagnostic), according to ISO 11731:2017 [9]. The plates were incubated at 35±2 °C with 2.5% CO_2 for 15 days.

Several samples collect from hot and cold-water distribution system displayed colonies growth on GVPC medium. The presumptive *Legionella* colonies were sub-cultured on buffered charcoal–yeast extract (BCYE) agar with L-cysteine (Cys+) and without L-cysteine (Cys-), Wadowsky–Yee medium (MWY), tryptone soya agar with 5% sheep blood agar medium (Thermo Fisher Scientific Diagnostic) and chocolate-enriched agar medium (MEUS S.r.L.) (Fig. S1). *Legionella*-like colony growth was observed only on BCYE *Cys*+, GVPC and MWY. These colonies (at least five different colonies for each plate) were tested using a *Legionella* latex test kit (Thermo Fisher Scientific Diagnostic), based on manufacturer instructions, with negative results. The colonies were also stored at −80±2 °C in glycerol [9].

GENOME FEATURES

Three colonies randomly chosen and designated as strain 27fs60, 30fs61 and 30cs62^T were characterized by macrophage infectivity potentiator (*mip*) and RNA polymerase beta subunit (*rpoB*) gene sequencing [10, 11]. The DNA extraction was performed using InstaGene Purification Matrix (Bio-Rad), and DNA concentrations were determined using a Qubit fluorometer (Thermo Fisher Scientific). PCR was carried out to amplify the *mip* and *rpoB* genes according to Ratcliff *et al.* and Ko *et al.* [10, 11], the products were visualized by electrophoresis on 2% agarose gel stained with ethidium bromide, and sequenced by Sanger technology, returning reliable sequences for the *mip* (611 nt) and *rpoB* (329 nt) genes.

Based on the *mip* gene sequence, the isolates were closely related to the *L. quateirensis* strain ATCC 49507^T, sharing 98.2% of sequence identity. The result obtained with the partial *rpoB* gene sequence, showed an identity of 94.5% (Table 1). The *mip* and *rpoB* gene sequences of the three isolates were identical to each other.

The whole genome sequencing (WGS) of the three strains (27fs60, 30fs61 and 30cs62^T) was performed as described in our genome announcement study [12]. Briefly, 100 ng of genomic DNA were used for next-generation sequencing libraries preparation, using the Nextera XT DNA library prep kit (Illumina). Sequencing was performed on an Illumina NextSeq 500 platform (2×150 paired-end reads). Raw reads were assembled using TORMES version 1.2.0 [13], an automated pipeline for whole bacterial genome analysis, using the default parameters. TORMES performed a sequence quality filtering (PRINSEQ version 0.20.4) and a *de novo* genome assembly (SPAdes version 13.4.1) [14]. The TORMES workflow detected the following virulence genes: *enhA*, *flgG*, *htpB*, *icmD*/*dotP*, *icmJ*/*dotN*, *icmW*, *IspG* and *mip*; in contrast, other genes involved in several pathways of *Legionella* infection were missing. In Table S1, the list of detected and undetected virulence genes, with the percentage of sequence identify respect to the most virulent *L. pneumophila philadelphia* Serogroup 1 ATCC 33152^T and the phylogenetically closest *L. quateirensis* ATCC 49507^T strains is reported. The role of these genes, as well as the pathogenic role of the strains, will be further investigated using amoebae and macrophage cell line models.

The generated contigs were passed to CSAR [15], a scaffolding tool able to order and orient the contigs of the given draft genome based on one or more reference genomes of a related organism. A set of three genomes of different evolutionary related organisms were given as a reference: *L. pneumophila* OLDA CP016030.2, *L. fallonii* NZ_LN614827.1, *L. quateirensis* LNYR01000001.1 (number of contigs: 50) and the best result was retained (*L. fallonii* NZ_LN614827.1). A further refinement was carried out remapping the reads on the CSAR scaffolds using the Geneious Prime 2021.1.1 software (www.geneious.com) [16]. BUSCO version 5.0.0. was useded to evaluate the completeness of the three genome assemblies. The analysis indicates that the three genomes are near-complete, with values between 93.5 and 99.2% (Table 2). The final draft genomes were submitted to the GenBank requiring the annotation by the PGAP pipeline [17]. Results regarding assembling and annotation are summarized in Table 3. Briefly, the genomes total lengths were 4211919, 3709497 and 4136543 bp for 27fs60, 30fs61 and 30cs62^T, respectively, with a G+C content of approximately 39 mol%.

Table 2. BUSCO quality analysis

BUSCO results (%) (no. of genes)	27fs60	30fs61	30cs62^T
Complete	99.2 (123)	93.5 (116)	95.2 (118)
Single-copy complete	99.2 (123)	93.5 (116)	95.2 (118)
Duplicated complete	0.0 (0)	0.0 (0)	0.0 (0)
Fragmented	0.8 (1)	0.8 (1)	0.8 (1)
Missing	0 (0)	5.7 (7)	4.0 (5)
Total no. of BUSCO genes	124	124	124

The discrepancy of 500 Kbp discovered between the three strains, calculated between the longer (27fs60) and shorter (30fs61) strain, could be attributed to technical and analytical factors, occurred during the analysis. Anyway, the alignment of three strains genomes, provided by BLAST Ring Image Generator (version 0.95) software using 27fs60 as reference genome, displayed that in the other two strains (30fs61 and 30cs62) some regions were missed (Fig. S2).

The analysis of discrepancy in the genome size was performed by the Prokka annotation pipeline (version 1.14.6) and Python (version 3.8.1) software (Tables S2 and S3). Considering that the results presented refer to draft genomes, a future step will be to cover the entire genome using PacBio's third-generation sequencing technique.

The OrthoANIu package [18] was used to measure the intra and inter species genome similarities by average nucleotide identity (ANI) among the assembled draft genomes. Comparison returned the following values: 99.88% between 27fs60 and 30fs61, 99.92% between 27fs60 and 30cs62^T, 99.92% between 30fs61 and 30cs62^T, confirming that the three strains belong to the same species and are identical to each other (Table 2). Further, FastANI [19] through DFAST [20] was performed against 13000 prokaryotic reference genomes from the National Center for Biotechnology Information (NCBI), to assess the taxonomic identity. The identity percentage of the closest strain for 27fs60, 30fs61 and 30cs62^T were: 91.32% *L. quateirensis* NCTC 12376^T, 91.45% *L. quateirensis* ATCC 49507^T and 91.45% *L. quateirensis* ATCC 49507^T, respectively. ANI was also measured based on BLAST+ (ANIb) and MUMmer (ANIm) using JSpeciesWS [21]. The analysis confirmed the FastANI results with the following values: 27fs60 ANIb 90.62% compared with *L. quateirensis* NCTC 12376^T, ANIm 91.86% with *L. quateirensis* ATCC 49507^T; 30fs61 ANIb 90.64% with *L. quateirensis* ATCC 49507^T, ANIm 91.88% with *L. quateirensis* ATCC 49507^T; 30cs62^T ANIb 90,59% with *L. quateirensis* ATCC 49507^T, ANIm 91.87% with *L. quateirensis* ATCC 49507^T.

The relatedness of our strains to known type strains was further analysed via digital DNA–DNA hybridization (dDDH). The method was implemented via the Genome-to-Genome Distance Calculator 2.1 (GGDC) web service (http://ggdc.dsmz.de), retaining default parameters, using BLAST+ [22] as a local alignment tool. The GGDC uses a genome BLAST distance phylogeny to infer genome-to-genome distances between pairs of entirely or partially sequenced genomes. Our strains were compared with the assembly of *L. quateirensis* strain ATCC 49507^T (AC LNYR01000001.1), the closest related strain based on the previous ANI outcomes. The analysis returned the following results: for 27fs60 a DDH estimate [generalized linear model (GLM)-based] of

Table 3. Genome statistics data from the NCBI

Attribute	27fs60	30fs61	30cs62^T
No. of raw reads	4184062	3851726	3626424
Average read length (bp)	149	149	149
Coverage (×)	142	131	124
Total Length (bp)	4211919	3709497	4136543
No. of contigs	23	37	32
G+C content (mol%)	39.00	39.10	39.00
N50 (bp)	312097	166809	176017
No. of coding sequences	3542	3155	3491
No. of rRNA genes	3	3	3
No. of tRNA genes	41	37	39

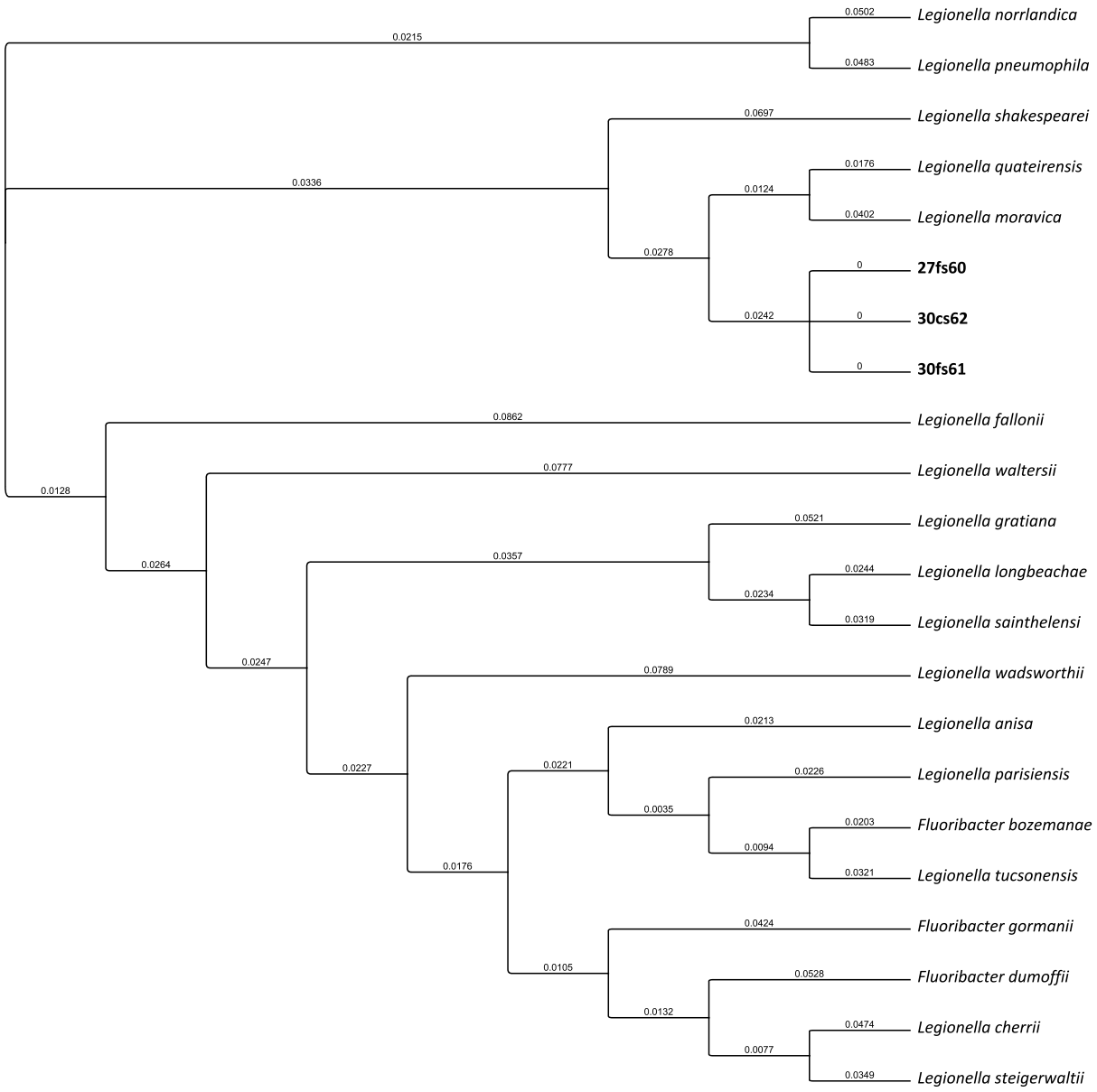

Fig. 1. Phylogenetic tree of the three strains (27fs60, 30fs61 and 30cs62^T) and closely related *Legionella* species based on concatenation of four genes (*16S rRNA*, *mip*, *rpoB*, *rnpB*). Branch lengths are transformed to be equal like a cladogram. Branch labels display the substitutions per site.

44.10% [probability that DDH >70% (i.e. same species) of 6.81% (via logistic regression)]; for 30fs61 a DDH estimate (GLM-based) of 44.00% [probability that DDH>70% (i.e., same species) of 6.79% (via logistic regression)]; for 30cs62^T a DDH estimate (GLM-based) of 44.10% (probability that DDH >70% (i.e., same species) of 6.93% (via logistic regression)). The genomic DNA of the two respective organisms revealed an *in silico* DDH similarity of below 70%, regarding them as distinct species [23]. Considering the thresholds established for ANI (95%) and dDDH (70%) analysis, our results led us to consider these strains as new species [23, 24].

Moreover, based on WGS results, the phylogenetic analyses were performed both with single and concatenated gene (*16S rRNA*, *mip*, *rpoB*, *rnpB*) sequences. The tree was built using implemented software in Geneious Prime genome browser (Geneious Prime 2021.1.1; www.geneious.com) [16]. The sequences were aligned using the MUSCLE algorithm [25]. The subsequent multiple sequence alignment was passed to FastTree [26], using the Jukes-Cantor (default parameters) as a genetic distance model. The concatenated tree (Fig. 1) showed that strains 27fs60, 30fs61 and 30cs62^T were strictly related to each other, and belonged to a monophyletic group comprising *L. quateirensis*, *L. moravica*, *L. shakespeari* and *L. fallonii* strains. Particularly, they formed

Fig. 2. Strain 30cs62^T from a water sample on glycine–vancomycin–polymyxin B–cycloheximide (GVPC) agar medium.

a sister group with the *L. quateirensis* ATCC 49507^T. The branching structure of the individual gene trees also supported the differentiation of our three strains from *L. quateirensis*, and the relationship with the other species included in the major clade described above (Figs S3–S6).

PHYSIOLOGY AND CHEMOTAXONOMY

For the physiological and chemotaxonomic characteristics, the 30cs62^T was taken as a model.

The subculture was performed on BCYE C*ys* + agar plates and the growth at different temperatures was evaluated: 32, 35 and 37 °C with 2.5% of CO_2.

The strain was analysed with Gram- and Ziehl Neelsen-staining and autofluorescence was evaluated under a Woods lamp (long-wave length UV light at 365 nm). Morphological features and the presence of flagellum were investigated by using a Zeiss EM109 transmission electron microscope (TEM). Images were captured using a Nikon Dmx 1200F digital camera and ACT-1 software (Nikon Corporation). Cells were negatively stained with 2% uranyl acetate.

Biochemical parameters for enzyme activities were investigated. The oxidase and catalase tests were performed by Oxidase test strips (Biolife) and a catalase colorimetric activity kit (Thermo Fisher Scientific Diagnostic), respectively. The hippurate and gelatinase reactions were evaluated using Diatabs kit and nutrient gelatin medium, respectively (Biolife). Moreover, the strain reaction was evaluated using BBL Crystal Enteric/Non-Fermenter ID kit (Becton Dickinson Systems) and a Remel RapID NF Plus system (Thermo Fisher Scientific Diagnostic), following manufacturer's procedure.

The detection of β-lactamase production was evaluated using Oxoid nitrocefin solution (Thermo Fisher Scientific Diagnostic).

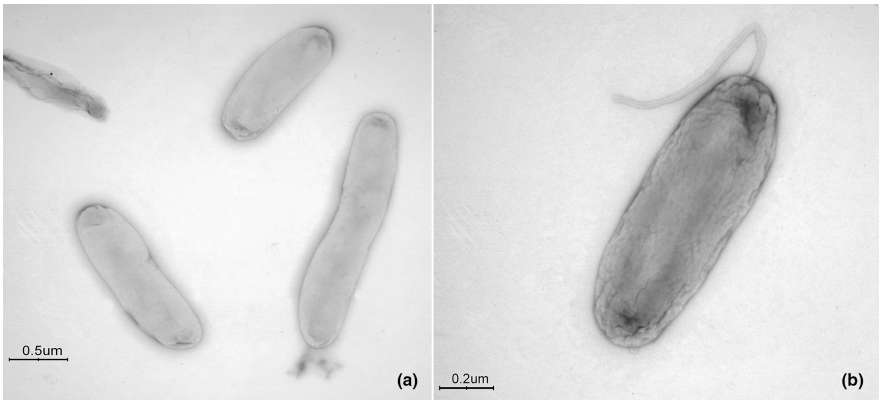

Fig. 3. Transmission electron microscopy (TEM) images of strain 30cs62^T grown on BCYE Cys+ agar for 2 days at 35 °C with 2.5% CO_2. (a) Aflagellate 30cs62^T form and (b) flagellate 30cs62^T form. Bar, (a) 0.5 µm and (b) 0.2 µm. Magnification: (a) ×12000 and (b) ×20000.

Table 4. Biochemical characteristics of strain 30cs62^T

+, Positive; –, negative; w, weakly positive.

Substrates	Result	Substrates	Result
ARA (arabinose)	–	GLU (glucose)	–
MNS (mannose)	–	PRO (proline-β-naphthylamide)	w
SUC (sucrose)	–	PYR (pyrrolidine-β-naphthylamide)	–
MEL (melibiose)	–	GGT (g-glutamyl β-naphthylamide)	+
RHA (rhamnose)	–	TRY (tryptophane β-naphthylamide)	–
SOR (sorbitol)	–	IND (tryptophane)	+
MNT (mannitol)	–	NO_3 (sodium nitrate)	+
ADO (adonitol)	–	GLR (p-nitrophenly β-glucuronide)	–
GAL (galactose)	–	NAG (ρ-nitrophenyl-*N*-acetyl-β,D-glucosaminide)	–
PHO (p-nitrophenyl phosphate)	+	GGL (γ-L-glutamyl *p*-nitroanilide)	–
BGL (p-nitrophenyl α-β-glucoside)	–	ESC (aesculin)	–
NPG (p-nitrophenyl β-galactoside)	–	PHE (p-nitro-DL-phenylalanine)	–
BPH (p-nitrophenyl bis-phosphate)	–	URE (urea)	–
BXY (p-nitrophenyl xyloside)	–	CIT (citrate)	–
AAR (p-nitrophenyl α-arabinoside)	–	MLO (malonate)	–
PHC (p-nitrophenyl phosphorylcholine)	–	TTC (tetrazolium)	–
ADH/ARG (arginine)	–	LYS (lysine)	–
TRD (aliphatic thiol)	+	GLY (glycine)	+
PHS (ρ-nitrophenyl-phosphoester)	+	BANA (*N*-benzyl-arginine-β-naphthylamide)	–
αGLU (ρ-nitrophenyl-α,D-glucoside)	–	EST (triglyceride)	+
βGLU (ρ-nitrophenyl-β,D-glucoside)	–	INO (inositol)	–
ONPG (ρ-nitrophenyl, β,D-galactoside)	–		

Antimicrobial susceptibility testing was carried out by MIC Test Strip (Liofilchem s.r.l.), on isolates grown on BCYE Cys + agar at 35 °C in a humidified atmosphere, according to manufacturer's instructions. The antibiotics tested were: azithromycin, erythromycin, ciprofloxacin, rifampicin, tigecycline and imipenem. EUCAST (European Committee on Antimicrobial Susceptibility Testing) guidance documents, provided by European Society of Clinical Microbiology and Infectious Diseases (ESCMID) were used for the interpretation of minimal inhibition concentration (MIC) data [27].

Moreover, the three strains were tested by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) MS technique with MALDI Biotyper system (Bruker Daltonik GmbH) following manufacturer's instruction. The acquired spectra were analysed using Bruker database (version BDAL 7854). Although the isolates were classified as belonging to the genus *Legionella* according to ISO 11731:2017, the MALDI-TOF MS technique returned a non-identification, neither at genus level.

The isolation on GVPC agar from water samples showed road shape colonies, white frosted with concentric circles, with an approximate diameter of 6 mm (Fig. 2). The sub-culture of the strain 30cs62^T on BCYE Cys + displayed a growth after 2 days of incubation at 35±2 °C with 2.5% CO_2. The colonies were convex, grey and road shape, with an approximate diameter of 1 mm (Fig. S7). Moreover, the best growth was observed at 32 °C, respect to 35 and 37 °C, without differences between the presence of CO_2, in respect to microaerophilia.

Strain 30cs62^T was Gram- and Ziehl Neelsen-stain-negative. Light microscopy observation showed the motility of strain 30cs62^T. The strains observed under long-wave length UV light (365 nm) did not show a fluorescence (Fig. S8). The cells of strain 30cs62^T in TEM images were 0.4±0.1 µm wide, 1.7±0.4 µm long and not all the cells had flagella (Fig. 3). The transitory flagella presence could be

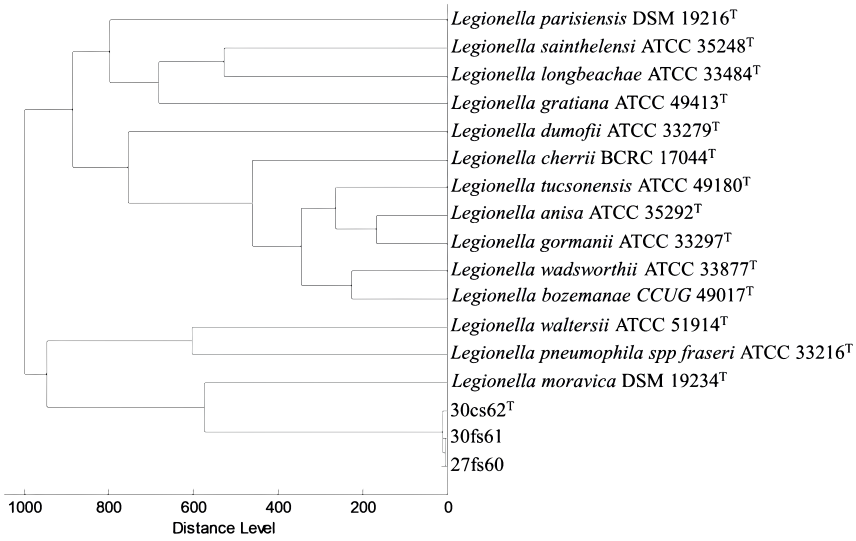

Fig. 4. Dendrogram based on whole-cell MALDI-TOF mass spectra of strains 27fs60, 30fs61 and 30cs62^T and *Legionella* reference strains present in the instrument database.

explained by the bacteria growth phase that is strictly linked to several environmental factors, such as temperature, medium viscosity and nutrient availability (e.g. amino acids and fatty acids) according to Appelt and Heuner [28].

Strain 30cs62^T was positive for oxidase, weakly positive for catalase and a positive reaction for hippurate and gelatinase tests was observed. Moreover, a beta-lactamase production was observed. The results of the biochemical parameters tested are shown in Table 4.

The results for antibiotic susceptibility are shown in Table S4. Based on the epidemiological cut-off established for *L. pneumophila* by EUCAST [27], strain 30cs62^T does not exhibit resistance.

The dendrogram developed on MALDI-TOF MS results (Fig. 4), demonstrated a clear separation between the three strains with respect to the closely related *Legionella* reference strains available in the instrument database.

DESCRIPTION OF *LEGIONELLA BONONIENSIS* SP. NOV.

Legionella bononiensis (bo.no.ni.en'sis. L. fem. adj. *bononiensis*, pertaining to the city of Bologna, which was called 'Bononia' in the Roman period).

Cells are Gram-negative, Ziehl Neelsen-negative and rod-shaped when grown on BCYE Cys+, GVPC or MWY agars. Able to grow at temperatures between 32–37 °C (optimal, 32 °C). Positive for oxidase, hippurate and gelatinase, weakly positive for catalase, with beta-lactamase production. The BBL crystal and Remel RapID tests show positive reactions for γ-glutamyl β-naphthylamide, tryptophane, sodium nitrate, p-nitrophenyl phosphate, aliphatic thiol, glycine, ρ-nitrophenyl-phosphoester and triglyceride. Weakly positive for proline-β-naphthylamide.

This species was isolated from hot and cold hotel water distribution system, in the Emilia-Romagna region, Italy. The type strain, 30cs62^T (=ATCC TSD-262^T=DSM 112526^T), has a G+C content of 39.0 mol%.

Funding information
This work received no specific grant from any funding agency.

Acknowledgements
The authors would like to thank Dr. Graziella Ciani, Dr. Francesco Bisognin and Dr. Miriam Cordovana for their support and suggestions.

Authors and contributors
Conceptualization: L.G., M.R.P. and S.C. Data curation: L.G., S.S. and M.R.P. Formal analysis: L.G., M.R.P., M.M., N.Z. and A.G. Funding acquisition: S.C. Investigation: L.G. Methodology: L.G., M.R.P. and S.C. Project administration: S.C. Resources: S.C. Software: S.S., M.O. and L.G. Supervision: S.C. Writing – original draft: L.G. Writing – review and editing: L.G. and S.C.

Conflicts of interest
The authors declared that there are no conflicts of interest.

References

1. Newton HJ, Ang DKY, van Driel IR, Hartland EL. Molecular pathogenesis of infections caused by *Legionella pneumophila*. *Clin Microbiol Rev* 2010;23:274–298.
2. Parte AC, Road W. LPSN–List of Prokaryotic Names with Standing in Nomenclature. *Nucleic Acids Res* 2014;42:D613–6.
3. Fields BS, Benson RF, Besser RE. *Legionella* and Legionnaires' disease: 25 years of investigation. *Clin Microbiol Rev* 2002;15:506–526.
4. Rota MC, Caporali MG, Bella A, Scaturro M, Giannitelli S, *et al*. The legionellosis surveillance system in Italy: 2019 results. *Boll Epidemiol Naz* 2020;1:32–38.
5. Centers for Disease Control and Prevention. Legionella (Legionnaires' Disease and Pontiac Fever) – Causes, How it Spreads, and People at Increased Risk; 2020. www.cdc.gov/legionella/about/causes-transmission.html
6. Correia AM, Ferreira JS, Borges V, Nunes A, Gomes B, *et al*. 'Probable person-to-person transmission of Legionnaires' disease. *N Engl J Med* 2016;374:497–498.
7. Borges V, Nunes A, Sampaio DA, Vieira L, Machado J, *et al*. *Legionella pneumophila* strain associated with the first evidence of person-to-person transmission of Legionnaires' disease: a unique mosaic genetic backbone. *Sci Rep* 2016;6:1–11.
8. ISO. UNI EN ISO 19458:2006 – Water quality — Sampling for microbiological analysis. Geneva, Switzerland: ISO; 2006. www.iso.org/standard/33845.html
9. ISO. ISO 11731:2017 Water quality – Enumeration of Legionella. Geneva, Switzerland: ISO; 2017. www.iso.org/standard/61782.html
10. Ratcliff RM, Lanser JA, Manning PA, Heuzenroeder MW. Sequence-based classification scheme for the genus *Legionella* targeting the *mip* gene. *J Clin Microbiol* 1998;36:1560–1567.
11. Ko KS, Lee HK, Park M-Y, Lee K-H, Yun Y-J, *et al*. Application of RNA polymerase beta-subunit gene (*rpoB*) sequences for the molecular differentiation of *Legionella* species. *J Clin Microbiol* 2002;40:2653–2658.
12. Girolamini L, Salaris S, Orsini M, Pascale MR, Mazzotta M, *et al*. Draft genome sequences of *Legionella* presumptive novel species isolated during environmental surveillance in artificial water systems. *Microbiol Resour Announc* 2021;10:e00307-21.
13. Quijada NM, Rodríguez-Lázaro D, Eiros JM, Hernández M. TORMES: an automated pipeline for whole bacterial genome analysis. *Bioinformatics* 2019;35:4207–4212.
14. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, *et al*. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19:455–477.
15. Chen K-T, Liu C-L, Huang S-H, Shen H-T, Shieh Y-K, *et al*. CSAR: a contig scaffolding tool using algebraic rearrangements. *Bioinformatics* 2018;34:109–111.
16. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, *et al*. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 2012;28:1647–1649.
17. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, *et al*. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 2016;44:6614–6624.
18. Yoon S-H, Ha S-M, Lim J, Kwon S, Chun J. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek* 2017;110:1281–1286.
19. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018;9:1–8.
20. Tanizawa Y, Fujisawa T, Nakamura Y. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* 2018;34:1037–1039.
21. Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* 2016;32:929–931.
22. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, *et al*. BLAST+: architecture and applications. *BMC Bioinformatics* 2009;10:1–9.
23. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14:60.
24. Kim M, Oh HS, Park SC, Chun J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* 2014;64:346–351.
25. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;32:1792–1797.
26. Price MN, Dehal PS, Arkin AP. FastTree 2–approximately maximum-likelihood trees for large alignments. *PLoS One* 2010;5:e9490.
27. EUCAST. Antimicrobial Susceptibility Testing of *Legionella pneumophila*. Vaxjo, Sweden; 2021. www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/General_documents/Legionella_guidance_document_20160909.pdf
28. Appelt S, Heuner K. The flagellar regulon of Legionella—a review. *Front Cell Infect Microbiol* 2017;7:1–13.