

Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN
SCIENZE DELLA TERRA, DELLA VITA E DELL'AMBIENTE

Ciclo 35°

Settore Concorsuale: 06/M1 - IGIENE GENERALE E APPLICATA, SCIENZE
INFERMIERISTICHE E STATISTICA MEDICA

Settore Scientifico Disciplinare: MED/42 – IGIENE GENERALE E APPLICATA

**Development of a Whole-Genome Sequencing (WGS)
approach for the characterization of *Legionella*
environmental isolates in water systems**

Presentata da: Luna Girolamini

Coordinatore Dottorato

Prof.ssa Maria Giovanna Belcastro

Supervisore

Prof.ssa Sandra Cristino

Co-supervisore

Prof. Marco Passamonti

Esame finale anno 2023

Abstract

Legionella is a Gram-negative bacterium that represent a public health issue, with heavy social and economic impact. Therefore, it is mandatory to provide a proper environmental surveillance and risk assessment plan to perform *Legionella* control in water distribution systems in hospital and community buildings.

The thesis joins several methodologies in a unique workflow applied for the identification of non-*pneumophila* *Legionella* species (n-pL), starting from standard methods as culture and gene sequencing (*mip* and *rpoB*), and passing through innovative approaches as MALDI-TOF MS technique and whole genome sequencing (WGS). The results obtained, were compared to identify the *Legionella* isolates, and lead to four presumptive novel *Legionella* species identification. One of these four new isolates was characterized and recognized at taxonomy level with the name of *Legionella bononiensis* (the 64th *Legionella* species).

The workflow applied in this thesis, help to increase the knowledge of *Legionella* environmental species, improving the description of the environment itself and the events that promote the growth of *Legionella* in their ecological niche. The correct identification and characterization of the isolates permit to prevent their spread in man-made environment and contain the occurrence of cases, clusters, or outbreaks. Therefore, the experimental work undertaken, could support the preventive measures during environmental and clinical surveillance, improving the study of species often underestimated or still unknown.

Extended Abstract

Aims: This dissertation joins several methodologies in a unique workflow, including routine analysis and research approach, in order to enhance the knowledge of the less studied species of environmental *Legionella* spp. and, in addition, underline the role of a proper environmental surveillance and risk assessment plan in the *Legionella* control in water systems.

Study Area: Emilia-Romagna Region.

Methods: 1) In Chapters 1 and 2 the focus is on the comparison between the standard methods and a faster and innovative technique (MALDI-TOF MS) for the identification of *Legionella*, other than the development of a new classification scheme regarding the molecular identification of *Legionella*. 2) Chapters 3 to 5 described the draft genomes of three presumptive novel species of *Legionella*, based on whole genome sequencing (WGS) approach 3) Chapter 6 shows the taxonomical characterization of the 64th species of *Legionella* through biochemical, serological, morphological, and molecular techniques. 4) In Chapter 7 the infectivity and antibiotic sensitivity test of the three novel *Legionella* species was described.

Results: 1) Chapter 1 suggests that MALDI-TOF MS technique is able to identify non-*pneumophila* *Legionella* species (n-pL). With results in line with *mip* gene sequencing, considered the “gold standard”. Moreover, through *mip*-gene sequencing two strains as “presumptive” novel *Legionella* species were discovered. Chapter 2 shows that *rpoB*-gene sequencing presents more discriminatory power with respect to *mip*-gene sequencing for n-pL providing a new classification scheme for their identification and suggesting its introduction for the *Legionella* identification scheme. In addition, the *rpoB* sequencing results confirmed the presence in the data set of two strains belonging to “presumptive” novel species, already discovered by *mip* gene sequencing. Moreover, thanks to *rpoB* sequencing other two strains showed differences in gene sequences, not previously detected by *mip* gene, leading to a total of four presumptive novel *Legionella* species.

2) Chapter 3, 4 and 5 confirmed the presumptive novel species as such.

3) Chapter 6 shows the taxonomical characterization of the 64th species of *Legionella* named *Legionella bononiensis*.

4) In Chapter 7, the three novel species were evaluated regarding their infectivity on *Amoebas* and macrophage cells, moreover, the susceptibility to main antibiotics used to treat legionellosis were tested.

Conclusions: The knowledge of the environment and the events that promote the changes in the ecological niche where *Legionella* live, can help to prevent their spread in man-made environment and contain the occurrence of cases, clusters, or outbreak. For this reason, an optimal environmental surveillance and a proper risk assessment plan are necessary to understand the *Legionella* dynamic of contamination during the time. Therefore, the correct identification and characterization of the isolates are essential. In this context, the workflow describes in this dissertation could help to better characterize the little recognize and studied non-*pneumophila* *Legionella* species, especially in term of their infectivity and antibiotic susceptibility.

Keywords: genotyping, *Legionella bononiensis* sp. nov., *Legionella* identification, non-*pneumophila* *Legionella* species, MALDI-TOF MS, *mip* and *rpoB* gene sequencing, polyphasic taxonomy, water distribution system, whole genome sequencing (WGS).

Preface and outline

The research presented in this dissertation is the product of a three-year PhD project at the Biological Geological and Environmental Science (BiGeA) Department of the Alma Mater Studiorum, University of Bologna, based at the research group of Environmental Microbiology and Molecular Biology Laboratory (MabLab), supervised by Prof. Sandra Cristino.

The general aim of the thesis has been performing the characterization of *Legionella* environmental species through new approaches as MALDI-TOF MS and whole-genome sequencing (WGS) developed for the isolates come from water distribution systems.

The dissertation is structured as papers collection (see “List of Publications”), and it follows the following outline. The thesis is organized in a general overview of the relevant scientific literature, the aim of the research, an overview of all material and methods applied during the PhD project, seven chapters and a general conclusion. The chapters consist of six scientific papers published and one that will be submitted in relevant peer-reviewed international open access journals.

The thesis is divided in four parts:

- **Part 1:** It includes two chapters: Chapters 1 and 2. Chapter 1 focused on the comparison between the routine standard methods and a faster and innovative technique for the identification of *Legionella*. The routine analysis was based on culture technique, the serological identification by agglutination test and the *mip*-gene sequencing. Whereas the innovative technique was based on MALDI-TOF MS approach. This Chapter showed that MALDI-TOF MS was able to produce a fast identification of *Legionella* with a good rate of concordance respect to the routinely techniques. In addition, Chapter 2 provides the comparison between the gold standard for the identification of non-*pneumophila* *Legionella* species (*mip*-gene sequencing) and another gene sequencing (*rpoB*). The results obtained showed a more discriminatory power of *rpoB*-gene sequencing with respect to the gold standard, other than the development of a new classification scheme regarding the molecular identification of *Legionella*.
- **Part 2:** The second part comprises 3 chapters (Chapters 3 to 5) in which the drafts genome of three presumptive novel species of *Legionella* were described. Chapter 3 focused on 27fs60, 30fs61 and 30cs62^T *Legionella* strains and their molecular characteristics provided by WGS technique. Chapter 4 described the molecular characteristics of 31fl33^T strain provided by WGS technique. Whereas Chapter 5 shows the workflow introduced to discover

a novel species of *Legionella* (8cVS16^T and 9fVS26 strains) with molecular characteristics of the novel strains.

- **Part 3:** The third part contains Chapter 6, where the taxonomic characterization of the 64th species of *Legionella*, called *Legionella bononiensis* (27fs60, 30fs61 and 30cs62^T strains) was described.
- **Part 4:** The fourth part includes the last chapter, Chapter 7. It will be submitted as “original research” in relevant peer-reviewed international open access journal. It contains the results obtained during my abroad period three months spend in France, at the French National Reference Center of *Legionella*, Hospices Civils of Lyon. Chapter 7 presents the new *Legionella* strains infectivity on Amoebas and macrophages models other than the results of susceptibility antibiotic tests performed on the three-novel species.

Considering that the articles are supposed to be an independent contribution, but on related topics, there is some similarity among the different articles. Following the last chapter, the conclusion presents the implication of my research on *Legionella* field, other than the potential future implications. In the final part of the dissertation, presented as Appendix A, the supplementary materials related to Chapter 6 (figures and tables) were displayed. The references cited in the “General Background” and “Materials and Methods” sections are listed below the two sections. Instead, each chapter has the independent reference section.

During these three-years of my PhD, I had the opportunity to improve my knowledge and skills. Results from my PhD project were presented at the 10th International Conference on *Legionella*. Moreover, I acted as co-supervisor on two master thesis (Anna Kamila Jedrusiak, Alex Conti) and one bachelor thesis (Federica Marino), as well as assisting with didactic activity for the course of Hygiene, teaching some lesson of epidemiology, other than I acted as tutor for teaching support activities in the Academic Years 2020/2021 and 2021/2022 (Bachelor’s degree, Scienze Biologiche (L-13), University of Bologna, years 2019-2022).

Furthermore, during these three years I published sixteen scientific papers, eight of them as first author. Four of these papers are strictly related to the main topic of my PhD project.

Luna Girolamini

Bologna, Italy, January 2023

List of publications

The results of this thesis are presented in the following papers based on the purpose and objectives of the doctoral project:

- Paper 1:** Evaluation of MALDI–TOF Mass Spectrometry in Diagnostic and Environmental Surveillance of *Legionella* Species: A Comparison With Culture and *Mip*-Gene Sequencing Technique
- Paper 2:** New Insight regarding *Legionella* Non-*Pneumophila* Species Identification: Comparison between the Traditional *mip* Gene Classification Scheme and a Newly Proposed Scheme Targeting the *rpoB* Gene
- Paper 3:** Draft Genome Sequences of *Legionella* Presumptive Novel Species Isolated during Environmental Surveillance in Artificial Water Systems
- Paper 4:** Draft Genome Sequence of *Legionella* Species Isolated from Drinking Water in an Italian Industry
- Paper 5:** Combining Traditional and Molecular Techniques Supports the Discovery of a Novel *Legionella* Species During Environmental Surveillance in a Healthcare Facility
- Paper 6:** *Legionella bononiensis* sp. nov., isolated from a hotel water distribution system in northern Italy

Contents

Abstract	2
Extended Abstract.....	3
Preface and outline.....	5
List of publications	7
Contents	8
List of abbreviations	10
1. General background.....	15
1.1 Historical background and discovery of the <i>Legionella</i>	15
1.2 <i>Legionella</i> taxonomy	15
1.3 Morphological and phenotypic characteristics.....	15
1.4 Ecology	16
1.4.1 Environmental reservoirs	17
1.4.2 <i>Legionella</i> natural and artificial life cycle	18
1.5 Pathogenicity and infection pathway	18
1.6 <i>Legionella</i> involvement in human disease.....	21
1.6.1 Legionnaires' diseases	21
1.6.2 Pontiac Fever.....	22
1.6.3 Risk factors	22
1.6.4 Research, diagnosis, and treatment.....	23
1.6.4.1 Human samples.....	23
1.6.4.2 Environmental samples	26
1.6.4.3 Identification of <i>Legionella</i> isolates	27
1.6.4.4 LD Therapy	28
1.6.5 Epidemiology.....	29
1.6.5.1 European epidemiological trend	29
1.6.5.2 Italian epidemiological data	32
1.7 Prevention and control of <i>Legionella</i> infections.....	33
1.7.1 Control measures	38
2. Scope of the study	41
3. Material and methods	42
3.1 Sample collection.....	42
3.2 Routinely <i>Legionella</i> microbiological analysis and serological identification	44
3.3 Research approach developed for the study.....	47
3.3.1 Identification by MALDI–TOF MS (Chapter 1).....	47

3.3.2	Identification of <i>Legionella</i> spp. by <i>mip</i> and <i>rpoB</i> genes sequencing (Chapters 1-6)	49
3.3.3	Phylogenetic and Allelic Diversity Analysis (Chapters 1-6)	51
3.3.4	Elaboration of matrices for the <i>rpoB</i> gene: definition of the intra- and interspecies intervals of variation (Chapter 2)	51
3.3.5	Identification of <i>Legionella</i> by Whole-Genome Sequencing (WGS) (Chapter 3-6)	52
3.3.6	Characterization of a novel <i>Legionella</i> species (Chapter 6)	53
References		54
Chapter 1		65
	<i>Evaluation of MALDI–TOF Mass Spectrometry in Diagnostic and Environmental Surveillance of Legionella Species: A Comparison With Culture and Mip-Gene Sequencing Technique</i>	65
Chapter 2		91
	<i>New Insight regarding Legionella Non-Pneumophila Species Identification: Comparison between the Traditional mip Gene Classification Scheme and a Newly Proposed Scheme Targeting the rpoB Gene</i>	91
Chapter 3		126
	<i>Draft Genome Sequences of Legionella Presumptive Novel Species Isolated during Environmental Surveillance in Artificial Water Systems</i>	126
Chapter 4		132
	<i>Draft Genome Sequence of Legionella Species Isolated from Drinking Water in an Italian Industry</i>	132
Chapter 5		138
	<i>Combining Traditional and Molecular Techniques Supports the Discovery of a Novel Legionella Species During Environmental Surveillance in a Healthcare Facility</i>	138
Chapter 6		161
	<i>Legionella bononiensis</i> sp. nov., isolated from a hotel water distribution system in northern Italy.	161
Chapter 7		180
	<i>Summary of preliminary results obtained during the three months abroad in French National Reference Center of Legionella, Hospices Civils of Lyon, France.</i>	180
4. Conclusions		190
	Suggestions for future work	190
Appendix A:		191
Supplementary material - Chapter 6		191
Acknowledgements		205

List of abbreviations

§	As explain in paragraph
A	Acid sample
ANI	Average nucleotide identity
APT	Amoebae plate test
ASL	Local health authority
AST	Antibiotic susceptibility test
ATCC	American Type Culture Collection
BCYE	Buffered Charcoal Yeast Extract
BiGeA	Biological Geological and Environmental Science
BLAST	Basic local alignment search tool
BMD	Broth microdilution
BUSCO	Benchmarking Universal Single-Copy Orthologs
CDC	Centres for Disease Control
cfu	Colonies forming units
CI	Confidence interval
CTs	Cooling towers
CWS	Cold-water samples
Cys	L-Cysteine
dDDH	DNA–DNA digital hybridization
D.Lgs.	Legislative decree
DFA	Direct fluorescent antibody

DMIPi	Department of Infectious Diseases
DMSO	Dimethylsulfoxide
dNTPs	Deoxynucleoside triphosphates
Dot/Icm	Defective organelle trafficking or intracellular multiplication
DSA	Direct immunofluorescence
DSM	Deutsche Sammlung von Mikroorganismen und Zellkulturen
DVR	Risk Assessment Document
ECOFF	Epidemiological cut-off
ELDSNet	European Legionnaires' Disease Surveillance Network
ER	Endoplasmic reticulum
ESCMID	European Society of Clinical Microbiology and Infectious Diseases
ESGLI	ESCMID Study Group for <i>Legionella</i> Infections
EU	European
EUCAST	European Committee on Antimicrobial Susceptibility Testing) guidance documents, provided by ESCMID
EWGLI	European Working Group for <i>Legionella</i> Infection
F	Filtrated sample
FBS	Fetal bovine serum
FTIR	Fourier-transform infrared spectroscopy
GBDP	Genome Blast Distance Phylogeny
GGDC	Genome-to-Genome Distance Calculator
GLM	Generalized linear model
GVPC	Glycine Vancomycin Polymyxin Cycloheximide

H	Heated sample
H₂O₂/Ag⁺	Hydrogen peroxide and silver salt
HCF	Healthcare facilities
HWS	Hot-water samples
IFA	Indirect immunofluorescent assay
IFI	Indirect immunofluorescence
IMVS	Institute of Medical and Veterinary Science
IPTG	Isopropyl β-d-1-thiogalactopyranoside
ISO	International Standard Organization
ISS	Italian National institute of Health
LCV	<i>Legionella</i> containing vacuole
LD	Legionellosis
Lp	<i>Legionella pneumophila</i>
Lp1	<i>Legionella pneumophila</i> serogroup 1
LPS	Lipopolysaccharide
MABs	Monoclonal antibodies
MabLab	Environmental Microbiology and Molecular Biology Laboratory
MALDI-TOF MS	Matrix-assisted laser desorption ionization-time-of-flight mass spectrometry
MIC	Minimal inhibition concentration
<i>mip</i>	Macrophage infectivity potentiator
MOI	Multiplicity of infection
MSA	Multiple sequence alignment
MUSCLE	Multiple sequence comparison by log-expectation

NBRC	NITE Biological Research Center, National Institute of Technology and Evaluation
NCBI	National Center for Biotechnology Information
NCTC	National Collection of Type Cultures, Central Public Health Laboratory
n-pL	Non- <i>pneumophila</i> <i>Legionella</i> species
OD₆₀₀	Optical density at 600nm
PGAP	Prokaryotic Genome Annotation Pipeline
PMA	Phorbol 12- myristate 13-acetate
PPE	Personal protective equipment
Ps	<i>Pseudomonas aeruginosa</i>
PTFE	Polytetrafluoroethylene
RPMI	Roswell Park Memorial Institute
rpoB	Beta subunit of RNA polymerase
SBT	Sequence-based typing
SE	Sensitivity
SGs	Serogroups
SP	Specificity
SST4	type IV secretion system
ST1	Sequence type 1
TEM	Transmission electron microscope
TLRs	Toll-Like Receptor
TSA	Tryptic soy agar
UAT	Urinary diagnostic test

UN	Untreated sample
UNI EN	Italian National Unification and European Committee
WDS	Water distribution system
WHO	World Health Organization
WGS	Whole-genome sequencing
WSP	Water safety plan
WT	Wild type

1. General background

1.1 Historical background and discovery of the *Legionella*

Legionella was discovered and identified in 1976. On June 1976, during the 58th annual convention of the American Legion, at the Bellevue Stratford Hotel of Philadelphia, PA, USA, a pneumoniae epidemic broke out. The American Legion was an organization of veterans of the U.S. military forces, called "Legionnaires". Among the 4400 participants, there were 182 cases of pneumonia and 29 of them died [1].

The first epidemiologic investigation did not be able to identify the cause of the pneumonia epidemic. Only on the December 1976, Doctor McDade from US Centres for Disease Control (CDC), was able to isolate the cause of the epidemic, named the bacteria discovered as *Legionella pneumophila* (*Lp*). The name of the bacteria derives from the name of the veterans, as well as the name of the disease: "Legionnaire's disease" [2]. The source of spread of the bacteria was identified in the hotel air condition system where the Legionnaires had stayed [3].

1.2 *Legionella* taxonomy

Legionella are the only genera of the Legionellaceae family. To date, the genera *Legionella* includes 64 species and about 70 serogroups (SGs) [4, 5]. Among them, not all the species are dangerous for human health and *Lp* species is the most studied species, and they are included 16 serogroups. *Lp* SG1 (*Lp*1) appears to be the most associated with human infections, and accounts for about 90% of them [6, 7].

1.3 Morphological and phenotypic characteristics

Legionella are a Gram-negative coccobacillus of the Legionellaceae family. They are asporogenous, strictly aerobically and mobile thanks to the presence of one or more flagella in polar or lateral position. The size shows a wide range from 0.3 to 0.9 μm and the length ranges are from 2 and 20 μm . Instead, in tissue and clinical samples, they have a coccobacillary form, measuring 1–2 μm [8]. *Legionella* metabolism consists of amino acids, such as cysteine, arginine, isoleucine, and methionine. Otherwise, it is unable to ferment sugars. In addition, the growth of the bacterium is favoured by the presence of iron salts [9]. For these reasons, their growth in culture condition is complex. The growth medium is based on L-Cysteine (Cys), iron, charcoal, yeast extract and alpha-ketoglutarate. On this medium, *Legionella* colonies are convex, roundish and with ground-glass

appearance, sharp margins, with a white to grey colour. Moreover, some species (i.e., *L. anisa*, *L. bozemanae*, *L. gormanii* and *L. dumoffii*) could be shown a blue-white autofluorescence, others (i.e. *L. rubrilucens*) show a red fluorescence if exposed under a Woods lamp (long wavelength UV light at 365 nm).

1.4 Ecology

Legionella are mesophilic bacteria, they can survive in the water between 5.7 to 63 °C, but the range of 25-40 °C represents the optimal growth temperature [10, 11]. Consequently, to decrease the risk of infections, it is highly recommendable to set the temperature of the artificial water distribution systems (WDS) above the 50 °C [12–14]. Although *Legionella* can survive some minutes/hours at temperature ≥ 50 °C (Figure 1), and some species can replicate at very low temperature, the monitoring of water temperature in the man-made WDS remains one of the most important procedure to undertake to prevent and control *Legionella* infection [15].

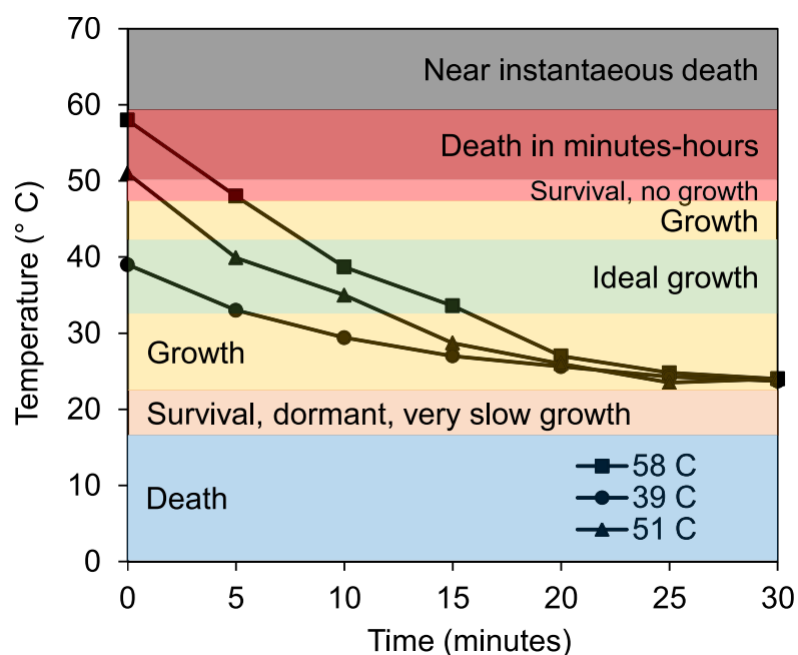


Figure 1 Effect of the water temperature on *Legionella* at distal taps with stagnation time [16].

Another important factor to *Legionella* survival is pH. *Legionella* lives in both acidic and alkaline environments, in a wide range of pH between 2.0 and 9 [11, 16, 17].

1.4.1 Environmental reservoirs

Legionella spp. are present in natural fresh water and in man-made WDS. In these environments, it is able to intracellularly replicate within a 15 species of amoeba (i.e., *Acanthamoeba*, *Hartmanella*, *Valkampfia*, *Naegleria*, etc.) and three species of ciliates (i.e., *Tetrahymena*, *Cyclidium*, and *Paramecium*) [18, 19]. In particular, *Legionella* exploits the ability of amoebae to produce cysts escaping the disinfection treatments and acquiring virulence factors to increase the infectivity power [20]. In the WDS, *Legionella* are commonly found within biofilms, where they can find nourishment and protection, that promoting its proliferation, adhesion to surfaces and spread in the environment [20].

The principal natural environments are rivers, thermal springs, lakes, etc., other than soil (i.e., *L. longbeachae*) [10, 21]. Among artificial environments, there are showers, whirlpools, faucets, baths, cooling towers (CTs), condensers, humidifiers, medication nebulizer devices, etc. [22–24]. These environments are considered the real sources of human risk of *Legionella* infection [6] (Figure 2).

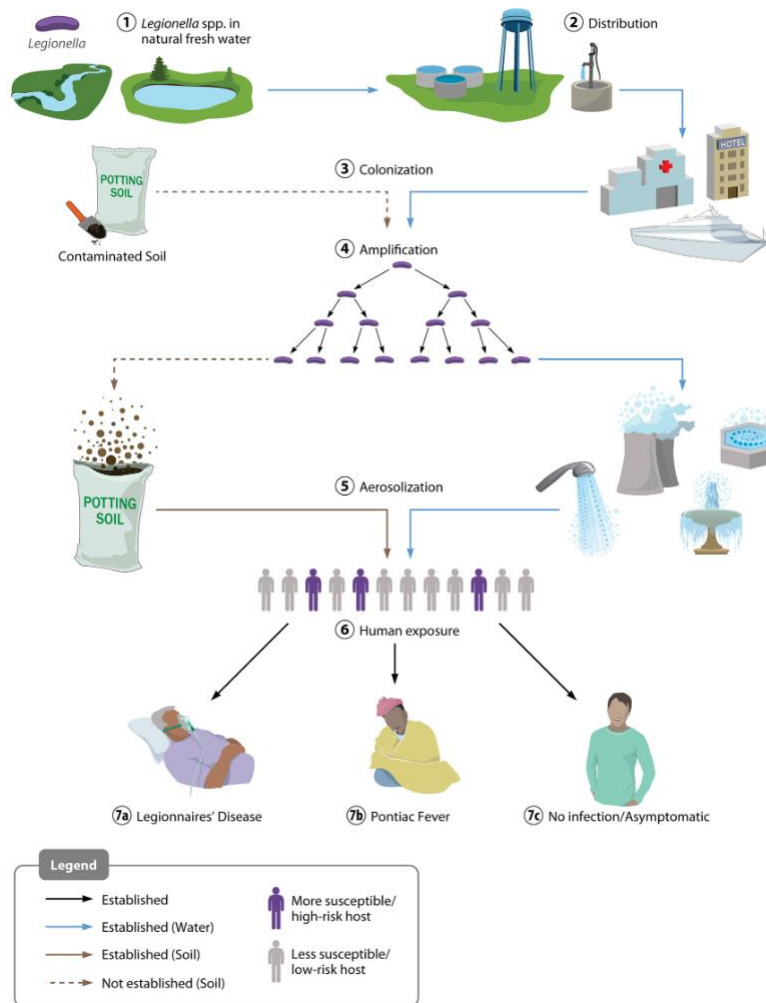


Figure 2 The way of *Legionella* dissemination from water for human and development of LDs [24].

1.4.2 *Legionella* natural and artificial life cycle

In natural condition, usually, *Legionella* live inside the biofilm communities, where they are preyed upon by amoebas [20]. Inside of the amoebas *Legionella* can resist to their digestion and can replicate. Killing their hosts, *Legionella* come back to the aquatic environment. The life cycle of *Legionella* is based on two different phases: replication and transmission (Figure 3).

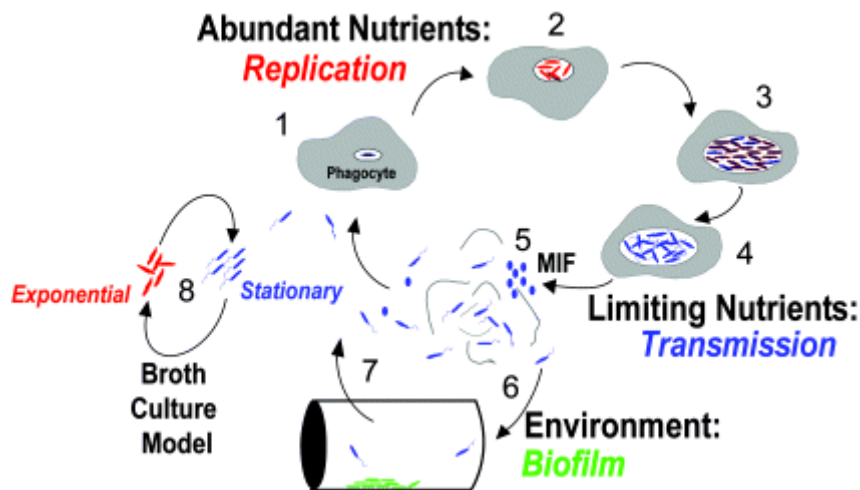


Figure 3 Life cycle of *L. pneumophila* [26].

In the replication form (inside the amoebas, ciliates and macrophages), *Legionella* are aflagellated and sodium resistant, show an elongated form, with suppression of some virulence genes [25]. When the presence of nutrient became inadequate, *Legionella* changing their form assuming a transmissible form. They repress the multiplication form promoting some genes to escape from the host cell and looking for a new niche showing a high virulence. In this phase *Legionella* change their conformation in a non-replicative form, flagellated and sensitive to sodium. *Legionella*, in this form, express the gene of motility, phagocyte necrosis, and resistance to cellular stress [25].

1.5 Pathogenicity and infection pathway

The human being is an accidental host of *Legionella*, which was infected by inhalation of the contaminated water droplets [19, 26, 27]. To date, only one case of interhuman transmission of *Legionella* has been reported [28, 29]. Through the contaminated aerosols inhalation, *Legionella* reach the lungs, as far as alveolar macrophages, which could be infected [30, 31]. The alveolar macrophages could phagocyte *Legionella*; after 2-4h this phagocyte mutates and show some ribosomes on its surface [32, 33]. Inside the ribosome-lined vacuoles, *Legionella* is able to replicate [34]. The vacuoles harbouring *Legionella* (*Legionella* containing vacuole (LCV)) and the macrophage

endoplasmic reticulum (ER), indicate that *Legionella* have come up with a way to enter in the ER lumen, altering the transport of the phagosome. Once inside, *Legionella* are into an organelle rich in peptides, their primary carbon source, and are protected against the lysosome degradation [35]. *Legionella* can cause an acute and severe, sometimes fatal, pneumonia known as Legionnaires' disease or Legionellosis (LD) [35]. In Europe, the 83% of human diseases are caused by *Lp1* [36]. Instead, less than 10% of human infections are caused by non-*pneumophila* *Legionella* species (*n-pL*), especially *L. longbeachae*, *L. micdadei*, *L. bozeman*, *L. dumoffii*, and *L. feeleii*, which are repeatedly isolated from hospitalized patients, whereas *L. anisa*, *L. wadsworthii*, and *L. cincinnatiensis* are only rarely found [37, 38]. The reported differences could be attributed to the specificity and sensitivity of the antigenic urinary diagnostic test (UAT) for *Lp1* [6, 39]. *Legionella* presents various virulence factors and genes (secretion systems, macrophage infectivity potentiator (*mip*), lipopolysaccharide (LPS), flagellin, etc.) that permit the interactions with the host cells, in order to develop the infection. The crucial steps of the cellular infection cycle of *Lp*, are the attachment and invasion of host cells. Several numbers of factors are implicated in these processes, including pili [40, 41]. Both in the natural and in human beings, the intracellular growth is promoted by Dot/Icm (Defective organelle trafficking or intracellular multiplication) type IV secretion system (SST4) and the mechanism of infection is similar in amoebas, ciliates and macrophages [42, 43]. The SST4 is useful to deliver some macromolecules across kingdom barriers, and it is the first contact between *Legionella*, host-cells and phagosome biogenesis [44, 45]. Thanks to the Dot/Icm effectors, *Legionella* can subvert the host functions to invade the hosts, replicate inside and escape from them [46, 47] (Figure 4).

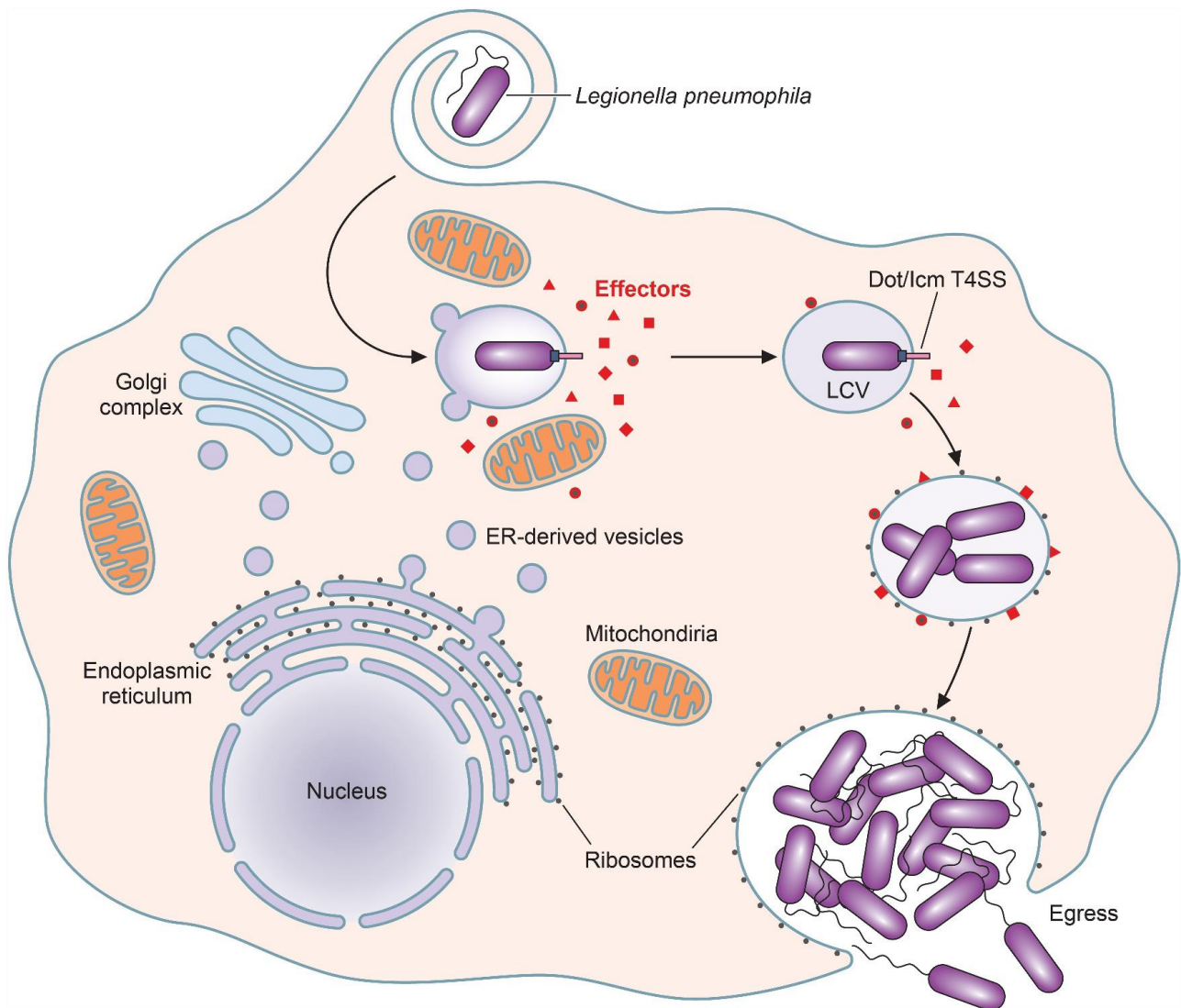


Figure 4 Life cycle of *Lp* within eukaryotic host cells [48]

The repertoire of these effectors is very conserved among different *Lp*, the same thing is not observed between other *Legionella* species because in most of the effectors lack orthologs [49–51]. Except to the effector diversity, at the interspecies level, the Dot/Icm secretion system is highly conserved [30, 51]. For this reason, a simple deletion of one gene that coding the SST4 effectors does not alter the functionality of the intracellular multiplication [52]. Instead, LPS, that found in the majority of the Gram negative bacteria, plays a fundamental role in the recognising of the *Legionella* by the immune system [53, 54]. *Legionella*, during the virulent form, show flagella. The flagella have a key role in the mobility of the Gram-negative bacteria and support the contact with the hosts. This happens thanks to the flagellin, a flagellar protein, recognized by the Toll-Like Receptor (TLRs) present in the alveolar macrophages surface. In addition, play a fundamental role in the inhibition of phagosome maturation and in the secretion of the pro-inflammatory cytokine in the alveolar

macrophages [55, 56]. The other important virulence gene is *mip*. It codifies to a membrane protein that is implicated in the entry process of the *Legionella* in the host cells [57].

1.6 *Legionella* involvement in human disease

Legionella plays several roles in human health and are responsible for a variety of clinical conditions known as Legionellosis. This term includes three different syndromes: Legionnaires' disease, Pontiac Fever and Pittsburgh's diseases other than an asymptomatic condition (Figure 4).

1.6.1 Legionnaires' diseases

The Legionnaires' disease is the most dangerous and potentially fatal interstitial pneumonia caused by *Legionella*. Usually, the incubation time is between 2-10 days, but can be also more longer over 15 days [9, 58–60]. Most of all are immunocompromised and hospitalized patients [61, 62]. The initial symptoms are high fever, weakness, follow by malaise and fatigue, anorexia, and lethargy. Continuing by mild and non-productive cough [62, 63]. During the progression of the disease, could appears blood-streaked sputum and haemoptysis. Sometimes in patients a pleuritic chest pain associated with haemoptysis may be misdiagnosed with pulmonary embolism [64]. Moreover, some abdominal pain and gastrointestinal symptoms are typically reported in patients [65]. About half of the patients are affected by neurologic disorders (e.g., headache, confusion, agitation, delirium, and hallucinations) [66]. Other common symptoms are: hyponatremia, decreased serum phosphorus, hypophosphatemia, elevated liver-associated enzymes, increased creatine kinase (MM fraction), leucocytosis or leukopenia, thrombocytopenia, high erythrocyte sedimentation rate and C-reactive protein levels, disseminated intravascular coagulation, serum ferritin promotion, elevated lactate dehydrogenase, myoglobinuria, pyuria, and microscopic haematuria [9, 67]. In rare cases, the infection can extend outside the lungs to other organs, through the bloodstream or lymphatic system. In conclusion, the common non-respiratory manifestations include cellulitis, sinusitis, peritonitis, pyelonephritis, pancreatitis, splenomegaly and spleen rupture, pericarditis, myocarditis, wound infections, endocarditis, arthritis, and central nervous system infections [68, 69]. Legionnaires' disease was caused not only by *Lp*, but also by *n-pL* (As explained in paragraph (§) 1.5). The clinical and radiographic characteristics of the LD caused by *n-pL* species are very similar to those caused by *Lp* [9].

1.6.2 Pontiac Fever

The disease's name derived from Pontiac, Michigan, USA. In 1968 an epidemic event broke out among employees of the Oakland County Health Department. Most of them developed a flu-like illness for 2-3 days. All of the 144 employees were hospitalized [70]. The source was identified in the water of condenser drippier, where was isolated a bacteria that some years later it was associated with Legionellaceae family [71]. The incubation period of Pontiac fever is between 5 hours and 3 days, but generally it is characterized by a short duration. The main symptoms include chills, myalgia, fever, headache, arthralgia, malaise, non-productive cough, abdominal pain, fatigue, and a sore throat [9]. Fortunately, it is self-limited and non-fatal illness. The only laboratory detectable alteration is leucocytosis [70]. The most common etiologic agents of this disease are: *Lp*, *L. anisa*, *L. feeleeii*, and *L. micdadei* [72–76]. During the last two decades, Pontiac fever has been poorly documented, the few documented cases show a complete recovery of the patients within two or five days [77, 78].

1.6.3 Risk factors

The risk to contract the LD is related to different factors [9, 74, 77, 81–86]:

- **Human factors:** where the individual susceptibility of the exposed person and the time and intensity of the exposition play a fundamental role. Among the individual susceptibility factors there are the following factors: age > 50 y, male sex, chronic respiratory or heart disease, glucocorticoid treatment, diabetes, organ transplants, long-term smoking, chronic lung disease, lung cancer, immunological statement, etc. For example, the transplant recipients have the highest risk of the nosocomial LD.
- **Bacterial and environmental factors:** the important bacterial factors are the virulence of the strain and bacteria load. Instead, regarding the environment, play an important role the chemical, physical and microbiological parameters of the water, as temperature, nutrients, presence of amoebae, algae and biofilm, that provide bacteria protection and subsistence.
- **Anthropological factors:** the design of the pipeline system also plays a role in the risk factors. In the presence of water stagnation and obstruction phenomena, increase the scale formation, limestone deposits, wear and corrosion processes. These factors offering to *Legionella* the protection from the biocidal activity of the disinfectants where were in use. Moreover, other factors are represented by some cooling systems like cooling towers or

evaporative condenser that are used as gas-liquid heat exchanger, in order to reduce liquids temperature, rather than the exposition and using of hot water during the staying in hotels rooms with discontinuous occupancy or into private homes, increase the risk of LD.

Some risk factors may also affect the course of the disease. In fact, the hospitalization and lethality rate for the LD are significantly higher for immunocompromised people or people that present concomitant pathologies, with respect to healthy people. The 89.9% of LD died people show a more frequency of pre-existing diseases, compared to those who survive (66.9%) [87].

1.6.4 Research, diagnosis, and treatment

Nowadays exist different way to research *Legionella*, to do the diagnosis and treat the symptoms of LD. Depending on the samples, environmental or clinical/human, the workflow to follow is different.

1.6.4.1 Human samples

LD is a pneumonia with symptoms common to other microorganisms linked to pneumonia (§ 1.6.1-1.6.3) [81]. The LD disease is subjected to class II mandatory notification between infectious disease (§1.6.6.2) [88] moreover *Legionella* is classified as Group 2 among pathogens according to Decree 81/2008 (X Title of D.Lgs.81/2008 as amended and supplemented) [89]. For these reasons, the research and diagnosis *Legionella* in human samples is important. The routine analysis plays a fundamental role to identify the microorganism in order to choose the correct antibiotic therapy reducing the antibiotic resistance mechanisms, the time of hospitalization and health care costs. In addition, during the clusters or epidemic event the epidemiological investigation is based on the match between clinical and environmental isolates. Respect to the species of *Legionella*, the sensitivity (SE) or specificity (SP) of the clinical tests is different: the *Lp1* tests show higher SE and/or SP respect to the other *Lp* or *n-pL* species.

The methods are divided in two categories:

- Direct methods: to identify the presence of the microorganism
 - Isolation of the bacterium by culture technique;
 - Detection of bacterial DNA by PCR;
 - Urinary antigen tests (UATs);
 - Detection of the bacterium in tissues or body fluids by immunofluorescence.
- Indirect methods: to identify the presence of a specific immune response
 - detection of antibodies on sera in the acute and convalescent phase of the disease;

All of these methods are better described in Italian Guidelines [81].

1.6.4.1.1 Culture method

This method is considered the gold standard for the LD, because it permits to keep isolated colonies for further studies (e.g., compare the colonies with the environmental ones, in order to identify the source of infection). The culture method is described in Italian Guideline, Annex 2 [81] and is based on the ISO 11731:2017. The isolation required specific medium and long-time incubation (4-15d). Generally, this method is conducted on lower respiratory tract of lung samples (e.g., pleural fluid, sputum, bronchial and tracheal aspirates; on samples of lung parenchyma taken as a result of a biopsy or autopsy; on blood samples by blood culture). The sampling needs to be performed before the antibiotic therapy, in order to avoid false negative results. Successively, the colonies need to be identified (§1.6.5.3) with agglutination latex test (§ 1.6.5.3.1), direct immunofluorescence (DFA) (§1.6.5.1.4) or sequencing of representative *Legionella* genes (§1.6.5.3.3).

1.6.4.1.2 Real-time PCR

The analysis of real-time PCR (or qPCR) is based on the amplification of specific genes of *Legionella*, [90, 91]. The target involved in this analysis usually are *mip*, *16S rRNA* and *5S rDNA* genes [90, 91]. This method is extremely advantageous with respect to cultural methods, due to the time of analysis and the low quantity of the samples but lacks in that the strain will be destroyed. Moreover, real-time is able to recognize all *Legionella* species.

This technique is not validated, due to some technical limits, i) the results are expressed as genomic unity per liter and not as colonies forming units (cfu) per liter. Unfortunately, the scientific community is not able to convert these measure units. ii) this method is able to recognize a high number of positive samples and/or high quantification values. Due to the detection of DNA of death bacteria and the DNA of *Legionella* which live intra-amoebae [92, 93]. To date, qPCR is carrying out during the epidemiological analysis to test simultaneously a high number of samples. Obviously, in case of positive results, that samples must be analyzed also by cultural methods.

1.6.4.1.3 Urinary antigen test (UAT)

This method was introduced in 1995. Since this year, the UAT supports the diagnosis of LD, with a consequent increased of disease notification, showing an easy and fast performance [94]. This test is currently validated only for *Lp1*, although in some cases show a positivity for other SGs, but it does not work with n-pL [95–97]. Confirmation ends to be obtained by using other diagnostic

methods (e.g., culture, serology). The technique is based on research of a soluble antigen of *Legionella* present in the urine within 1-3 days of the presence of symptoms. Although the pick is between 5-10 days, the presence can be found even after weeks, or in case of immunocompromised patients, even after one year or also after antibiotic treatments [98, 99]. Like all tests, UAT has benefits and drawbacks. Among the benefits are ease of use, speedy execution, and ease of reading, particularly in the early stages of the LD. Among the disadvantages, there are the impossibility to distinguish among the different phases of the LD, and the fact that the antigen of *Legionella* could remain for a long time in the urine.

Unfortunately, have been reported some false positive results [94, 100, 101]. Therefore, to increase the test sensitivity, the urine should be boiled before the test [39]. According to Svarrer et al. [39] in combination with UAT an additional diagnostic test (cultural, serological and PCR examination) should be carried out.

1.6.4.1.4 Direct immunofluorescence

Direct immunofluorescence (DFA) is a immunological method used to detect the presence of the *Legionella* in clinical samples (e.g., sputum, tracheal aspirates, and lung biopsies) [102], but it is also used for the environmental sample. The method is based on the identification of the link between the antigen(s) of *Legionella* and species-specific or serotype-specific monoclonal antibodies labelled with a fluorochrome. The antigen-antibody linkage is visible as fluorescent corpuscles under microscope.

Despite it is able to return a diagnosis of LD in a few hours, this method has a lesser validity than the cultural method. In fact, a comparison between cultural and DFA methods during the LD diagnosis in positive patients, shows a percentage of positivity between 25 and 70% of DFA respect to cultural method. Instead, the SP of DFA depends on the antisera utilized and of the dimension of the samples; in the better condition the SP >99.9% [81]. As happened to UAT, when utilized on sputum the DFA could show a positive result even after 2-4d of the starting time of antibiotic treatment or in case of cavitary pneumonia [99, 103].

During the last year, the diagnosis of LD on clinical samples are doing more with UAT respect to cultural or DFA methods. This has led to an increasing detection of *Lp1* and an under diagnosis of another SG or *n-pL*.

1.6.4.1.5 Indirect immunofluorescence

The indirect immunofluorescence (IFI), as a DFA, is a immunological method used to detect the presence of the *Legionella* in clinical samples. The SE and SP of the test were evaluated only for *Lp1* [99], this fact due to false negative results. The diagnosis is carried out through two samplings of serum to avoid the false positive results: the first in the acute phase, and the second during the convalescent phase of the disease. The positive diagnostic results show a fourfold or greater increase in antibody titer between the first and second sera drawn. If only one serum is positive, the result will only return as presumptive diagnosis. Usually, a significant increase in antibody titer occurs in 1-9 weeks after disease onset. The principal limit of the method is the inability to determine the antibody class and thus differentiate between current or previous infection. Other limit is represented by the cross reactivity between *Legionella*, *Campylobacter* e *Pseudomonas* species [104, 105] and by the false negative results due to the late seroconversion (e.g., in individuals with compromised immune defenses or who have advanced age, during which a natural decline in immune response occurs). Due to all these limits, a negative result does not exclude an LD diagnosis.

1.6.4.2 Environmental samples

The research and identification of *Legionella* in the environment are fundamental in the presence of the case(s) to compare the clinical strain with the one present in the environment, in order to identify the source of infection and to be able to undertake any corrective measures. Moreover, the *Legionella* environmental surveillance is mandatory in some environment, according to International and national guidelines [81, 106].

The gold standard for *Legionella* detection in the environmental sample is the culture method. This method requires long time (2-15 d) of incubation, and additional tests for the identification of the *Legionella* species and SGs.

From 2004, the year of publication of *Legionella* genome, it was possible to carry out various molecular techniques. The molecular techniques, although do not replace the culture method, permit to reduce the time of analysis and to obtain accurate species and SGs identifications thanks to the higher SE and SP.

1.6.4.2.1 Culture method

The culture method is the gold standard for the research of *Legionella* in the environmental samples following UNI EN ISO 11731:2017 (§ 3.2) [107]. The sample must be analyzed within 4 days after sampling.

Successively, the colonies found, must be further analyzed for species identification with agglutination latex test (§ 1.6.5.3.1). Moreover, the sequencing of some *Legionella* genes is performed for *Legionella* identification.

1.6.4.2.2 Real-time PCR

The qPCR is already well described in paragraph § 1.6.5.1.2.

1.6.4.3 Identification of *Legionella* isolates

All the colonies isolated by culture method, both in clinical or environmental samples, must be identified in order to characterize them and found a correlation between the clinical and environmental samples during an epidemiological investigation.

The identification could be carried out with phenotypical or genotypical techniques. Among the phenotypical technique, there are the agglutination latex test and DFA. Regarding the genotypical technique the gold standard for the identification of *Legionella* species is the *mip* gene sequencing. The phenotypical technique shows a low SE and SP, and they do not show information useful for the epidemiological surveillance. This information is however produced by molecular methods.

1.6.4.3.1 Agglutination latex test

The agglutination is a serological test, based on the identification of the antigen of species o SG by the utilizing of monoclonal or polyclonal antibody. One of the most used kit is able to recognize *Lp1* or *Lp 2-14* or seven species of n-*pL*.

1.6.4.3.2 MALDI-TOF MS

The identification of *Legionella* can be evaluated also with MALDI-TOF MS technique. This technique makes it possible to accurately measure the molecular weight of macromolecules of biological interest and determine their identity based on their mass-to-charge ratio. Using the database contained in the instrument software, spectral profiles, typical for each species, are identified.

Despite being high discrimination power, and its large use in clinical routine nosocomial infections, it is not commonly used in both clinical and environmental *Legionella* surveillance. Moreover, in the

literature its application in *Legionella* identification application is described only in few dated works [108–110].

1.6.4.3.3 Sequencing of *Legionella* genes

The identification of *Legionella* could be carried out also with molecular techniques. These techniques are able to recognize all the species of *Legionella*. The identification is performed by the *mip* gene sequencing following ESGLI (ESCMID (European Society of Clinical Microbiology and Infectious Diseases) Study Group for *Legionella* Infections) protocol [111, 112]. The genomic typing of *Lp* strains is carried out by SBT. This technique was developed by members of the European Legionnaires' Disease Surveillance Network (ELDSNet) and it is a rapid, simple, discriminatory and portable method [113]. SBT is able to study the sequence of seven housekeeping and virulence *Legionella* genes (*flaA*, *pilE*, *asd*, *mip*, *mompS*, *proA*, and *neuA*) [114]. The presence of online database allows the storage and the sharing of the identified sequences. The database was available on the EWGLI website (<http://www.ewgli.org/>), and the sequences were compared with the ones presents in the ELDSNet database to assign the ST allelic profile (http://www.hpa-bioinformatics.org.uk/legionella/legionella_sbt/php/sbt_query_frontpage.php). To date, this technique is considered the gold standard for genotyping of *Lp* [115]. SBT, generally, is performed only during epidemiological investigations after the notification of a single case, or cluster or epidemic event. The routine laboratories, are not confident with technique, considering the high personal specialization required.

1.6.4.4 LD Therapy

Legionella are a Gram-negative intracellular bacteria, for this reason the antibiotic treatment must be able to concentrate and be active at the intracellular level [116] and be distributed and persist inside of the infected tissue.

Generally, the antibiotics used are quinolones, macrolides, rifampicin and, with lower efficiency, tetracyclines [87]. In particular, quinolones (i.e., levofloxacin) work better than macrolides. Among the latter, azithromycin appeared superior to clarithromycin, and both of these drugs proved to be superior to erythromycin [81, 87, 117]. In particularly severe cases or in the case of immunocompromised patients, these antibiotics are used in combination.

1.6.5 Epidemiology

The exact global incidence of Legionnaires' disease is unknown. This is due to the awareness levels, diagnostic methods and reports differ from country to country [118].

Although in recent years there is a greatly improved of the epidemiological surveillance of LD, this disease remains under-diagnosed and also under-notified [81]. The improvement was thanks to increased awareness among physicians and the availability of a simple, noninvasive diagnostic test (i.e., UAT).

The mortality rate of the LD ranges is between 5% and 10% and depends on the treatment and the initial clinical condition of the affected person [119].

1.6.5.1 European epidemiological trend

The European (EU) surveillance of LD was started in 1995. Historically, four countries (Italy, France, Germany, and Spain) reported the majority number of cases in the EU area (Figure 5).

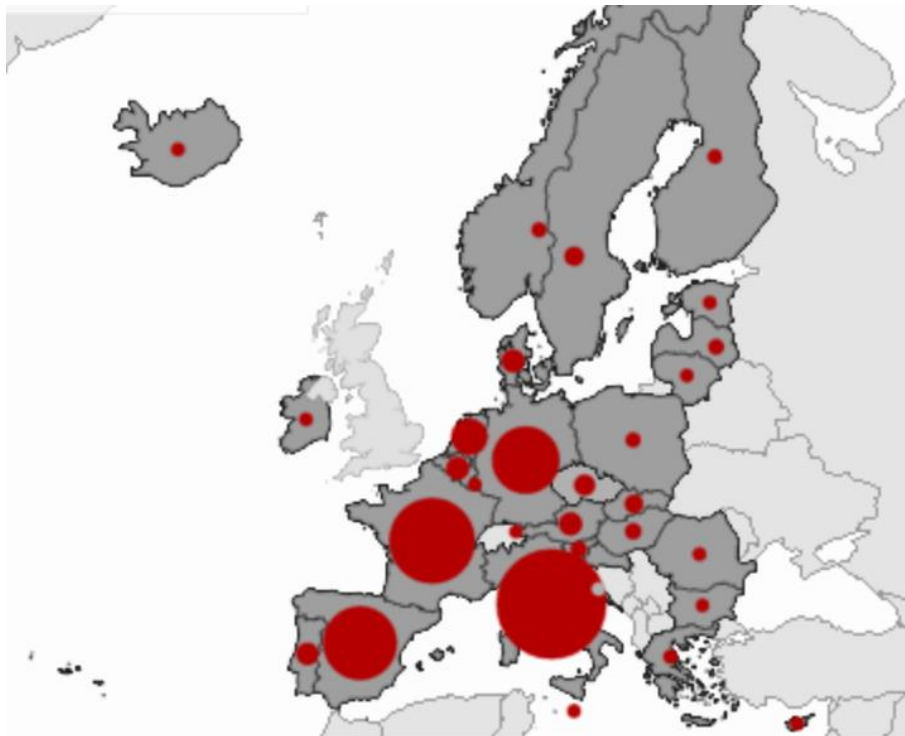


Figure 5 Distribution of reported cases in Europe in 2021

These four countries, in 2021, have made up 74.5% of all EU cases, making this microorganism an alarming emerging pathogen (Figure 6). Regardless, LD remains an important cause of potential and preventable morbidity and mortality in Europe [36, 120].

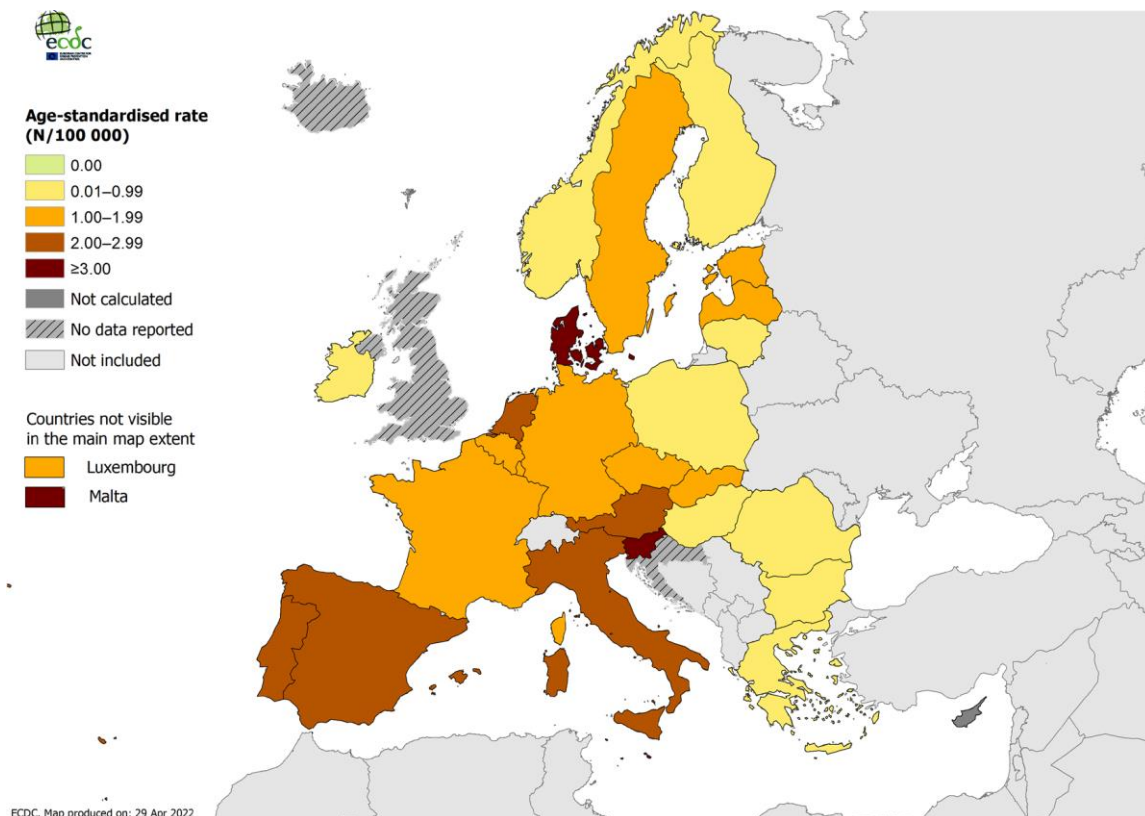


Figure 6 Distribution of cases of LD per 100 000 inhabitants by country, EU/EEA, 2020

Up to 2019 the trend of the LD was in increment; however, in 2020 the reported cases decreased (Figure 7).

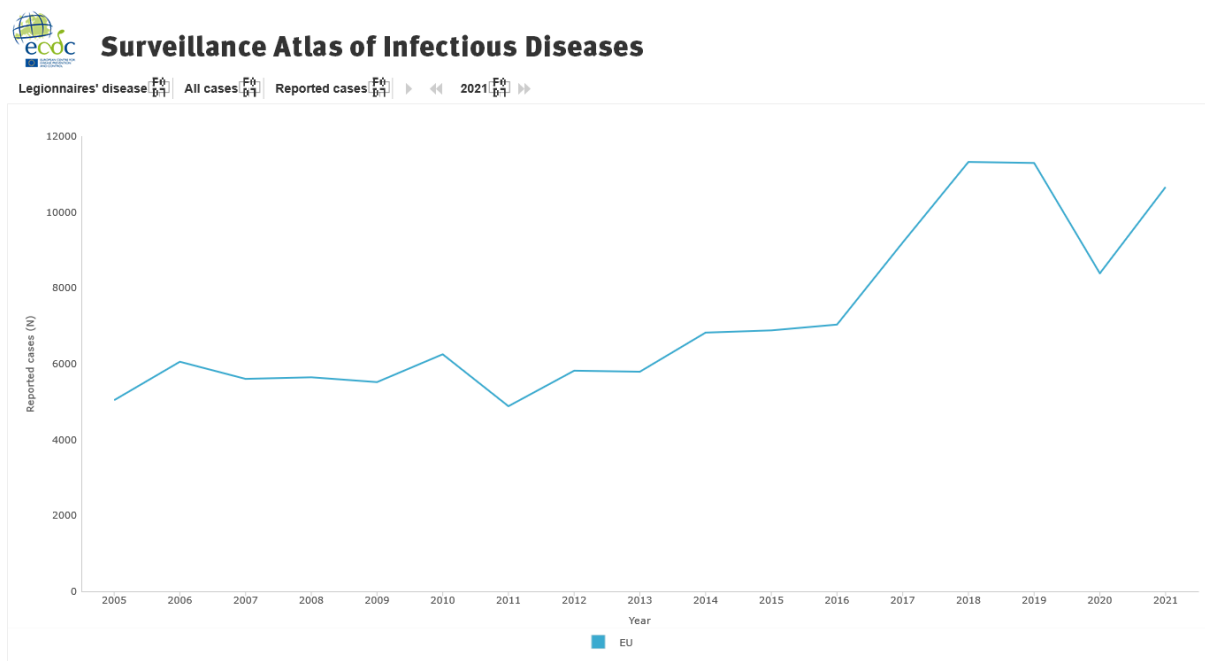


Figure 7 Reported cases of LD per year (2005–2021) in Europe

The COVID-19 pandemic may influence this decrement, due to the similarity between LD and COVID-19 symptoms. Another reason could be associated with the lower number of travels and other

restrictive measures applied (i.e., mask). Verhasselt et al. [121] highlighted that about 20% of COVID-19 patients have supposedly contracted *Lp* co-infection, due to the presence of a positive anti-*Legionella* IgM antibody titre. This suggests that COVID-19 patients were able to contract LD both during hospitalization and in the following months, as a consequence of the residual pulmonary impairment found in these individuals and lower LD risk prevention during the pandemic peak period [122]. Also, the renovation work, carried out for the renovation of some hospitals or clinics to increase beds has resulted in the detachment of biofilms inside the water pipeline systems and release of microorganisms [123–125].

The largest number of cases in 2020 (87%) was reported by UAT, the trend is generally like 2019 (88-90%). Eleven percent of these cases were confirmed by culture or PCR test. Among these, the 5% of the cases were caused by n-*pL* (i.e., *L. anisa*, *L. bozemanii*, *L. longbeachae*, *L. micdadei*, etc.) and more than 80% by *Lp1*.

According to ECDC, in 2021, a prevalent number of the cases (77.2%) was notified in the community, following by 6.1% of domestic travel cases, while nosocomial cases were 5.4% (Figure 8) [120].

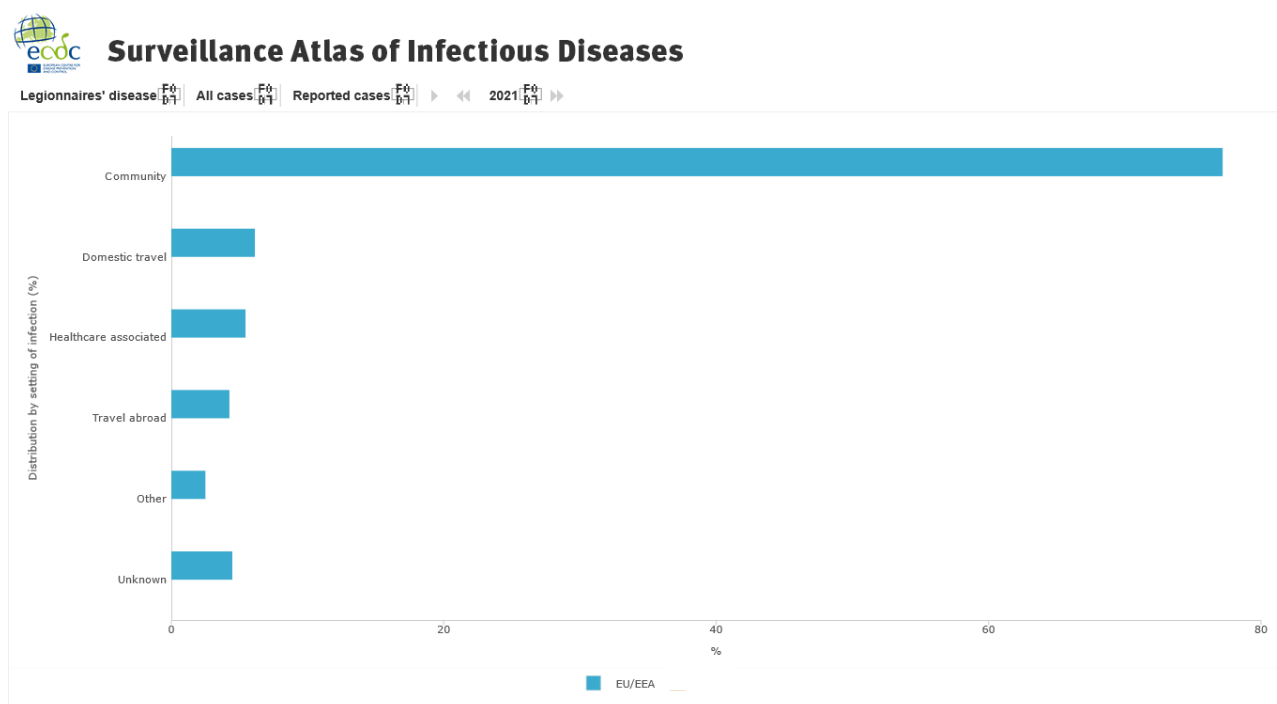


Figure 8 LD distribution by setting of infection in EU/EEA, 2021

The priority continues to be the assistance to the countries with fewer number of notification rates to improve both the diagnosis and LD report to public health authorities [36].

1.6.5.2 Italian epidemiological data

In Italy, LD has statutory notifiable disease in class II, following DM 15/12/1990 and subsequent additions and modifications [88].

Therefore, within 48 h of diagnosis, the surveillance form completed by clinicians, must be sent to the local health authority (ASL), Department of Infectious Diseases (DMIPI) of the Italian National Institute of Health (ISS) [81].

In Italy, the incidence rate of LD in 2021 amounted to 46.0 cases per million of inhabitants (Figure 9).

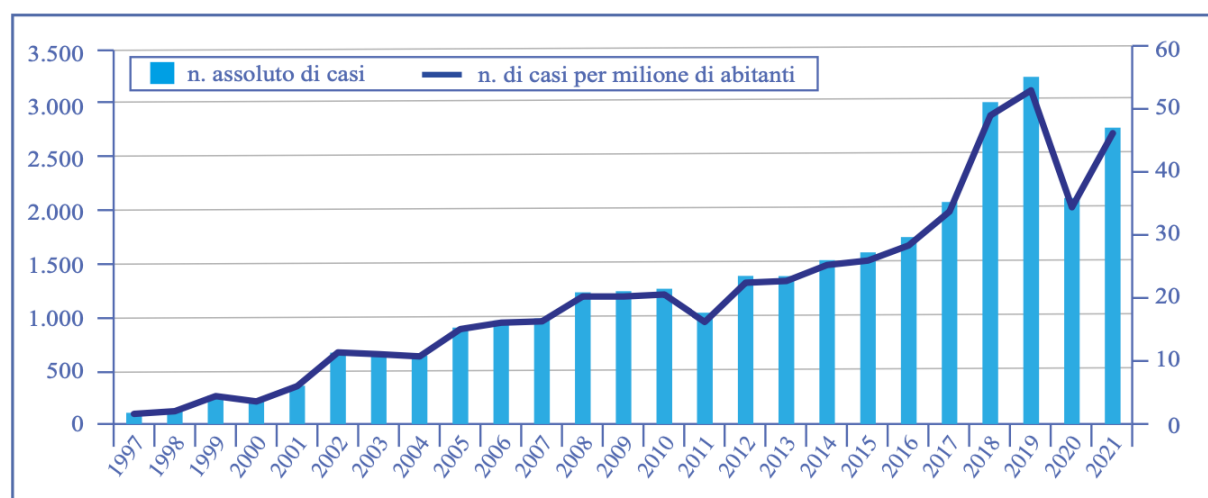


Figure 9 Trends in the number of cases and incidence rate of LD in Italy, 1997-2021

There is a different distribution among north, center and south of Italy, with a decrease gradient from north to south (71,9 and 11,6 cases per million of inhabitants respectively).

The European trend is confirmed with 79.5% of cases notified by 6 regions: Lombardy, Veneto, Emilia-Romagna, Tuscany, Lazio, and Piedmont, and the remaining 20.5% of from the remaining 15 regions.

The most frequently used diagnostic method has been UAT (97.4%), followed by 1.5% of PCR test and 0.3% of culture technique. All these cases were caused by Lp; the prevalent SG was 1, followed by 6, 3, and 8.

According to data from the Italian Institute of Health, among the 2726 cases notified, 83.7% were classified as community-acquired, 8.9% were travel-associated, 6.8% were hospital and other health care facilities linked (Figure 10) [120, 126].

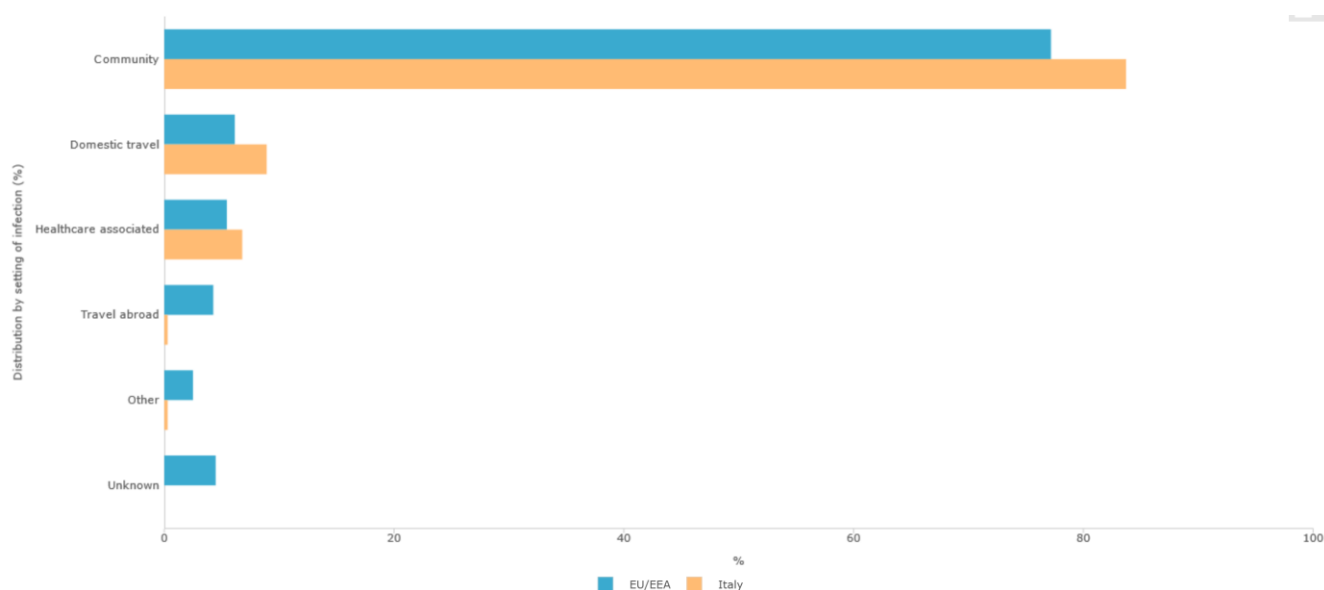


Figure 10 Comparison of LD distribution by setting of infection in EU/EEA and Italy, 2021

Nevertheless, the percentage of cases is higher in the community with respect to nosocomial cases, there is a big difference about the lethality rate between them 11.2% and 40.5%, respectively [126].

1.7 Prevention and control of *Legionella* infections

The strategies to control the LD in the community, recreational, civic, and industrial sphere start from prevention. The prevention should start from the correct design and implementation of the pipeline systems (e.g., potable water, hot water system, process water, etc...). The synergistic action among the engineering, microbiology and construction leads to the proper implementation of a good water safety plan.

For this reason, the environmental prevention and monitoring are essential, not only in nosocomial environment, but in all the environments where the human may enter into contact with *Legionella* (e.g., industries, hospitals, healthcare facilities, spa facilities, hotels, campsites, swimming pools, recreational fountains, gardening activities, etc.).

In this area, play a key role the Italian National and Regional Guidelines for Surveillance and Control of Legionellosis and the directive on "Prevention and Protection of Workers: D.Lgs. 81/2008" application of European directive (UE) 2019/1936 [81, 89, 127]. *Legionella* is classified in the group two of pathogenic agents in the Title X of those laws.

To safeguard human health is mandatory the development of a "Risk Assessment Plan" (DVR), articulated into three sequential and correlate phases (risk assessment, risk management and risk

communication) that must be applied in all the structure or activity where the human health could be at risk.

Going into detail in the DVR, the three sequential and correlate phases are as follow:

- **Risk assessment:** is an investigation that evaluates the facility (intended use, type of hosts, etc.) and the WDS and the conditions that could link the actual or potential presence of *Legionella* in the systems, to the possibility of infection. The information obtained through this investigation must be reported to the facility manager or his or her supervisor, who will then inform all those involved in the control and prevention of *Legionella* infections;
- **Risk management:** represents the interventions and procedures aimed at permanently removing or constantly containing the criticisms identified in the risk assessment phase. All the maintenance or preventive interventions implemented, must be the result of a strategy that considers all system characteristics and possible interactions in the balance of the system;
- **Risk communication:** are all actions aimed at informing, training, and raising awareness among those affected by the potential risk (e.g., stakeholders, control personnel, people exposed, hosts, etc.).

For these reasons, information and training is an essential element in ensuring the proper implementation and functioning of the Guideline for the prevention and control of LD.

As explained in the previous paragraph, the risk assessment evaluates the risk inside of the buildings or systems in case of *Legionella* contamination.

The risk assessment must be carried out and develop by competent multidisciplinary working group, which must be able to best evaluate the system, in order to identify which are the criticalities present and identifying the points of the network and the areas that are able to represent a potential risk for both guests and workers present within the building.

The first step to follow is the inspection of the system through its updated scheme.

A good WDS should have a temperature below 20°C if it is used with cold water or above of 50°C if it is used with hot water, must not have dead spots or blind branches, suitable for stagnation of water and they must be made of suitable materials.

The risk assessment must be reviewed and duly documented each year in the hospital or healthcare facilities, while each two years for the other structure. In case the WDS is subject to modifications, or the *Legionella* analysis provides a positive result with high value that requires an immediate

action, the DVR must be overhauled. Here, plays a role, the risk management. In fact, it is very important to redact, with the help of qualified technical personnel, a timeline plan for the system management and maintenance, where must be included all the sanification procedure and the measures to be implemented for the containment of the risk, with their relative periodicity.

Another important point is the risk communication, because only with the information and formation of the staff is possible to apply all the good practice and measures that permit to reduce the risk of the contamination.

Among the good practices and measures to be applied, there are:

- Maintain a correct water temperature (hot: >50°C, cold: <20°C);
- Avoid dead branches in the system;
- Avoid biofilms formation on showerheads or faucet outlet, utilizing anti-scale products;
- Clean, de-scaling and disinfect cold water storage tanks using chlorine at least once a year;
- Clean, de-scaling and disinfect hot water storage tanks (inspected every month) at least twice a year;
- Use appropriate disinfection treatments, if there is a significant increase in microbial load;
- Etc.

The Italian Guideline for the prevention and control of LD [81] required a DVR for tourist-receptive structures, spas, nosocomial facilities and in general for all those environments (e.g., cooling towers, evaporative condensers, etc.) where there is a biological risk associated with the presence of *Legionella*. Each of these structures has different methods, timing, and sampling points for the detection of *Legionella*, different concentration limits of the *Legionella* and different types of intervention to undertake if these limits are exceeded. In table 1 and 2 there are two examples of nosocomial facilities and cooling tower types of intervention.

Table 1. Types of intervention indicated by concentration of *Legionella* (cfu/L) in water systems with LD risk, carried out in hospital/health facilities.

<i>Legionella</i> concentration (cfu/L)	Type of intervention
Up to 100	Nothing
Between 101 and 1,000	<u>In the absence of cases:</u> - If less than 30% of the samples are positive, the water system must be retested, at least by the same sampling points, after verifying that the current risk control practices are correctly applied. If the result is confirmed, a risk assessment review shall be carried out to identify the necessary additional corrective measures. - If more than 30% of the samples are positive, the water system must be retested, at least from the same sampling point, after verifying that the current risk control practices are correctly applied. If the result is confirmed, a disinfection and a risk assessment review shall be carried out to identify the necessary additional corrective measures.
	<u>In the presence of cases:</u> Regardless of the number of positive samples, carry out a risk assessment review and carry out a disinfection of the plant.

<i>Legionella</i> concentration (cfu/L)	Type of intervention
Between 1,001 and 10,000	<u>In the absence of cases:</u> - If less than 20% of the samples are positive, the water system must be retested, at least by the same dispensers, after verifying that the current risk control practices are correctly applied. If the result is confirmed, a risk assessment review shall be carried out to identify the necessary additional corrective measures. - If more than 20% of the samples are positive, the water system must be disinfected, and a risk assessment review carried out to identify the necessary additional corrective measures. The water system must be resampled, at least in the same sampling points that showed positive results. Increased clinical surveillance is recommended, especially for risk patients. Avoid the use of water from the water system for showers or ablutions that may cause aerosol formation.
	<u>In the presence of cases:</u> Regardless of the number of positive samples, the system must be disinfected, and the risk assessment reviewed to identify the necessary additional corrective measures. The water system must be resampled after disinfection, at least from the same sampling point that showed positive results.
More than 10,000	Both in the presence and absence of cases, the system must be disinfected (replacing positive terminals) and revised risk assessment. The water system must be resampled, at least by the same positive outcomes.

Table 2. Type of intervention indicated for concentrations of *Legionella* (cfu/L) in cooling systems with evaporative towers or evaporative condensers

Legionella concentration (cfu/L)	Type of intervention
Up to 1,000	Verify that current risk control practices are correctly applied.
Between 1,001 and 10,000	The water installation shall be resampled, after verifying that the practical risk control currents are correctly applied and after increasing the dosage of an appropriate biocidal product. If the result is confirmed, a risk assessment review shall be carried out to identify the necessary additional corrective measures.
Between 10,000 and 100,000	Carry out a disinfection with an appropriate biocidal product and review the risk assessment to identify the necessary additional corrective measures, such as any mechanical cleaning of the plant basin to support the disinfection.
More than 100,000	Stop the system, disinfect with an appropriate biocidal product, and review the risk assessment to identify the necessary additional corrective measures, such as any mechanical cleaning of the system basin to support the disinfection. Restart the system when sampling results after disinfection return to <1000 ufc/L levels.

1.7.1 Control measures

Regarding the control of *Legionella* proliferation, to ensure a high quality of the water some procedures have been undertaken. Among them, there are shock treatment, the UV light, chemical disinfectants, etc. [128, 129].

The shock treatment is performed to reach high temperatures (70-80°C) starting from the hot water storage heater and which propagates through the rest of the hot WDS, where the temperature at

the distal outlet reach at least 60-65°C for three consecutive days [130, 131]. The disadvantage is the transitory effect and a low rooting on the biofilm [132].

The UV radiation light is another method for the control of *Legionella* contamination. UV light (254 nm) inactivates bacteria, including *Legionella*, by producing, in their DNA sequence, some thymine dimmers, thereby inhibiting their replication [133, 134]. This method is used at the point of use and could be used in synergy with thermal shock and chemical disinfectants methods. The advantage is that methods are easy to install, safe for the pipelines, and does not show an adverse effect on the taste or potability of the water. The disadvantage is represented by the required maintenance procedure and does not affect the biofilm and its functionality only on the point of use.

Concerning chemical disinfectant, the most famous treatments are chlorination, monochloramine and hydrogen peroxide and silver salt ($\text{H}_2\text{O}_2/\text{Ag}^+$).

Chlorination is the most common treatment used in the water, including the cold one. It works like an oxidizing agent, reacts with the cell membrane. The shock hyperchlorination (20-50 mg/L of residual chlorine concentration at distal outlets for 1-2h) or continuous treatment (1-2 mg/L of residual chlorine concentration) are carried out *Legionella* control [133]. The advantage is that in all the pipeline system the residual chlorine works. The disadvantages are the difficulties to maintain those concentrations given the volatility of chlorine, its corrosive effect on the pipelines, other than the reduction of effectiveness during the time [135].

Monochloramine (NH_2Cl) is the product of ammonia and hypochlorous acid ($\text{NH}_3 + \text{HOCl}$). This compound shows more effect on biofilm and to prevent the outbreak of LD, less complaints of taste and odour with respect to chlorine [133, 136]. Like chlorine, monochloramine can cause corrosion of the pipeline systems.

$\text{H}_2\text{O}_2/\text{Ag}^+$ shows a synergic action against the bacteria, and it is stable at high temperature [133, 137]. The silver salts form a bacteriostatic film inside the pipeline and H_2O_2 is completely soluble in water. The recommended concentration is 20-25 mg/L with a temperature between 40-50°C [115, 138].

The advantages are that this compound does not interact with the pipeline, shows low eco-toxicity, no odour or colour [139]. Among the disadvantage there is the advice to do not use this disinfectant in water systems that supplying the dialysis units [133]. Other than it is potentially influenced by

several factors (i.e., temperature, pH or the presence of substances that hamper its reactivity [140]) and does not show the same functionality on all the *Legionella* species [141].

2. Scope of the study

The presence of a wide natural reservoir makes impossible to eradicate *Legionella*, both in WDS and air conditioning systems. For this reason, a preventive approach is fundamental to control the risk and to minimize the number of cases. In Italy, the National Guidelines for Prevention and Control of Legionellosis and directive on the “Prevention and Protection of Workers: D.Lgs. 81/2008” suggest to perform an appropriate surveillance not only in nosocomial environments, but also in civil environments or in climatization systems, cooling towers, fountains, and in all the environments or buildings where the human beings could be at risk of LD.

Only identifying *Legionella* species and their interaction with the human is possible to do a better prevention.

Here, it plays a fundamental role the development of a WGS approach that permits to increase knowledge of *Legionella* species, especially in regard to the n-pL because they are the less studied and consequently less associated with human disease with respect to *Lp*, although they are the most abundant in the environment.

The Knowledge acquired on *Legionella* species may help the engineers, health authorities, scientific community, and technical staff to formulate new preventive strategies in order to implement the correct measures to apply during the prevention, other than to help the health authorities to better recognize the *Legionella* species during cases or epidemic events.

The WGS approach can help to identify some *Legionella* species that are difficult to isolate and recognize during routine culture technique, improving the knowledge of the ecological status, pathogenicity, and possible resistance of antibiotics of *Legionella* isolates, leading to a better estimation of the real risk of *Legionella* infection.

In conclusion, with this dissertation I would like to give my personal contribution to the scientific community during the surveillance and prevention control stage of LD. If all the teams that carrying out the prevention, could get to know all *Legionella* species and strains that colonize WDS under their control will be very relevant during an epidemiological study and in addition, would support the rapid implementation of the interventions to be carried out, avoids loss of economic resources and time.

3. Material and methods

This study was conducted in the Bologna area among hospitals (and their cooling towers), healthcare facilities, industries, and hotels which, according to Italian Guidelines [81], a water safety plan (WSP) involving a *Legionella* surveillance program was implemented [84, 137, 142–144]. Some hospitals' and healthcare facilities' WDS were treated in continuous with a disinfectant based on a stabilized combination of $\text{H}_2\text{O}_2/\text{Ag}^+$ (H_2O_2 34%, wt/wt and Ag^+ salts 0.003%, wt/wt). Only one healthcare facility's WDS was treated by the UV lamp.

3.1 Sample collection

According to the *Legionella* risk assessment plan, 2 liters of hot- and cold-water were collected in post-flushing modality (1 min) in sterile polytetrafluoroethylene (PTFE) bottles containing a sodium thiosulfate solution (20 mg/L). The sampling, transport and storage of the samples were performed following UNI EN ISO 19458:2006 [145]. Depends on the risk assessment plan, each building was sampled from one to four times per years. The sampling was then repeated when the results were not in line with the Legislation limits (§1.7) [81].

According to Italian Guidelines, the sampling points were selected considering the facilities' size, workers' exposure to contaminated aerosol, the facility's epidemiological data, and, in the case of hospitals or healthcare facilities, it could also consider the number of inpatient rooms and risk level of patients. Since that the *Legionella* was transmitted by inhalation of the contaminated aerosol, the sampling must be performed in each outlet and/or devices supplied by water, able to produce aerosol. The presence of specialized operators was required. The operators must not belong to a risk category and must always wear the appropriate personal protective equipment (PPE) (e.g., mask, gloves), to protect themselves and maintain the sample's sterile condition [81].

To perform a proper sampling plan, is necessary:

- Flush the water for at least one minute;
- Flam with a flame (when technically possible) the inside and outside of the point of sampling; if it is not possible, disinfect with 1% hypochlorite or 70% ethanol, leaving the disinfectant to act for at least 60 seconds;
- Flush the water again for at least 1 minute to remove any residual disinfectant; during this flushing measure the temperature by placing the digital probe thermometer in the flush of water and waiting as long as necessary to reach a nearly constant value;

- When presents measure, through a commercial kit, the concentration of the disinfectant in the water;
- Collect samples [81].

The sampling usually started from the technical room, following by the hot and cold WDS for each floor, usually more hot-water samples (HWS), than cold-water samples (CWS), considering the high rate of *Legionella* proliferation in hot water. The HWSs were sampled in the vicinity of, mid-way to, and away from the technical room, and CWS were sampled away from the water supply point [137]. In the technical room were chosen the following points: water supply point, hot water output, hot water return line, and hot water storage tanks. Regarding the WDS the sampling point including faucets, sinks, showerheads and toilet showerheads in bathrooms or impatient rooms or staff facilities. Of course, the sampling points were sampled by the criterion of rotation.

Samples must be uniquely identified and reported in the sampling scheme (Figure 11).

Place and date of sampling

Sampling ID structure

ID sample	Sampling point	Temperature °C	<u>Disinfectant</u> mg/L	<u>Observation of the sampling</u>
1				
2				
3				
4				
Etc.				

Figure 11 *Example of sampling scheme*

Moreover, during the sampling, the disinfectant residues and temperature values were measured in distal outlets and recorded. Through a digital thermometer for probes (XS Temp 7 Vio PT 100 Thermometer from –200 to +999 °C; Eutech Instruments Pte Ltd., Singapore) the temperature (°C) was measured. Regarding the disinfectant, it was measured (mg/L) on-site by a commercial kit that reveals the residual H₂O₂ component used a colorimetric test. This measurement is semi-quantitatively result, based on production of a blue oxidation product resulting from peroxidase

activity to transfer peroxide oxygen to an organic redox indicator. This blue product can be compared with a color scale present on the kit in a range between 1 and 200 mg/L [137].

The samples were transported out of the light, separating HWS from CWS, and preserved in the fridge at 5 ± 3 °C until the analysis. During the same day or maximum the day after of the sampling, the analysis of the samples was carried out.

3.2 Routinely *Legionella* microbiological analysis and serological identification

All the samples were analyzed for the presence of both *P. aeruginosa* (Ps) and *Legionella*.

The research of Ps permit to understand the dynamics of microbiological contamination related to biofilm presence, of which Ps is one of the major components, other than determining the health risk for patients [81, 141, 144, 146]. Moreover, Ps is a competing microorganism of *Legionella* and its growth in culture could lead to false negative results regarding *Legionella* results. The data regarding Ps were not shown in this dissertation.

Legionella analyses were performed by culture method according to UNI EN ISO 11731:2017 [107].

In Figure 12 was displayed a workflow scheme (Figure 12).

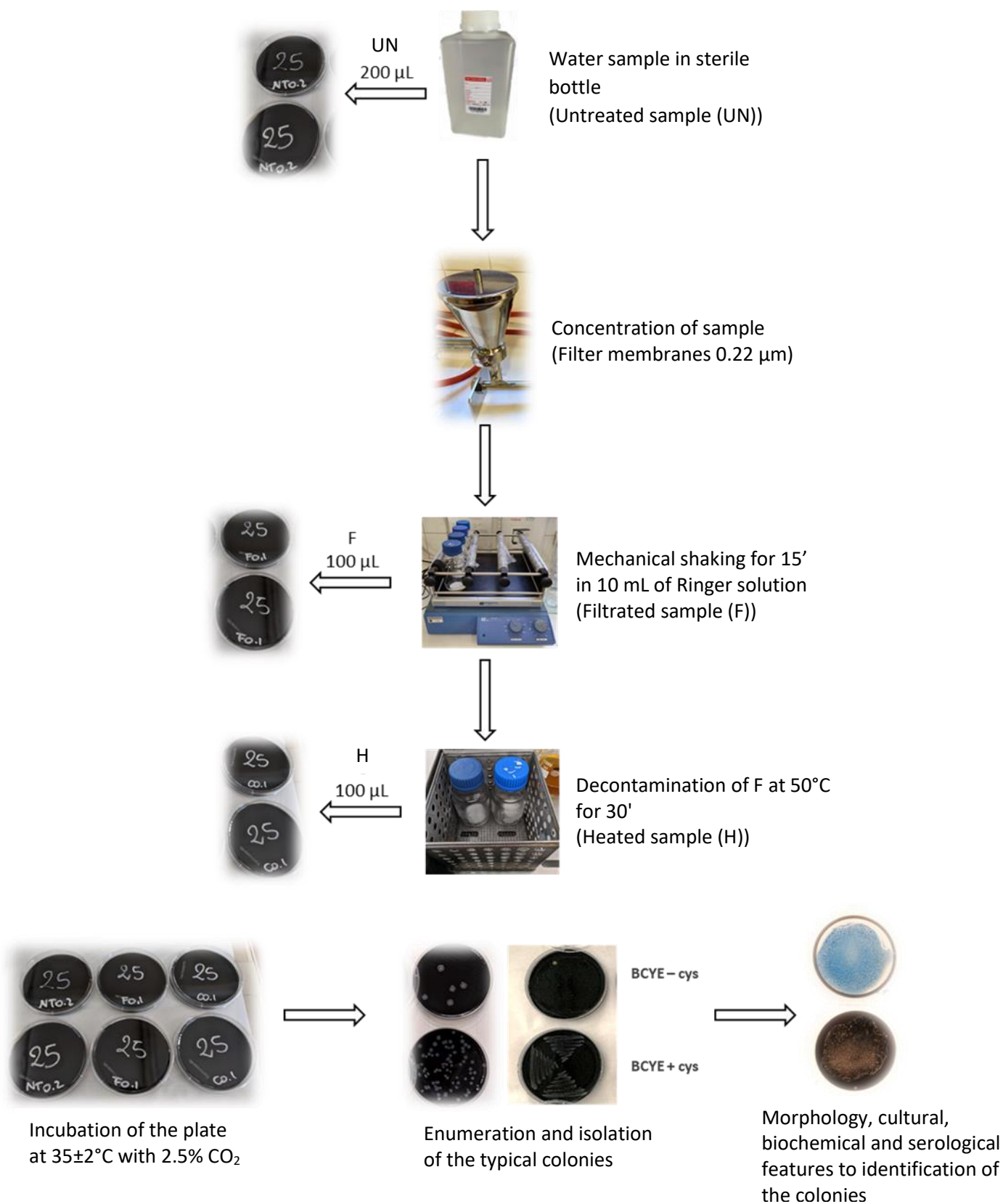


Figure 12 Summary of *Legionella* analysis according to UNI EN ISO 11731:2017.

Briefly, each water sample contained in sterile bottle (“untreated sample” (UN)) was concentrated using 0.22 µm polycarbonate pre-sterilized filter membranes (Sartorius Stedim Biotech, Göttingen, Germany) through a filtering ramp system, obtained a concentrate sample. The latter was then subjected to a mechanical shaking for 15’ in 10 mL of Ringer’s solution (“filtrated” sample (F)).

Each “F” sample was then heated for 30 min at 50 °C, in order to inhibit interfering microflora, with this step the “heated” (H) sample was obtained. In case the samples are highly contaminated (e.g., cooling tower samples or dental unit samples), the filtrate samples were treated by 1:10 dilution buffered acid solution (HCl-KCl, pH 2.2) and leave at room temperature for 5 min before the seeded (“acid” sample (A)).

Before the previous step, for each sample, 0.2 mL of the UN sample was plated onto GVPC agar plates (*Legionella* Glycine Vancomycin Polymyxin Cycloheximide selective medium, Thermo Fisher Scientific, Oxoid Ltd., Basingstoke, UK) and 0.1 mL of each F, H and A sample were plated in duplicate onto GVPC agar. All plates were then incubated aerobically at 35.5±2 °C in presence of 2.5% of carbon dioxide (CO₂) in a humid environment for up to 15 days.

The plates were examined each three days since the analysis, for a total of 15 days.

The colonies with typical or atypical *Legionella* morphology of, were enumerated and then sub-cultured on buffered charcoal yeast extract (BCYE) agar, both with (BCYE Cys+) and without (BCYE Cys-) L-Cysteine (Thermo Fisher Scientific, Oxoid, Ltd., Basingstoke, U.K.).

Only the colony that has grown on BCYE Cys+ are considered *Legionella* species. There is the exceptions represented by two species: *L. oakridgensis* and *L. spiritensis* that are able to grow also in the absence of L-Cysteine [107].

All the isolates grown on BCYE Cys+ were then serologically identified by agglutination test (*Legionella* latex test kit; Thermo Fisher Scientific, Oxoid Ltd.).

The agglutination test permits to identify among *Lp*1 and *Lp* SG2-14 and 7 species of n-pL (*L. anisa*, *L. bozemanii* 1 e 2, *L. dumoffii*, *L. gormanii*, *L. longbeachae* 1 e 2, *L. jordanis* e *L. micdadei*) which are involved in human disease. The isolates that show positive results for *Lp* SG2-14 were then also tested for a single SG by specific polyclonal latex reagents (Biolife, Milan, Italy) in order to identify the SG.

According to the ISO 11731, given that 2L of water was filtered the limit of the technique was 50 cfu/L. For this reason, the result of the samples that do not show *Legionella* growth in the culture medium was expressed as <50 cfu/L [107]. Instead, according to Italian Guideline, the risk level was

set to >100 cfu/L considering the “risk level” the *Legionella* concentration that required intervention measures [81].

During the routine analysis the agglutination test is the last step of *Legionella* analysis. The report usually can report a generic indication of “*Legionella*”, “*L. pneumophila*” or “*Legionella* non-pneumophila species”. This is required by Guidelines that are focused on the quantification of *Legionella* presence.

Because of the agglutination test is not able to identify all n-pL species, it requires additional methods.

3.3 Research approach developed for the study

In the present paragraph of the dissertation, the research approach used during the three years of PhD was described. Although during epidemiological analysis *mip* gene sequencing is performed to identify n-pL, in this paper it is included in the "research approach" as the second level of identification approach followed on all *Legionella* isolates.

The workflow described is firstly based on the comparison between *mip* gene sequencing and a most famous identification technique, routinely used in clinical approach for nosocomial infections and microorganism identification, based on matrix-assisted laser desorption ionization-time-of-flight mass spectrometry (MALDI-TOF MS). In the second step, the results obtained were confirmed by whole genome sequencing (WGS) approach.

3.3.1 Identification by MALDI-TOF MS (Chapter 1)

MALDI-TOF MS is an unconventional approach for *Legionella*, it is an innovative, rapid, and inexpensive technique for species-level microbial identification through the analysis of ribosomal protein patterns. Thanks to the MALDI-TOF MS those laboratories were able to replace or combine most traditional, biochemical or molecular techniques [147–149].

Despite this technique is routinely introduced in clinical laboratories to identify microorganisms associated with infectious disease, it is not applied for *Legionella* identification [108–110, 150–153]. As a first step of my project, all isolates grown on BCYE Cys + with positive, negative, or ambiguous results returned by agglutination test were evaluated by the MALDI Biotyper system (Bruker Daltonik GmbH, Bremen, Germany). Microflex LT mass spectrometer (2,000–20,000 Da, linear positive mode) was utilized for the spectra acquisition and processing. This instrument is able to

recognize the *Legionella* spectra through MALDI Biotyper Compass 4.1 software, whose library (version BDAL 7854) included the spectra of 39 *Legionella* strains, as shown in Table 3.

Table 3 Number of *Legionella* spectra included in MALDI Biotyper software library

<i>L. anisa</i>	<i>L. feeleeii</i>	<i>L. longbeachae</i>	<i>L. rubrilucens</i>
(n = 9)	(n = 8)	(n = 9)	(n = 1)
<i>L. beliardensis</i>	<i>L. geestiana</i>	<i>L. maceachernii</i>	<i>L. sainthelensi</i>
(n = 1)	(n = 1)	(n = 3)	(n = 3)
<i>L. birminghamensis</i>	<i>L. gormanii</i>	<i>L. micdadei</i>	<i>L. santicrucis</i>
(n = 6)	(n = 2)	(n = 5)	(n = 3)
<i>L. bozemanii</i>	<i>L. gratiana</i>	<i>L. moravica</i>	<i>L. tucsonensis</i>
(n = 10)	(n = 5)	(n = 1)	(n = 2)
<i>L. brunensis</i>	<i>L. hackeliae</i>	<i>L. oakridgensis</i>	<i>L. wadsworthii</i>
(n = 2)	(n = 1)	(n = 2)	(n = 1)
<i>L. cherrii</i>	<i>L. impletisoli</i>	<i>L. parisiensis</i>	<i>L. waltersii</i>
(n = 4)	(n = 1)	(n = 1)	(n = 1)
<i>L. cincinnatiensis</i>	<i>L. israelensis</i>	<i>L. pneumophila</i>	<i>L. worsleiensis</i>
(n = 2)	(n = 1)	(n = 3)	(n = 1)
<i>L. dresdenensis</i>	<i>L. jamestowniensis</i>	<i>L. pneumophila ssp fraseri</i>	<i>L. yabuuchiae</i>
(n = 1)	(n = 1)	(n = 4)	(n = 1)
<i>L. dumoffii</i>	<i>L. jordanis</i>	<i>L. pneumophila ssp pascullei</i>	<i>Legionella sp.</i>
(n = 7)	(n = 2)	(n = 2)	(n = 1)
<i>L. erythra</i>	<i>L. lansingensis</i>	<i>L. pneumophila ssp pneumophila</i>	
(n = 1)	(n = 1)	(n = 6)	

The experiments were performed comparing two different methods: direct smear vs. full extraction method. Usually, the full extraction method is recommended for Gram-positives species, Actinomycetales, and fungi, because they are on the thick cell wall. No significant differences were found within the two methods in term of identification and quality of spectra.

For this reason, to identify presumptive *Legionella* isolates the direct smear method was applied.

Briefly, starting from a fresh culture (24-48h of incubation), a small amount of bacterial biomass was picked and spotted in duplicate onto a MALDI Biotyper target. After that, the latter were overlaid with 1 mL of MALDI Biotyper matrix solution and left drying before the instrument measurement. The results were expressed as Log (score) as follows: ≥ 2.0 (“high confidence level”), between 1.7 and 1.99 (“low confidence level”) and a score between 0.00 and 1.69 indicate a non-reliable identification. The literature reports a second level of identification that could be applied when the Log (score) is between 1.4 and 1.69. In this case the identification could be considered reliable if the first four proposed identification results are identical [154, 155]. Moreover, reliable results can be obtained when the difference in the Log (score) values between the first and second-best matches is ≥ 0.3 [156]. The data obtained were elaborated by GraphPad Prism software version 8.0.1 for Windows (GraphPad Software, San Diego, California, United States) to determine the SE, SP, and confidence interval (CI) at the 95% level of significance.

3.3.2 Identification of *Legionella* spp. by *mip* and *rpoB* genes sequencing (Chapters 1-6)

InstaGene Purification Matrix (Bio-Rad, Hercules, CA) was utilized for DNA extraction, and Qubit fluorometer (Thermo Fisher Scientific, Paisley, United Kingdom) was used for DNA quantification. PCR analysis for *n-pL* isolates was performed to determine the gene sequences of *mip* and *rpoB* as described by Ratcliff et al. and Ko et al. [111, 157], respectively.

Degenerate primers [111] and modified by **M13** tailing were utilized for *mip* sequencing to avoid noise in the DNA sequence [Legmip_f_**M13F** **TGTA AACGACG GCCAGTGGGRATTVTTTATGAAGATGARAYTGG** and Legmip_r_**M13R** **CAGGAAACAGCTATGACCTCRTTNGGDCCDATNGGNCCDCC**] [158].

mip-gene amplification was carried out in a 50 μ L reaction containing DreamTaq Green PCR Master Mix 2 $\times$ (Thermo Fisher Diagnostic Basingstoke, United Kingdom) and 40 pmol of each primer; 100 ng of the DNA extracted from the presumptive *Legionella* colonies was added as a template. Specifically, the *mip* amplicons were sequenced using tailed **M13** forward and reverse primers to obtain a complete coverage of the region of interest.

Amplification of *mip* gene was performed in a thermocycler under the following conditions: i) pre-denaturation for 3 min at 96°C; ii) 35 cycles consisting of 1 min at 94°C for denaturation, 2 min at 58°C for annealing, and 2 min at 72°C for extension; iii) by a final extension at 72°C for 5 min; iv) held the reaction mixtures at 4°C.

Instead, *rpoB* gene amplification was performed in a 50- μ L reaction volume containing 100 ng of template DNA, 40 pmol of each primer (RL1 5'- GATGATATCGATCAYCTDGG-3'; RL25'- TTCVGGCGTTTCAATNGGAC-3'), 1 U of Taq polymerase, and a PCR mixture consisting of PCR buffer 10 $\times$, 1.5 mM MgCl₂, and 250 μ M deoxynucleoside triphosphates (dNTPs).

Amplification of *rpoB* was performed in a thermocycler under the following conditions: i) 35 cycles, and each cycle consisted of 30 s at 94°C for denaturation, 30 s at 55°C for annealing, and 30 s at 72°C for extension; ii) final extension at 72°C for 10 min; iii) held the reaction mixtures at 4°C.

PCR products were visualized by electrophoresis on 2% agarose gel.

Following purification by ExoSAP-IT™ PCR Product Cleanup kit (Applied Biosystems, Foster City, CA), DNA was sequenced using BigDye Chemistry and analyzed on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA). Raw sequencing data were assembled using CLC Main Workbench 7.6.4 software.

Final purification was carried out by using BigDye X Terminator™ Purification kit (Applied Biosystems) and then samples were undergone to capillary electrophoresis on ABI PRISM 3100 Genetic Analyzer. Raw sequencing data were assembled using CLC Main Workbench 22.0.2 software (QIAGEN, Hilden, Germany). The sequences obtained were compared to several databases. The *mip* gene sequences were compared to the sequences deposited in the *Legionella mip*-gene sequence database using a similarity analysis tool. EWGLI have established an accessible web database (http://bioinformatics.phe.org.uk/cgi-bin/Legionella/mip/mip_id.cgi) that contains sequence data from described species and allows the identification of *n-pL*. Species-level identification was performed on the basis of a similarity score of 98-100% compared to the sequences in the database [112] and considering the intra- and interspecies intervals of variation previously described by Ratcliff et al. [111]. The 205 nucleotide *mip* sequences generated for this study were submitted to GenBank, at the follow accession numbers: MW021138, MW052863-MW053066.

Concerning *rpoB* gene, the sequences were compared to type strain sequences deposited in the National Center for Biotechnology Information (NCBI) from several culture collections, including the American Type Culture Collection (ATCC), National Collection of Type Cultures, Central Public Health Laboratory (NCTC), NITE Biological Research Center, National Institute of Technology and Evaluation (NBRC), and Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSM). According to Adékambi et al. and Ko et al., the cutoff used for *rpoB* gene sequence-based identification was fixed at 94-95% similarity percentage using an *rpoB* gene fragment of 300 to 600 bp [157, 159].

3.3.3 Phylogenetic and Allelic Diversity Analysis (Chapters 1-6)

To estimate the relationship among the *Legionella* species involved in the dissertation, a multiple sequence alignment (MSA) and a phylogenetic tree were performed on the *mip* and *rpoB* gene sequences and the three positive controls. For each taxon, identified as previously described, the reference *mip* and *rpoB* genes sequences of the corresponding type strains from different culture collections were retrieved and added to the analysis.

When required, manual editing was performed on the sequences, trimming them to the same length as the reference sequence. The nucleotide sequences were aligned by the multiple sequence comparison by log-expectation (MUSCLE) program [160]. The obtained MSA was passed to FastTree [161], a tool for inferring approximate maximum likelihood phylogenetic trees. FastTree uses Jukes-Cantor as a genetic distance model and the Shimodaira-Hasegawa test to estimate the reliability of each split in the tree [162]. Branch lengths were transformed to be equal, as in a cladogram. Branch labels display the substitutions per site. Both MUSCLE and FastTree were performed in Geneious Prime 2021.1.1¹, retaining the default settings.

3.3.4 Elaboration of matrices for the *rpoB* gene: definition of the intra- and interspecies intervals of variation (Chapter 2)

Legionella type strains retrieved from the NCBI, were used to determine the ranges of the intra- and interspecies intervals of variation for the *rpoB* gene, resulting in a pairwise identity matrix for the entire gene with a length from 4,101 to 4,143 bp (The matrices developed is reported in the results section of Chapter 2, Fig. 1a and b) and for the 329-bp selected region (The matrices developed is reported in the results section of Chapter 2, Fig. 2a and b), corresponding to the amplicon suggested by Ko et al. [157].

The list of type strains used in the study is reported in Table 3 (Materials and methods section of Chapter 2). The matrices were built using the MUSCLE program [160] in Geneious Prime 2021.1.1, retaining the default settings. The developed matrices permitted us to obtain a minimum and a maximum value of variation to establish intra- and interspecies intervals of divergence for the identification of the environmental isolates used in the present dissertation.

¹ <https://www.geneious.com>

3.3.5 Identification of *Legionella* by Whole-Genome Sequencing (WGS) (Chapter 3-6)

Among all the *n-pL* strains studied, the interval of interspecies found suggest the presence of a presumptive novel species. In this dissertation, three of them were investigated: 30cs62^T, 31fl33^T and 8cVS16^T by WGS approach.

One hundred nanograms of DNA was used to prepare the library for the next-generation sequencing (NGS) carried out by Illumina Nextera XT DNA Library Preparation kit (Illumina, New England Biolabs, Ipswich, MA, USA). The Illumina NextSeq 500 platform (2 × 150 bp paired-end reads) was used to perform the sequencing. Raw reads were assembled using TORMES v.1.2.0 [163], an automated pipeline for whole bacterial genome analysis, using the default parameters. TORMES performed a sequence quality filtering (PRINSEQ v. 0.20.4) and a de novo genome assembly (SPAdes v. 13.4.1) [164]. The generated contigs were passed to CSAR v1.1.1 [165], 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. Depending on the strain, different *Legionella* species was selected as a reference sequence was selected as a reference sequence, based on the output of KmerFinder [166–168] that predicts prokaryotic species based on the number of overlapping k-mers, i.e., 16-mers, between the query genome and genomes in a reference database (NCBI). Regarding 30cs62^T the reference sequence was *L. fallonii* (GenBank accession number NZ_LN614827.1), for 31fl33^T was *L. hackeliae* strain ATCC 35250 (GenBank accession no. LN681225.1) and for 8cVS16^T was *Legionella* sp. PC1000 (GenBank accession no. NZ_CP059400.1).

A further refinement was carried out by remapping the reads on the CSAR scaffolds, using the Geneious Prime 2022.0.2 software [169]. Benchmarking Universal Single-Copy Orthologs (BUSCO) v.5.0.0 [170] was performed to evaluate the completeness of the two genome assemblies. The final draft genomes were submitted to the GenBank requiring the annotation by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP v.4.3) [171]. The OrthoANI package [172] was used to measure the intra- and inter-species genome similarities by average nucleotide identity (ANI) among the assembled draft genomes. Further, FastANI [173] through DFAST [174] was performed against 13000 prokaryotic reference genomes from NCBI to assess the taxonomic identity. It was also measured based on BLAST+ (ANIb) and MUMmer (ANIm) using JSpeciesWS [175]. The relatedness of our strains to *Legionella*-known type strains was further analyzed by applying a digital DNA–DNA hybridization (dDDH) to the closest related strain based on the previous ANI outcome. 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+ [176] as a local alignment tool. The GGDC uses a Genome Blast Distance Phylogeny (GBDP) to infer genome-to-genome distances between pairs of entirely or partially sequenced genomes. Below the similarity threshold of the in silico DDH (70%) and the ANI analysis (95%), two strains are considered distinct species [177, 178].

3.3.6 Characterization of a novel *Legionella* species (Chapter 6)

Until now, the taxonomy characterization was performed only for strain 30cs62^T.

The taxonomy characterization was based on some cultural, biochemical, morphological tests to describe the new species. To test the best temperature able to promote the strain growth, the subcultures of the strain were performed on BCYE Cys+ agar plates and incubated at different temperatures: 32, 35 and 37 °C with 2.5 % of CO₂. The strain was also analyzed with Gram- and Ziehl Neelsen-staining and autofluorescence was evaluated under a Woods lamp (long-wavelength 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).

To evaluate the biochemical characteristics, the strain reaction was investigated 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 instruction. The detection of β -lactamase production was evaluated using Oxoid nitrocefin solution (Thermo Fisher Scientific Diagnostic). Moreover, to understand the antimicrobial susceptibility, the strain was tested carried out the 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 ESCMID) tested for some *Lp* isolates were used for the interpretation of minimal inhibition concentration (MIC) data [179].

References

1. Fraser D, Tsai T, Orenstein W, Parkin W, Beecham H, Sharrar R, et al. Legionnaires' disease: description of an epidemic of pneumonia. *N Engl J Med* 1977; **297**: 1189–1197.
2. Brenner D, Steigerwalt A, McDade J. Classification of the Legionnaires' disease bacterium: *Legionella pneumophila*, genus novum, species nova, of the family Legionellaceae, familia nova. *Ann Intern Med* 1979; **90**: 656–658.
3. McDade J, Shepard C, Fraser D, Tsai T, Redus M, Dowdle W. Legionnaires' disease: isolation of a bacterium and demonstration of its role in other respiratory disease. *N Engl J Med* 1977; **297**: 1197–1203.
4. Parte AC, Sardà Carbasse J, Meier-Kolthoff JP, Reimer LC, Göker M. List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. *Int J Syst Evol Microbiol* 2020; **70**: 5607–5612.
5. Bacterio.net L. List of prokaryotic names with standing in nomenclature (LPSN), Legionella species list, moves to the DSMZ. <https://lpsn.dsmz.de/genus/legionella>. .
6. Fields BS, Benson RF, Besser RE. *Legionella* and legionnaires' disease: 25 years of investigation. *J Clin Microbiol* 2002; **15**: 506–526.
7. Salinas MB, Fenoy S, Magnet A, Vaccaro L, Gomes TD, Hurtado C, et al. Are pathogenic *Legionella pneumophila* a common bacteria in Water Distribution Networks? *Water Res* 2021; **196**: 117013.
8. Diederer BMW. *Legionella* spp. and Legionnaires' disease. *J Infect* 2008; **56**: 1–12.
9. Jomehzadeh N, Moosavian M, Saki M, Rashno M. *Legionella* and legionnaires' disease: An overview. *J Acute Dis* 2019; **8**: 221–232.
10. Fliermans CB, Cherry WB, Orrison LH, Smith SJ, Tison DL, Pope DH. Ecological distribution of *Legionella pneumophila*. *Appl Environ Microbiol* 1981; **41**: 9–16.
11. Ohno A, Kato N, Yamada K, Yamaguchi K. Factors Influencing Survival of *Legionella pneumophila* Serotype 1 in Hot Spring Water and Tap Water. *Appl Environ Microbiol Env Microbiol* 2003; **69**: 2540–2547.
12. Serrano-Suárez A, Dellundé J, Salvadó H, Cervero-Aragó S, Méndez J, Canals O, et al. Microbial and physicochemical parameters associated with *Legionella* contamination in hot water recirculation systems. *Environ Sci Pollut Res* 2013; **20**: 5534–5544.
13. Bonetta S, Bonetta S, Ferretti E, Balocco F, Carraro E. Evaluation of *Legionella pneumophila* contamination in Italian hotel water systems by quantitative real-time PCR and culture methods. *J Appl Microbiol* 2010; **108**: 1576–1583.
14. Rasheduzzaman M, Singh R, Haas CN, Gurian PL. Required water temperature in hotel plumbing to control *Legionella* growth. *Water Res* 2020; **182**: 115943.
15. Arvand M, Jungkind K, Hack A. Contamination of the cold water distribution system of health care facilities by *Legionella pneumophila*: Do we know the true dimension? *Eurosurveillance* 2011; **16**: 1–6.
16. Wadowsky RM, Wolford R, McNamara AM, Yee RB. Effect of temperature, pH, and oxygen level on the multiplication of naturally occurring *Legionella pneumophila* in potable water. *Appl Environ Microbiol* 1985; **49**: 1197–1205.
17. Katz SM, Hammel JM. The Effect of Drying, Heat, and pH on the Survival of *Legionella pneumophila*.

Ann Clin Lab Sci 1987; **17**: 150–156.

18. Newton HJ, Ang DKY, Van Driel IR, Hartland EL. Molecular pathogenesis of infections caused by *Legionella pneumophila*. *Clin Microbiol Rev* 2010; **23**: 274–298.
19. Boamah DK, Zhou G, Ensminger AW, O'Connor TJ. From many hosts, one accidental pathogen: The diverse protozoan hosts of *Legionella*. *Front Cell Infect Microbiol* 2017; **7**.
20. Lau HY, Ashbolt NJ. The role of biofilms and protozoa in *Legionella* pathogenesis: Implications for drinking water. *J Appl Microbiol* 2009; **107**: 368–378.
21. Potočnjak M, Magdalenic Z, Dijan M, Rebić D, Gobin I. Environmental factors affecting the survival of soil dwelling *Legionella longbeachae* in water. *Ann Agric Environ Med* 2016; **23**: 452–455.
22. Hines SA, Chappie DJ, Lordo RA, Miller BD, Janke RJ, Lindquist HA, et al. Assessment of relative potential for *Legionella* species or surrogates inhalation exposure from common water uses. *Water Res* 2014; **56**: 203–213.
23. Mercante JW, Winchell JM. Current and emerging *Legionella* diagnostics for laboratory and outbreak investigations. *Clin Microbiol Rev* 2015; **28**: 95–133.
24. Carlson KM, Boczek LA, Chae S, Ryu H. Legionellosis and recent advances in technologies for *Legionella* control in premise plumbing systems: A review. *Water (Switzerland)* 2020; **12**: 1–22.
25. Molofsky AB, Swanson MS. Differentiate to thrive: Lessons from the *Legionella pneumophila* life cycle. *Mol Microbiol* 2004; **53**: 29–40.
26. Allegra S, Leclerc L, Massard PA, Girardot F, Riffard S, Pourchez J. Characterization of aerosols containing *Legionella* generated upon nebulization. *Sci Rep* 2016; **6**.
27. Schoen ME, Ashbolt NJ. An in-premise model for *Legionella* exposure during showering events. *Water Res* 2011; **45**: 5826–5836.
28. Correia AM, Ferreira JS, Borges V, Nunes A, Gomes B, Capucho R, et al. Probable Person-to-Person Transmission of Legionnaires' Disease. *N Engl J Med* 2016; **374**: 497–498.
29. Borges V, Nunes A, Sampaio DA, Vieira L, Machado J, Simões MJ, 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.
30. Gomez-Valero L, Chiner-Oms A, Comas I, Buchrieser C, Hershberg R. Evolutionary Dissection of the Dot/Icm System Based on Comparative Genomics of 58 *Legionella* Species. *Genome Biol Evol* 2019; **11**: 2619–2632.
31. Horwitz MA, Silverstein SC. Legionnaires' disease bacterium (*Legionella pneumophila*) multiples intracellularly in human monocytes. *J Clin Invest* 1980; **66**: 441–450.
32. Horwitz MA. Formation of a novel phagosome by the legionnaires' disease bacterium (*Legionella pneumophila*) in human monocytes. *J Exp Med* 1983; **158**: 1319–1331.
33. Swanson MS, Isberg RR. Association of *Legionella pneumophila* with the macrophage endoplasmic reticulum. *Infect Immun* 1995; **63**: 3609–3620.
34. Katz SM, Hashemi S. Electron microscopic examination of the inflammatory response to *Legionella pneumophila* in guinea pigs. *Lab Invest* 1982; **46**: 24–32.
35. Roy CR, Tilney LG. The road less traveled: Transport of *Legionella* to the endoplasmic reticulum. *J Cell Biol* 2002; **158**: 415–419.
36. ECDC: European Centre for Disease Prevention and Control. Legionnaires' disease. Annual

epidemiological report for 2020. 2022.

37. Yu VL, Plouffe JF, Pastoris MC, Stout JE, Schousboe M, Widmer A, et al. Distribution of *Legionella* species and serogroups isolated by culture in patients with sporadic community-acquired legionellosis: An international collaborative survey. *J Infect Dis* 2002; **186**: 127–128.
38. Beauté J, Robesyn E. European Centre for Disease Prevention and Control - Legionnaires' disease in Europe 2015. 2017.
39. Svarrer CW, Lück C, Elverdal PL, Uldum SA. Immunochromatic kits Xpect *Legionella* and BinaxNOW *Legionella* for detection of *Legionella pneumophila* urinary antigen have low sensitivities for the diagnosis of Legionnaires' disease. *J Med Microbiol* 2012; **61**: 213–217.
40. Hoppe J, Ünal CM, Thiem S, Grimpe L, Goldmann T, Gaßler N, et al. *PilY1* promotes *Legionella pneumophila* infection of human lung tissue explants and contributes to bacterial adhesion, host cell invasion, and twitching motility. *Front Cell Infect Microbiol* 2017; **7**: 1–13.
41. Coil DA, Anné J. The role of *fimV* and the importance of its tandem repeat copy number in twitching motility, pigment production, and morphology in *Legionella pneumophila*. *Arch Microbiol* 2010; **192**: 625–631.
42. Otto GP, Wu MY, Clarke M, Lu H, Roger Anderson O, Hilbi H, et al. Macroautophagy is dispensable for intracellular replication of *Legionella pneumophila* in *Dictyostelium discoideum*. *Mol Microbiol* 2004; **51**: 63–72.
43. Hägele S, Köhler R, Merkert H, Schleicher M, Hacker J, Steinert M. *Dictyostelium discoideum*: A new host model system for intracellular pathogens of the genus *Legionella*. *Cell Microbiol* 2000; **2**: 165–171.
44. Watarai M, Derre I, Kirby J, Gowney JD, Dietrich WF, Isberg RR. *Legionella pneumophila* is internalized by a macropinocytotic uptake pathway controlled by the *dot/icm* system and the mouse *Lgn1* locus. *J Exp Med* 2001; **194**: 1081–1095.
45. Hilbi H, Segal G, Shuman HA. *Icm/Dot*-dependent upregulation of phagocytosis by *Legionella pneumophila*. *Mol Microbiol* 2001; **42**: 603–617.
46. Isberg RR, O'Connor TJ, Heidtman M. The *Legionella pneumophila* replication vacuole: Making a cosy niche inside host cells. *Nat Rev Microbiol* 2009; **7**: 13–24.
47. Hubber A, Roy CR. Modulation of host cell function by *Legionella pneumophila* type IV effectors. *Annu Rev Cell Dev Biol* 2010; **26**: 261–283.
48. Chauhan D, Shames SR. Pathogenicity and Virulence of *Legionella*: Intracellular replication and host response. *Virulence* 2021; **12**: 1122–1144.
49. Gomez-Valero L, Rusniok C, Jarraud S, Vacherie B, Rouy Z, Barbe V, et al. Extensive recombination events and horizontal gene transfer shaped the *Legionella pneumophila* genomes. *BMC Genomics* 2011; **12**.
50. Gomez-Valero L, Rusniok C, Rolando M, Neou M, Dervins-Ravault D, Demirtas J, et al. Comparative analyses of *Legionella* species identifies genetic features of strains causing Legionnaires' disease. *Genome Biol* 2014; **15**: 1–21.
51. Burstein D, Amaro F, Zusman T, Lifshitz Z, Cohen O, Gilbert JA, et al. Genomic analysis of 38 *Legionella* species identifies large and diverse effector repertoires. *Nat Genet* 2016; **48**: 167–175.
52. Shames SR, Liu L, Havey JC, Schofield WB, Goodman AL, Roy CR. Multiple *Legionella pneumophila* effector virulence phenotypes revealed through high-throughput analysis of targeted mutant

libraries. *Proc Natl Acad Sci U S A* 2017; **114**: E10446–E10454.

53. Kowalczyk B, Chmiel E, Palusinska-szys M. The Role of Lipids in *Legionella* -Host Interaction. 2021.
54. Palusinska-Szys M, Luchowski R, Gruszecki W, Choma A, Szuster-Ciesielska A, Lück C, et al. The Role of *Legionella pneumophila* Serogroup 1 Lipopolysaccharide in Host-Pathogen Interaction. *Front Microbiol* 2019; **10**: 1–17.
55. Molofsky AB, Shetron-Rama LM, Swanson MS. Components of the *Legionella pneumophila* flagellar regulon contribute to multiple virulence traits, including lysosome avoidance and macrophage death. *Infect Immun* 2005; **73**: 5720–5734.
56. Hawn TR, Verbon A, Lettinga KD, Zhao LP, Li SS, Laws RJ, et al. A Common Dominant TLR5 Stop Codon Polymorphism Abolishes Flagellin Signaling and Is Associated with Susceptibility to Legionnaires' Disease. *J Exp Med* 2003; **198**: 1563–1572.
57. Rasch J, Theuerkorn M, Ünal C, Heinsohn N, Tran S, Fischer G, et al. Novel cycloheximide derivatives targeting the moonlighting protein Mip exhibit specific antimicrobial activity against *Legionella pneumophila*. *Front Bioeng Biotechnol* 2015; **3**: 1–8.
58. Scaturro M, Girolamini L, Pascale MR, Mazzotta M, Marino F, Errico G, et al. Case Report: First Report of Fatal *Legionella pneumophila* and *Klebsiella pneumoniae* Coinfection in a Kidney Transplant Recipient. *Front Med* 2022; **9**: 1–6.
59. Den Boer JW, Yzerman EPF, Schellekens J, Lettinga KD, Boshuizen C, Van Steenberghe JE, et al. A Large Outbreak of Legionnaires' Disease at a Flower Show, the Netherlands, 1999. 2002; **8**: 37–43.
60. Boshuizen HC, Neppelenbroek SE, Van Vliet H, Schellekens JFP, Den Boer JW, Peeters MF, et al. Subclinical *Legionella* infection in workers near the source of a large outbreak of legionnaires disease. *J Infect Dis* 2001; **184**: 515–518.
61. Burillo A, Pedro-Botet ML, Bouza E. Microbiology and Epidemiology of Legionnaire's Disease. *Infect Dis Clin North Am* 2017; **31**: 7–27.
62. Chidiac C, Che D, Pires-Cronenberg S, Jarraud S, Campese C, Bissery A, et al. Factors associated with hospital mortality in community-acquired legionellosis in France. *Eur Respir J* 2012; **39**: 963–970.
63. World Health Organization, Water Sanitation and Health Team. Guidelines for drinking-water quality. Vol. 1: Recommendations, 3rd ed. 2004. Geneva, Switzerland.
64. Moosavian M, Fathollahzadeh B, Amoli K, Moazzami N. Isolation of *Legionella Pneumophila* Serogroups 1, 8 & 10 (Causative Agents of Legionnaires' Disease) from Water Sources in Tehran, Iran. 1998; **2**: 83–87.
65. Mulazimoglu L, Victor LY. Can Legionnaires disease be diagnosed by clinical criteria? A critical review. *Chest* 2001; **120**: 1049–1053.
66. Jespersen S, Sogaard OS, Schønheyder HC, Fine MJ, Østergaard L. Clinical features and predictors of mortality in admitted patients with community- and hospital-acquired legionellosis: A Danish historical cohort study. *BMC Infect Dis* 2010; **10**.
67. Castillo N, Rajasekaran A, Ali S. Legionnaires' disease: A review. *Infect Dis Clin Pr* 2016; **24**.
68. Fernández JA, López P, Orozco D, Merino J. Clinical study of an outbreak of Legionnaire's disease in Alcoy, southeastern Spain. *Eur J Clin Microbiol Infect Dis* 2002; **21**: 729–735.
69. Amodeo MR, Murdoch DR, Pithie AD. Legionnaires' disease caused by *Legionella longbeachae* and *Legionella pneumophila*: Comparison of clinical features, host-related risk factors, and outcomes. *Clin Microbiol Infect* 2010; **16**: 1405–1407.

70. Kaufmann A, McDade F, Patton JE, Bennett C, Skaliy P, Feeley JC, et al. Pontiac fever: isolation of the etiologic agent (*Legionella pneumophila*) and demonstration of its mode of transmission. *Am J Epidemiol* 1981; **114**: 337–347.
71. Glick TH, Gregg MB, Berman B, Mallison G, Rhodes Jr. WW, Kassanoff IRA. Pontiac Fever: An epidemic of unknown etiology in a health department: I. Clinical and epidemiologic aspects. *Am J Epidemiol* 1978; **107**: 149–160.
72. Fallon RJ, Stack BHR. Legionnaires' disease due to *Legionella anisa*. *J Infect* 1990; **20**: 227–229.
73. Fenstersheib MD, Miller M, Diggins C, Liska S, Detwiler L, Werner SB, et al. Outbreak of Pontiac fever due to *Legionella anisa*. *Lancet* 1990; **336**: 35–37.
74. Phin N, Parry-Ford F, Harrison T, Stagg HR, Zhang N, Kumar K, et al. Epidemiology and clinical management of Legionnaires' disease. *Lancet Infect Dis* 2014; **14**: 1011–1021.
75. Jones TF, Benson RF, Brown EW, Rowland JR, Crosier SC, Schaffner W. Epidemiologic Investigation of a Restaurant-Associated Outbreak of Pontiac Fever. *Clin Infect Dis* 2003; **37**: 1292–1297.
76. Wang C, Chuai X, Liang M. *Legionella feeleyi*: pneumonia or Pontiac fever? Bacterial virulence traits and host immune response. *Med Microbiol Immunol* 2019; **208**: 25–32.
77. Muder RR, Yu VL. Infection due to *Legionella* species other than *L. pneumophila*. *Clin Infect Dis* 2002; **35**: 990–998.
78. Edelstein PH, Roy CR. Legionnaires' disease and Pontiac fever. In: Elsevier Saunders (ed). *Bennett JE, Dolin R, Blaser MJ, editors. Masters BR. Mandell, Douglas, and Bennett's principles and practice of infectious diseases*. 2015. pp 2633–2644.
79. Griffin GK, Pope TLJ, Armstrong P. Pittsburgh Pneumonia Agent in a cadaveric renal transplant recipient. *Clin Nephrol* 1983; **20**: 263–266.
80. Mehta P, Patel JD, Milder JE. *Legionella micdadei* (Pittsburgh pneumonia agent). Two infections with unusual clinical features. *JAMA* 1983; **249**: 1620–1623.
81. Italian National Institute of Health. Guidelines for Prevention and Control of Legionellosis. Approvate in Conferenza Stato-Regioni Seduta Del 7 Maggio 2015. Italy, 2015. 2015.
82. Wolter N, Carrim M, Cohen C, Tempia S, Walaza S, Sahr P, et al. Legionnaires' Disease in South Africa, 2012–2014. 2016; **22**: 2012–2014.
83. Di Pippo F, Di Gregorio L, Congestri R, Tandoi V, Rossetti S. Biofilm growth and control in cooling water industrial systems. *FEMS Microbiol Ecol* 2018; **94**: 1–13.
84. Girolamini L, Brattich E, Marino F, Pascale MR, Mazzotta M, Spiteri S, et al. Cooling Towers Influence on an Urban Environment: A Predictive Model to Control and Prevent Legionella Risk and Legionellosis Events. *SSRN Electron J* 2022; **228**: 109891.
85. Špaleková M, Kotrbancová M, Fulová M, Šimonyiová D. Risk of Legionellosis from Exposure to Water Aerosol from Industrial Cooling Tower. *Int Arch Public Heal Community Med* 2019; **3**: 1–8.
86. Hamilton KA, Hamilton MT, Johnson W, Jjemba P, Bukhari Z, Lechevallier M, et al. Risk-Based Critical Concentrations of *Legionella pneumophila* for Indoor Residential Water Uses. *Environ Sci Technol* 2019; **53**: 4528–4541.
87. Cianciotto NP, Kwaik Y, Edelstein PH, Fields BS, Geary D, Harrison TG, et al. *Legionella* - State of the art 30 years after its recognition. 2006. AMS Press, Washington.
88. Italian Ministry of Health. Decreto Ministeriale 15 dicembre 1990. Sistema di sorveglianza sulle

malattie infettive. 6 . 1991. Gazzetta Ufficiale, Italy.

89. Italian Work Ministry. Legislative Decree 09.04.2008, n.81. Implementation of Article 1 of the Law of 3 August 2007, n. 123, Concerning the Protection of Health and Safety in the Workplace. OJ. General Series n.101 of 30-04-2008—Suppl. Ordinary n. 108. *15 Maggio 2008* 2008; **129**: 2008.
90. Cloud JL, Carroll KC, Pixton P, Erali M, Hillyard DR. Detection of *Legionella* species in respiratory specimens using PCR with sequencing confirmation. *J Clin Microbiol* 2000; **38**: 1709–1712.
91. Murdoch DR. Nucleic acid amplification tests for the diagnosis of pneumonia. *Clin Infect Dis* 2003; **36**: 1162–1170.
92. Alleron L, Merlet N, Lacombe C, Frère J. Long-Term Survival of *Legionella pneumophila* in the Viable But Nonculturable State After Monochloramine Treatment. *Curr Microbiol* 2008; **57**: 497–502.
93. Lee J V., Lai S, Exner M, Lenz J, Gaia V, Casati S, et al. An international trial of quantitative PCR for monitoring *Legionella* in artificial water systems. *J Appl Microbiol* 2011; **110**: 1032–1044.
94. Rota MC, Fontana S, Montaña-Remacha C, Scaturro M, Caporali MG, Vullo V, et al. Legionnaires' disease pseudoepidemic due to falsely positive urine antigen test results. *J Clin Microbiol* 2014; **52**: 2279–2280.
95. Benson RF, Tang PW, Fields BS. Evaluation of the Binax and Biotest urinary antigen kits for detection of Legionnaires' disease due to multiple serogroups and species of *Legionella*. *J Clin Microbiol* 2000; **38**: 2763–2765.
96. Olsen CW, Elverdal P, Jørgensen CS, Uldum SA. Comparison of the sensitivity of the *Legionella* urinary antigen EIA kits from Binax and Biotest with urine from patients with infections caused by less common serogroups and subgroups of *Legionella*. *Eur J Clin Microbiol Infect Dis* 2009; **28**: 817–820.
97. Okada C, Kura F, Wada A, Inagawa H, Lee GH, Matsushita H. Cross-reactivity and sensitivity of two *Legionella* urinary antigen kits, biotest EIA and binax NOW, to extracted antigens from various serogroups of *L. pneumophila* and other *Legionella* species. *Microbiol Immunol* 2002; **46**: 51–54.
98. Kohler RB, Winn WC, Wheat LJ. Onset and duration of urinary antigen excretion in Legionnaires disease. *J Clin Microbiol* 1984; **20**: 605–607.
99. Lück PC, Helbig JH, Schuppler M. Epidemiology and Laboratory Diagnosis of *Legionella* Infections. *J Lab Med* 2002; **26**: 174–182.
100. Deforges L, Legrand P, Tankovic J, Brun-buisson C, Lang P, Soussy C-J. Case of False-Positive Results of the Urinary Antigen Test for *Legionella pneumophila*. *Clin Infect Dis* 1999; **29**: 953–954.
101. Bailleul E, Magerman K, Mewis A, Peeters V, Rummens J-L, Cartuyvels R. False-positive result with BinaxNOW *Legionella* Antigen immunochromatographic (ICT) assay: response to Helbig et al. (2001). *J Med Microbiol* . 2004. England. , **53**: 173
102. Su HP, Tung SK, Tseng LR, Tsai WC, Chung TC, Tsung CC. Identification of *Legionella* species by use of an oligonucleotide array. *J Clin Microbiol* 2009; **47**: 1386–1392.
103. Stout JE, Yu VL. Experiences of the first 16 hospitals using copper-silver ionization for *Legionella* control: implications for the evaluation of other disinfection modalities. *Infect Control Hosp Epidemiol* 2003; **24**: 563–568.
104. Marshall LE, Boswell TCJ, Kudesia G. False positive *Legionella* serology in *campylobacter* infection: *Campylobacter* serotypes, duration of antibody response and elimination of cross-reactions in the indirect fluorescent antibody test. *Epidemiol Infect* 1994; **112**: 347–357.
105. Boswell TCJ. Serological cross reaction between *Legionella* and *Campylobacter* in the rapid

microagglutination test. *J Clin Pathol* 1996; **49**: 584–586.

106. World Health Organization. Legionella and the Prevention of Legionellosis. *Emerg Infect Dis* 2008; **14**: 1006a – 1006.
107. UNI EN ISO 11731:2017 Water quality - Enumeration of *Legionella*. 2017. Geneva, Switzerland.
108. Gaia V, Casati S, Tonolla M. Rapid identification of *Legionella* spp. by MALDI-TOF MS based protein mass fingerprinting. *Syst Appl Microbiol* 2011; **34**: 40–44.
109. Trnková K, Kotrbancová M, Špaleková M, Fulová M, Boledovičová J, Vesteg M. MALDI-TOF MS analysis as a useful tool for an identification of *Legionella pneumophila*, a facultatively pathogenic bacterium interacting with free-living amoebae: A case study from water supply system of hospitals in Bratislava (Slovakia). *Exp Parasitol* 2018; **184**: 97–102.
110. Svarrer CW, Uldum SA. The occurrence of *Legionella* species other than *Legionella pneumophila* in clinical and environmental samples in Denmark identified by *mip* gene sequencing and matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clin Microbiol Infect* 2012; **18**: 1004–1009.
111. 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.
112. Fry NK, Afshar B, Bellamy W, Underwood AP, Ratcliff RM, Harrison TG, et al. Identification of *Legionella* spp. by 19 European reference laboratories: Results of the European Working Group for *Legionella* Infections External Quality Assessment Scheme using DNA sequencing of the macrophage infectivity potentiator gene a. *Clin Microbiol Infect* 2007; **13**: 1119–1124.
113. Gaia V, Fry NK, Harrison TG, Peduzzi R. the Potential for True Portability in Legionellosis Outbreak Investigation. *Society* 2003; **41**: 2932–2939.
114. Hutchinson EJ, Joseph C BCE. A European surveillance scheme for travel associated legionnaire s disease. *Euro Surveill* 1996; **1**: 33–39.
115. Shuval H, Yarom R, Shenman R. An Innovative Method for the Control of *Legionella* Infections in the Hospital Hot Water Systems with a Stabilized Hydrogen Peroxide-Silver Formulation. *Int J Infect Control* 2009; **5**.
116. Horwitz MA. The legionnaires' disease bacterium (*Legionella pneumophila*) inhibits phagosome-lysosome fusion in human monocytes. *J Exp Med* 1983; **158**: 2108–2126.
117. Pedro-Botet L, Yu VL. *Legionella*: Macrolides or quinolones? *Clin Microbiol Infect* 2006; **12**: 25–30.
118. Cunha BA, Burillo A, Bouza E. Legionnaires' disease. *Lancet* 2016; **387**: 376–385.
119. WHO WHO. Legionellosis. <https://www.who.int/news-room/fact-sheets/detail/legionellosis>. Accessed 23 Dec 2022.
120. European Center for Disease Prevention and Control. Surveillance atlas of infectious diseases. <https://atlas.ecdc.europa.eu/public/index.aspx?Dataset=27&HealthTopic=30>. Accessed 27 Sep 2022.
121. Verhasselt HL, Buer J, Dedy J, Ziegler R, Steinmann J, Herbstreit F, et al. COVID-19 co-infection with *Legionella pneumophila* in 2 tertiary-care hospitals, Germany. *Emerg Infect Dis* 2021; **27**: 1535–1537.
122. Camões J, Lobato CT, Beires F, Gomes E. *Legionella* and SARS-CoV-2 Coinfection in a Patient With Pneumonia – An Outbreak in Northern Portugal. *Cureus* 2021; **13**: 2020–2022.
123. De Giglio O, Diella G, Lopuzzo M, Triggiano F, Calia C, Pousis C, et al. Impact of lockdown on the microbiological status of the hospital water network during COVID-19 pandemic. *Environ Res* 2020;

191: 110231.

124. Nagy DJ, Dziejewski DM, Codru N, Lauper UL. Understanding the distribution of positive *Legionella* samples in healthcare-premise water systems: Using statistical analysis to determine a distribution for *Legionella* and to support sample size recommendations. *Infect Control Hosp Epidemiol* 2021; **42**: 63–68.
125. Cristina ML, Spagnolo AM, Casini B, Baggiani A, Del Giudice P, Brusaferro S, et al. The impact of aerators on water contamination by emerging gram-negative opportunists in at-risk hospital departments. *Infect Control Hosp Epidemiol* 2014; **35**: 122–129.
126. Rota MC, Caporali MG, Bella A, Scaturro M, Giannitelli S, Ricci ML. I risultati del sistema di sorveglianza della legionellosi nel 2021. *Boll Epidemiol Naz* 2022; **2**: 9–16.
127. Emilia-Romagna Region. Regional Guidelines for Surveillance and Control of Legionellosis. Delibera Della Giunta Regionale 12 Giugno 2017, N. 828. 2017; **2012**: 1–136.
128. Blanc DS, Carrara P, Zanetti G, Francioli P. Water disinfection with ozone, copper and silver ions, and temperature increase to control *Legionella*: Seven years of experience in a university teaching hospital. *J Hosp Infect* 2005; **60**: 69–72.
129. Canals O, Serrano-Suárez A, Salvadó H, Méndez J, Cervero-Aragó S, Ruiz de Porras V, et al. Effect of chlorine and temperature on free-living protozoa in operational man-made water systems (cooling towers and hot sanitary water systems) in Catalonia. *Environ Sci Pollut Res* 2015; **22**: 6610–6618.
130. Ashbolt N. Environmental (Saprozoic) Pathogens of Engineered Water Systems: Understanding Their Ecology for Risk Assessment and Management. *Pathogens* 2015; **4**: 390–405.
131. Whiley H, Bentham R, Brown MH. *Legionella* persistence in manufactured water systems: Pasteurization potentially selecting for thermal tolerance. *Front Microbiol* 2017; **8**: 1–6.
132. Bédard E, Boppe I, Kouamé S, Martin P, Pinsonneault L, Valiquette L, et al. Combination of heat shock and enhanced thermal regime to control the growth of a persistent *Legionella pneumophila* strain. *Pathogens* 2016; **5**: 13–18.
133. European Working Group for *Legionella* Infections (ESGLI). European technical guidelines for the Prevention, Control and Investigation, of Infections Caused by *Legionella* species. 2017.
134. Rattanakul S, Oguma K. Inactivation kinetics and efficiencies of UV-LEDs against *Pseudomonas aeruginosa*, *Legionella pneumophila*, and surrogate microorganisms. *Water Res* 2018; **130**: 31–37.
135. Lin YE, Stout JE, Yu VL. Controlling *Legionella* in Hospital Drinking Water: An Evidence-Based Review of Disinfection Methods. *Infect Control Hosp Epidemiol* 2011; **32**: 166–173.
136. Mancini B, Scurti M, Dormi A, Grottola A, Zanotti A, Cristino S. Effect of monochloramine treatment on colonization of a hospital water distribution system by *Legionella* spp.: A 1 year experience study. *Environ Sci Technol* 2015; **49**: 4551–4558.
137. Girolamini L, Dormi A, Pellati T, Somaroli P, Montanari D, Costa A, et al. Advances in *Legionella* control by a new formulation of hydrogen peroxide and silver salts in a hospital hot water network. *Pathogens* 2019; **8**: 1–21.
138. Casini B, Aquino F, Totaro M, Miccoli M, Galli I, Manfredini L, et al. Application of hydrogen peroxide as an innovative method of treatment for *Legionella* control in a hospitalwater network. *Pathogens* 2017; **6**: 1–12.
139. Juven BJ, Pierson MD. Antibacterial effects of hydrogen peroxide and methods for its detection and quantitation. *J Food Prot* 1996; **59**: 1233–1241.

140. McDonnell G, Russell D. Antiseptics and Disinfectants: Activity, Action, and Resistance. *Am Soc Microbiol* 1999; **12**: 147–179.
141. Girolamini L, Salaris S, Pascale MR, Mazzotta M, Cristino S. Dynamics of *Legionella* Community Interactions in Response to Temperature and Disinfection Treatment: 7 Years of Investigation. *Microb Ecol* 2022; **83**: 353–362.
142. Girolamini L, Pascale MR, Mazzotta M, Spiteri S, Marino F, Salaris S, et al. Combining Traditional and Molecular Techniques Supports the Discovery of a Novel *Legionella* Species During Environmental Surveillance in a Healthcare Facility. *Front Microbiol* 2022; **13**: 1–11.
143. Mazzotta M, Girolamini L, Pascale MR, Lizzadro J, Salaris S, Dormi A, et al. The role of sensor-activated faucets in surgical handwashing environment as a reservoir of *Legionella*. *Pathogens* 2020; **9**: 1–17.
144. Girolamini L, Mazzotta M, Lizzadro J, Pascale MR, Dormi A, Salaris S, et al. Sit bath systems: A new source of *Legionella* infection. *PLoS One* 2020; **15**: 1–19.
145. UNI EN ISO 19458:2006 - Water quality — Sampling for microbiological analysis. 2006. Geneva, Switzerland.
146. Marion CR, Dela Cruz CS, Niederman MS, Kulkarni HS, Sockrider M. What is legionnaires' disease? *Am J Respir Crit Care Med* 2016; **193**: P5–P6.
147. Maier T, Klepel S, Renner U, Kostrzewa M. Fast and reliable MALDI-TOF MS-based microorganism identification. *Nat Methods* 2006; **3**.
148. Croxatto A, Prod'hom G, Greub G. Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. *FEMS Microbiol Rev* 2012; **36**: 380–407.
149. Singhal N, Kumar M, Kanaujia PK, Virdi JS. MALDI-TOF mass spectrometry: An emerging technology for microbial identification and diagnosis. *Front Microbiol* 2015; **6**: 1–16.
150. Moliner C, Ginevra C, Jarraud S, Flaudrops C, Bedotto M, Couderc C, et al. Rapid identification of *Legionella* species by mass spectrometry. *J Med Microbiol* 2010; **59**: 273–284.
151. Fujinami Y, Kikkawa HS, Kurosaki Y, Sakurada K, Yoshino M, Yasuda J. Rapid discrimination of *Legionella* by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Microbiol Res* 2011; **166**: 77–86.
152. He Y, Chang TC, Li H, Shi G, Tang YW. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry and database for identification of *Legionella* species. *Can J Microbiol* 2011; **57**: 533–538.
153. Dilger T, Melzl H, Gessner A. Rapid and reliable identification of waterborne *Legionella* species by MALDI-TOF mass spectrometry. *J Microbiol Methods* 2016; **127**: 154–159.
154. Christner M, Rohde H, Wolters M, Sobottka I, Wegscheider K, Aepfelbacher M. Rapid identification of bacteria from positive blood culture bottles by use of matrix-assisted laser desorption-ionization time of flight mass spectrometry fingerprinting. *J Clin Microbiol* 2010; **48**: 1584–1591.
155. Schubert S, Weinert K, Wagner C, Gunzl B, Wieser A, Maier T, et al. Novel, improved sample preparation for rapid, direct identification from positive blood cultures using matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry. *J Mol Diagnostics* 2011; **13**: 701–706.
156. Martiny D, Dediste A, Vandenberg O, Vandenberg O. Comparison of an in-house method and the commercial Sepsityper™ kit for bacterial identification directly from positive blood culture broths by matrix-assisted laser desorption-ionisation time-of-flight mass spectrometry. *Eur J Clin Microbiol Infect Dis* 2012; **31**: 2269–2281.

157. Ko KS, Lee HK, Park MY, Lee KH, Yun YJ, Woo SY, et al. Application of RNA polymerase β -subunit gene (*rpoB*) sequences for the molecular differentiation of *Legionella* species. *J Clin Microbiol* 2002; **40**: 2653–2658.
158. Mentasti M, Fry NK, Afshar B, Palepou-Foxley C, Naik FC, Harrison TG. Application of *Legionella pneumophila*-specific quantitative real-time PCR combined with direct amplification and sequence-based typing in the diagnosis and epidemiological investigation of Legionnaires' disease. *Eur J Clin Microbiol Infect Dis* 2012; **31**: 2017–2028.
159. Adékambi T, Drancourt M, Raoult D. The *rpoB* gene as a tool for clinical microbiologists. *Trends Microbiol* 2009; **17**: 37–45.
160. Edgar RC. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004; **32**: 1792–1797.
161. Price MN, Dehal PS, Arkin AP. Fasttree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 2009; **26**: 1641–1650.
162. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst Biol* 2010; **59**: 307–321.
163. Quijada NM, Rodríguez-Lázaro D, Eiros JM, Hernández M. TORMES: an automated pipeline for whole bacterial genome analysis. *Bioinformatics* 2019; **35**: 4207–4212.
164. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012; **19**: 455–477.
165. Chen KT, Liu CL, Huang SH, Shen HT, Shieh YK, Chiu HT, et al. CSAR: A contig scaffolding tool using algebraic rearrangements. *Bioinformatics* 2018; **34**: 109–111.
166. Hasman H, Saputra D, Sicheritz-Ponten T, Lund O, Svendsen CA, Frimodt-Møller N, et al. Rapid whole-genome sequencing for detection and characterization of microorganisms directly from clinical samples. *J Clin Microbiol* 2014; **52**: 139–146.
167. Larsen M V., Cosentino S, Lukjancenko O, Saputra D, Rasmussen S, Hasman H, et al. Benchmarking of methods for genomic taxonomy. *J Clin Microbiol* 2014; **52**: 1529–1539.
168. Clausen PTL, Aarestrup FM, Lund O. Rapid and precise alignment of raw reads against redundant databases with KMA. *BMC Bioinformatics* 2018; **19**: 1–8.
169. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 2012; **28**: 1647–1649.
170. Seppey M, Manni M ZE. BUSCO: assessing genome assembly and annotation completeness. *Methods Mol Biol* 2019; **1962**: 227–245.
171. Tatusova T, Dicuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, et al. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 2016; **44**: 6614–6624.
172. Yoon SH, Ha S min, Lim J, Kwon S, Chun J. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek, Int J Gen Mol Microbiol* 2017; **110**: 1281–1286.
173. 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.
174. Tanizawa Y, Fujisawa T, Nakamura Y. DFAST: A flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* 2018; **34**: 1037–1039.

175. 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.
176. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: Architecture and applications. *BMC Bioinformatics* 2009; **10**: 1–9.
177. 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**.
178. 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.
179. EUCAST. Antimicrobial Susceptibility Testing of *Legionella pneumophila*. 2021. Vaxjo, Sweden; 2021.

Chapter 1

Evaluation of MALDI–TOF Mass Spectrometry in Diagnostic and Environmental Surveillance of Legionella Species: A Comparison With Culture and Mip-Gene Sequencing Technique

Chapter published as:

Pascale MR, Mazzotta M, Salaris S, Girolamini L, Grottola A, Simone ML, Cordovana M, Bisognin F, Dal Monte P, Bucci Sabattini MA, Viggiani M and Cristino S (2020) Evaluation of MALDI–TOF Mass Spectrometry in Diagnostic and Environmental Surveillance of Legionella Species: A Comparison With Culture and Mip-Gene Sequencing Technique. Front. Microbiol. 11:589369. doi: 10.3389/fmicb.2020.589369

Abstract

Legionella spp. are widespread bacteria in aquatic environments with a growing impact on human health. Between the 61 species, *Legionella pneumophila* is the most prevalent in human diseases; on the contrary, *Legionella non-pneumophila* species are less detected in clinical diagnosis or during environmental surveillance due to their slow growth in culture and the absence of specific and rapid diagnostic/analytical tools. Reliable and rapid isolate identification is essential to estimate the source of infection, to undertake containment measures, and to determine clinical treatment. Matrix-assisted laser desorption ionization–time-of-flight mass spectrometry (MALDI–TOF MS), since its introduction into the routine diagnostics of laboratories, represents a widely accepted method for the identification of different bacteria species, described in a few studies on the *Legionella* clinical and environmental surveillance. The focus of this study was the improvement of MALDI–TOF MS on *Legionella non-pneumophila* species collected during *Legionella* nosocomial and community surveillance. Comparative analysis with cultural and *mip*-gene sequencing results was performed. Moreover, a phylogenetic analysis was carried out to estimate the correlations amongst isolates. MALDI–TOF MS achieved correct species-level identification for 45.0% of the isolates belonging to the *Legionella anisa*, *Legionella rubrilucens*, *Legionella feeleeii*, and *Legionella jordanis* species, displaying a high concordance with the *mip*-gene sequencing results. In contrast, less reliable identification was found for the remaining 55.0% of the isolates, corresponding to the samples belonging to species not yet included in the database.

The phylogenetic analysis showed relevant differences inside the species, regrouped in three main clades; among the *Legionella anisa* clade, a subclade with a divergence of 3.3% from the main clade was observed. Moreover, one isolate, identified as *Legionella quinlivanii*, displayed a divergence of 3.8% from the corresponding reference strain. However, these findings require supplementary investigation. The results encourage the implementation of MALDI–TOF MS in routine diagnostics and environmental *Legionella* surveillance, as it displays a reliable and faster identification at the species level, as well as the potential to identify species that are not yet included in the database. Moreover, phylogenetic analysis is a relevant approach to correlate the isolates and to track their spread, especially in unconventional reservoirs, where *Legionella* prevention is still underestimated.

Keywords

Legionella non-pneumophila species, *Legionella* identification, MALDI-TOF MS, MALDI Biotyper system, agglutination test, *mip*-gene sequencing

Introduction

Legionella spp. are Gram-negative bacteria ubiquitous in natural fresh water such as rivers, lakes, and thermal springs. They can also be found in moist soil and mud in which the association with amoebae, protozoa, and biofilm plays a key role in the life cycle of bacteria [1–3]. However, all cases, clusters, or outbreaks are linked to artificial freshwater environments, which represent a main reservoir of *Legionella* and a significant health risk from an epidemiological point of view since the main vehicle of microorganism diffusion and exposure to humans is the aerosol produced by some devices, such as condensers, showers, faucets, humidifiers, whirlpool baths, and medication nebulizer devices [4–6].

Legionella have been found in drinking water distribution systems, cooling towers, and water supply systems, where the water temperature is higher than the environment temperature, sustaining its growth [7]. In fact, *Legionella* grow and replicate between temperatures of 25 and 45°C, with optimal growth between 32 and 42°C [8]. In addition, the presence of biofilm in water distribution systems increases the risk of infection, due to the ability of *Legionella* spp. to obtain a high level of nutrients inside biofilm and protection from environmental stresses [9, 10]. The ability of *Legionella* to replicate inside protozoa cells allows bacteria to then also infect human cells, such as alveolar macrophages, causing the disease in humans known as Legionellosis—an aggressive form of pneumonia—and the Pontiac fever, a febrile and generally benign non-pulmonary disease form [5, 11, 12]. *Legionella pneumophila* (*L. pneumophila*) is the most common and the most studied *Legionella* pathogenic species in humans. This species is divided into 16 serogroups, and the majority of infection cases, confirmed by the use of specific diagnostic criteria (i.e., positive urinary antigen, antibody titer movement, and isolation of strain), are attributable to the *L. pneumophila* serogroup 1 (SG1), as confirmed by epidemiological data [13–17]. However, in addition to *L. pneumophila*, there are more than 60 known species [18], some of which are associated with human diseases. Therefore, the risk of infection related to *Legionella non-pneumophila* species [19–21], represents a serious problem in clinical settings (e.g., healthcare facilities, and hospitals), in addition to water for human consumption [21–23] and it should not be underestimated. Among them, *Legionella anisa* (*L. anisa*) is frequently isolated together with *L. pneumophila* in the hospital plumbing systems and can be used as a probable indicator of epidemic risk, beyond being associated with some cases of human infections [24–27]. Moreover, cases of endocarditis due to *L. anisa* and co-infection in HIV-associated pneumonia have been reported [28, 29]. In Australia and New

Zealand, *L. longbeachae* isolated from potting soil mixes, represents the main source of human infection with 30.4% of community-acquired Legionellosis [14, 30, 31]. Due to its abundance in the environment and the level of pathogenicity, different techniques are applied for the detection and identification of *Legionella* spp., each of which present both advantages and disadvantages. According to the Italian Guidelines for the prevention and control of Legionellosis [32], the culture technique represents the “gold standard” method for the isolation of *Legionella* strains, as well as for the typing, which is routinely performed by serological and molecular techniques such as the agglutination test, the direct fluorescent antibody (DFA) test, indirect immunofluorescent assay (IFA), sequence based typing (SBT), and amplification of the macrophage infectivity potentiator (*mip*) gene [33]. In detail during environmental investigations, the isolates from water samples using the culture technique are mainly identified by biochemical and serological tests such as the agglutination test. The agglutination test is routinely used to type *Legionella* isolates, but it is also associated with negative or ambiguous results, leading to inaccurate evaluations of *Legionella* [34]. Nevertheless, only an identification at the genus level can be achieved, and in most cases, the quality of the results depends on the experience of the laboratory staff. Since standard methods are lab-intensive, time-consuming, and prone to delivering false-negative results, several molecular techniques have been developed to detect *Legionella* [35, 36]. The European Working Group for *Legionella* Infection (EWGLI) developed the SBT approach for clinical and environmental *L. pneumophila* strain typing, currently, this technique is recognized as the gold standard for the genotyping of *L. pneumophila* strains [37, 38]. Regarding the typing of *Legionella* non-*pneumophila* species, Ratcliff et al. [39] presented a classification scheme for the *Legionella* genus based on the *mip* gene, encoding for a membrane protein referred to as the “macrophage infectivity potentiator,” that for its genetic stability and specificity is the most reliable and recommended method for identification at the species level [40–42]. Although sensitive and specific, molecular methods are expensive and require specialized laboratories and well-trained staff. Moreover, molecular techniques are affected by their inability to quantify the real risk to humans, as they do not allow for discrimination between dead and alive bacteria, and this leads to a lack of correlation between genomic units and bacterial loads [expressed in colony formant units (CFUs)]. Therefore, the culture-based approach is considered as a reference method in *Legionella* environmental surveillance [32, 43], although the need to support the culture technique with rapid and cost-effective methods to improve diagnosis and the adoption of the appropriate antibiotic treatment.

In recent years, matrix-assisted laser desorption ionization-time-of-flight mass spectrometry (MALDI–TOF MS) has emerged as an innovative, rapid, and inexpensive technique for species-level microbial identification through the analysis of ribosomal protein patterns. This technique has improved the routine practice of clinical microbiology laboratories, replacing most traditional biochemical or molecular techniques [44–46].

Although the use of MALDI–TOF MS is widespread across the world, few data have been published regarding its application for the identification of *Legionella* spp. in clinical and environmental samples [47–53]. Moreover, in Italy, to the best of our knowledge, MALDI–TOF MS is not widely used in the clinical or environmental laboratories where *Legionella* surveillance is carried out.

The aim of this study is to evaluate the MALDI Biotyper system for the identification of *Legionella* isolates in the environment. We focused our attention on *Legionella* non-*pneumophila* species isolated from different facilities (i.e., hospitals, healthcare settings, companies, and community areas). Moreover, some of these environments (e.g., homes, companies, and hotels) are supplied by water distribution systems that do not have a water safety plan that includes *Legionella* control measures (i.e., temperature controls, disinfection treatment, and maintenance water system programs), representing a risk of subsequent *Legionella* infections. The MALDI Biotyper system was compared to other identification methods, such as the agglutination test and *mip* gene sequencing, to evaluate its performance and its potential use in clinical and environmental *Legionella* surveillance.

Materials and methods

The isolates evaluated in this study were collected during the *Legionella* environmental surveillance programs of several municipal water distribution systems, supplying different environments usually associated with risk of *Legionella* infections such as hospitals, healthcare facilities, companies, and community areas (e.g., spas, private apartments, and hotels). The dataset used in the study is shown in Table 1.

Table 1 Distribution of isolates between man-made environments.

Number of isolates (n = 202)	Man-made environments
121	Hospitals
35	Private apartments
24	Hotels
16	Companies
3	Fitness centers
2	Wellness centers
1	Bath houses

Legionella Culture and Isolate Selection

Water sampling was performed according to UNI EN International Standard Organization (ISO) 19458:2006 [54] and Italian guidelines [32]. Two liters of a hot water sample were collected in sterile polytetrafluoroethylene (PTFE) bottles containing sodium thiosulphate solution (20 mg/L) and then stored at 4°C and processed within 24 h of collection. The isolation of *Legionella* was performed by the culture technique according to ISO 11731:2017 [55]. Briefly, 2 L of the sample was concentrated by filtration on a 0.22 mm polyethersulfone membrane (Sartorius, Bedford, MA, United States). Different aliquots (from 0.2 to 0.1 mL) of the untreated sample or the filtered, heated, and acid-treated sample were seeded on plates of the selective medium glycine-vancomycin–polymyxin B–cycloheximide (GVPC) (Thermo Fisher Diagnostic, Basingstoke, United Kingdom), and incubated at 35±2°C with 2.5% CO₂ for a maximum of 15 days. During the incubation period, the growth of *Legionella* was evaluated every 2 days, examining the plates for the presence of colonies with specific characteristics ascribable to *Legionella* spp. Suspected colonies were sub-cultured on buffered charcoal yeast extract (BCYE) agar supplemented with L-Cysteine (Cys +) and without L-

Cysteine (Cys-) supplementation (Thermo Fisher Diagnostics, Basingstoke, United Kingdom). Positive *Legionella* colonies grow on *Legionella* BCYE Cys + agar, but fail to grow on *Legionella* BCYE Cys- agar. Furthermore, to provide more evidence that the suspect colonies do not belong to the genus *Legionella*, they were isolated on a blood agar plate (Tryptone Soya Agar + 5% sheep blood) (Thermo Fisher Diagnostics, Basingstoke, United Kingdom): the colonies that grew on blood agar were considered as Cysteine non-dependent and were reported as non-*Legionella* spp.

Serological and biochemical typing

A total of 202 isolates of *Legionella* detected in the hot water sample growth on BCYE Cys +, without growth on blood agar and with or without a positive reaction to the latex agglutination test, were considered eligible for the study. All colonies grown on BCYE Cys +, identified presumably as *Legionella non-pneumophila* species, underwent serological typing by the latex agglutination test (*Legionella* latex test kit, Thermo Fisher Diagnostic, Basingstoke, United Kingdom). The test allowed to identify *L. pneumophila* SG1, *L. pneumophila* SG 2–14, and *Legionella non-pneumophila* species. Regarding the *Legionella non-pneumophila* group, the pool of antibody provided by the manufacturer recognized only a few species most commonly associated with clinical cases, such as *L. anisa*, *L. bozemanii* 1 and 2, *L. gormanii*, *L. longbeachae* 1 and 2, *L. dumoffii*, and *L. jordanis*. Three clinical isolates of *L. pneumophila* previously typed as SBT as sequence type 1 (ST1) were included as a positive control. Moreover, five isolates grown on GVPC medium, but which failed to grow on BCYE cys +, were inserted as the negative control. These five isolates were sub-cultured on tryptic soy agar (TSA) (Biolife, Milan, Italy) and the colonies were successively biochemically typed by the Remel RapID NF Plus system (Thermo Fisher Diagnostic) as *Brevundimonas diminuta* (n = 1), *Acinetobacter junii* (n = 2), and *Flavobacterium lindanitolerans* (n = 1), and by Rapid ANA II as *Streptococcus sanguinis* (n = 1), according to the manufacturer's instructions. To confirm the results obtained, the isolates were also analyzed by the MALDI Biotyper system.

Identification by Maldi–Tof MS

All 202 isolates grown on BCYE cys + with positive, negative, or ambiguous results of the *Legionella* agglutination test were analyzed by the MALDI Biotyper system (Bruker Daltonik GmbH, Bremen, Germany). Spectra acquisition and processing were performed using the Microflex LT mass spectrometer (2,000–20,000Da, linear positive mode) and the MALDI Biotyper Compass 4.1

software, whose library (version BDAL 7854) included the spectra of 39 *Legionella* strains, as shown in Table 2.

Table 2 Number of *Legionella* strains spectra included in matrix-assisted laser desorption ionization (MALDI) Biotyper software library.

<i>L. anisa</i>	<i>L. feeleeii</i>	<i>L. longbeachae</i>	<i>L. rubrilucens</i>
(n = 9)	(n = 8)	(n = 9)	(n = 1)
<i>L. beliardensis</i>	<i>L. geestiana</i>	<i>L. maceachernii</i>	<i>L. sainthelensi</i>
(n = 1)	(n = 1)	(n = 3)	(n = 3)
<i>L. birminghamensis</i>	<i>L. gormanii</i>	<i>L. micdadei</i>	<i>L. santicrucis</i>
(n = 6)	(n = 2)	(n = 5)	(n = 3)
<i>L. bozemanii</i>	<i>L. gratiana</i>	<i>L. moravica</i>	<i>L. tucsonensis</i>
(n = 10)	(n = 5)	(n = 1)	(n = 2)
<i>L. brunensis</i>	<i>L. hackeliae</i>	<i>L. oakridgensis</i>	<i>L. wadsworthii</i>
(n = 2)	(n = 1)	(n = 2)	(n = 1)
<i>L. cherrii</i>	<i>L. impletisoli</i>	<i>L. parisiensis</i>	<i>L. waltersii</i>
(n = 4)	(n = 1)	(n = 1)	(n = 1)
<i>L. cincinnatiensis</i>	<i>L. israelensis</i>	<i>L. pneumophila</i>	<i>L. worsleiensis</i>
(n = 2)	(n = 1)	(n = 3)	(n = 1)
<i>L. dresdenensis</i>	<i>L. jamestowniensis</i>	<i>L. pneumophila ssp fraseri</i>	<i>L. yabuuchiae</i>
(n = 1)	(n = 1)	(n = 4)	(n = 1)
<i>L. dumoffii</i>	<i>L. jordanis</i>	<i>L. pneumophila ssp pascullei</i>	<i>Legionella sp.</i>
(n = 7)	(n = 2)	(n = 2)	(n = 1)
<i>L. erythra</i>	<i>L. lansingensis</i>	<i>L. pneumophila ssp pneumophila</i>	
(n = 1)	(n = 1)	(n = 6)	

During the first part of the experiments a comparison between direct smear and full extraction methods was performed for a subset of strains (20/202, 10%), following manufacturer's instructions. No significant differences were found in term of identification and quality of spectra. Therefore, considering that *Legionella* is a Gram-negative bacterium, and the full extraction method is recommended for species with thick cell wall, such as Gram-positives, *Actinomycetales*, and fungi, the direct smear method was applied on the remain *Legionella* isolates. Briefly, a small amount of bacterial biomass, picked from a fresh plate culture (24–48 h of incubation) was spotted in duplicate

onto a MALDI Biotyper target, overlaid with 1 ml of MALDI Biotyper matrix solution, and allowed to air dry before the measurement. The identification result was considered reliable when the log (score) was ≥ 2.0 ("high confidence level") or between 1.7 and 1.99 ("low confidence level"). In addition, in case of log (score) results between 1.4 and 1.59, identification was considered reliable when the first four proposed results were identical [56, 57].

Identification was considered reliable also in those cases in which the difference in the log (score) values between the first and second-best matches was ≥ 0.3 [58]. In contrast, the presence of score of 0.00–1.69 indicated a nonreliable identification. Determination of the sensitivity, specificity, and confidence interval (CI) at the 95% level of significance was performed using GraphPad Prism software version 8.0.1 for Windows (GraphPad Software, San Diego, California, United States).

Identification of *Legionella* spp. By *mip*-gene sequencing

The DNA extraction was carried out using the InstaGene Purification Matrix (Bio-Rad, Hercules, CA), and DNA concentrations were determined using the Qubit fluorometer (Thermo Fisher Scientific, Paisley, United Kingdom). The PCR for all *Legionella non-pneumophila* species isolates was performed using the protocols for the *mip* gene suggested by EWGLI, as described by Ratcliff et al. [39]. The *mip*-gene amplification was carried out using degenerate primers and modified by M13 tailing to avoid noise in the DNA sequence [59]. The *mip*-gene amplification was carried out in a 50µL reaction containing DreamTaq Green PCR Master Mix 2× (Thermo Fisher Diagnostic Basingstoke, United Kingdom) and 40 picomoles of each primer; 100 ng of the DNA extracted from the presumptive colonies was added as template. Specifically, the *mip* amplicons were sequenced using tailed M13 forward and reverse primers (mip-595R-M13R caggaaacagctatgaccCATATGCAAGACCTGAGGGAAC and mip-74F-M13F tgtaaacgacggccagtGCTGCAACCGATGCCAC) to obtain a complete coverage of the sequenced region of interest [59]. Amplification was performed in a thermocycler under the following conditions: Pre-denaturation for 3 min at 96°C, then 35 cycles consisting of 1 min at 94°C for denaturation, 2 min at 58°C for annealing, and 2 min at 72°C for extension, followed by a final extension at 72°C for 5 min. The reaction mixtures were then held at 4°C. The PCR products were visualized by electrophoresis on 2% agarose gel and stained with ethidium bromide. Following purification, DNA was sequenced using BigDye Chemistry and analyzed on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA). Raw sequencing data were assembled using CLC Main Workbench 7.6.4 software. The sequences were compared to sequences deposited in the *Legionella mip*-gene sequence

database using a similarity analysis tool. EWGLI have established an accessible web database² that contains sequence data from described species and allows the identification of *Legionella* non-*pneumophila* species. The species-level identification was done on the basis of $\geq 98\%$ similarity to a sequence in the database [41]. For one strain with a sequence generically identified in the *mip*-gene sequence database as *Legionella* species (*L. spp.*), the sequence was also queried in the GenBank database using the basic local alignment search tool (BLAST), and it showed the best match with *Legionella* species H (*L. species H*) [39]. The 205 nucleotide *mip* sequences generated for this study were submitted to GenBank. The provided accession numbers were as follows: MW021138, MW052863-MW053066.

Phylogenetic and allelic diversity analysis

To estimate the relationship among the *Legionella* species found, a multiple sequence alignment (MSA) and a phylogenetic tree were performed on the 202 *mip*-gene sequences and the three positive controls. For each taxon, identified as previously described, the reference *mip* sequence of the correspondent American Type Culture Collection (ATCC) and Institute of Medical and Veterinary Science (IMVS) strain was retrieved and added to the analysis. When required, a manual editing was performed on the sequences, trimming them to the same length as the reference sequence. The nucleotide sequences were aligned by the multiple sequence comparison by log-expectation (MUSCLE) program [60], executed in Geneious Prime 2020.1.2³, retaining the default settings. The phylogenetic tree was built by the Geneious Tree Builder, using Tamura–Nei [61] as a genetic distant model and neighbor joining [62] as a tree building method, and then bootstrapped using 100 replicates.

² http://bioinformatics.phe.org.uk/cgi-bin/Legionella/mip/mip_id.cgi

³ <https://www.geneious.com>

Results

Legionella identification by the agglutination test

Overall, 34/202 (16.8%) isolates resulted positive and 158/202 (78.2%) negative in the agglutination test, while 10/202 (5.0%) isolates provided an ambiguous result. All of the strains that were included as controls delivered the expected results. Examples of the agglutination results are shown in Figure 1.

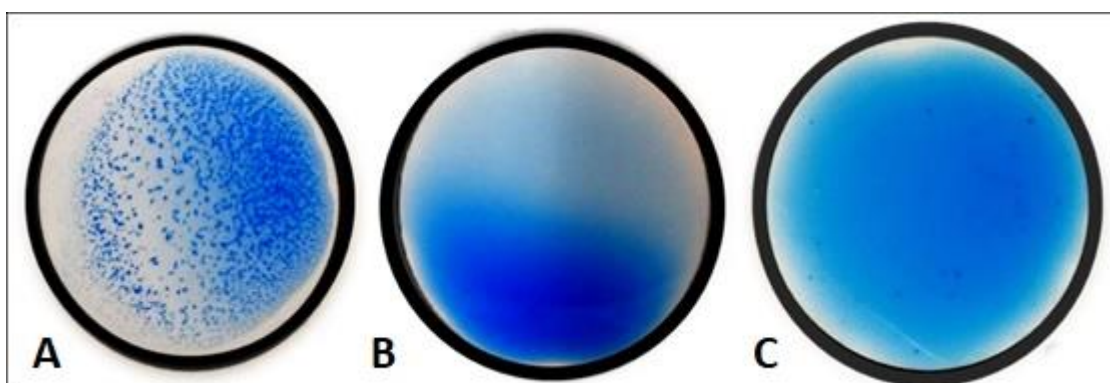


Figure 1 Positive (A), negative (B), and ambiguous (C) *Legionella non-pneumophila* species results of the agglutination test.

MALDI Biotyper system results

Applying the cut-offs recommended by the manufacturer, the MALDI Biotyper system identified at the genus level 90/202 (44.5%) isolates. Among them, 59/90 (65.5%) were identified at the species level with a high confidence level ($n = 50$ *L. anisa*; $n = 1$ *L. feeleii*; and $n = 8$ *L. rubrilucens*), and 31/90 (34.4%) were identified at the genus level with a low confidence level (*L. anisa*, *L. feeleii*, *L. jordanis*, and *L. rubrilucens*; Table 3); 112/202 isolates (56.0%) remain without identification.

Table 3 Isolates identification by MALDI Biotyper system according to the manufacturer's threshold and the revised sub-criteria.

Identification	Isolates	Manufacturer thresholds and number (n) of isolates			Total manufacturer identification (n)	*Sub-Criteria for score ≤ 1.69	Total (n)
		2.00-3.00	1.70-1.99	0.00-1.69*			
<i>Legionella</i>	<i>L. anisa</i>	50	2		52		52
	<i>L. feeleeii</i>	1	8		9	2*	11
	<i>L. jordanis</i>		1		1		1
	<i>L. rubrilucens</i>	8	20		28	38*	66
	Total	59	31		90	40*	130
Not Identified*				112*	112	72	72
Total					202		202

*[56-58]

The application of the additional interpretation criteria for identification results with a log (score) ≤ 1.69 enabled a further 40 strains to be identified from 112 un-identified isolates, achieving a total of 130/202 (64.4%) successful results. In contrast, 72 isolates remain unidentified. The summary of the MALDI Biotyper system results is shown in Table 3.

The isolates used as controls were identified with a high confidence level as *L. pneumophila* and other bacteria (i.e., *Acinetobacter junii*, *Brevundimonas dimunita*, *Flavobacterium lindanitolerans*, and *Streptococcus sanguinis*).

Comparison between the agglutination test and MALDI Biotyper system results

The 34 isolates with positive agglutination results were identified by the MALDI Biotyper system as *L. anisa*. The main isolates involved in the study (n = 158) displayed negative agglutination results, of which the MALDI Biotyper identified 89/158 (56.3%) of the isolates as *Legionella* genus, while the other 69/158 (43.7%) were not identified. Regarding the isolates with ambiguous agglutination results (n = 10), the MALDI Biotyper system correctly identified *Legionella* in seven cases (n = 4 *L. anisa* with a high score, and n = 3 *L. rubrilucens* with a revised low score); in contrast, 3/10 isolates were not identified. Full concordance between the two techniques was found for the control strains. However, the three positive controls identified by the agglutination reaction as *L. pneumophila* SG1, 3, and 6 were not distinguished at the serogroup level by the MALDI Biotyper system.

mip-gene sequencing

The results obtained by mip-gene sequencing with the respective range of matches to the reference strains (i.e., ATCC, IMVS-911, and IMVS-3376) are shown in Table 4.

Table 4 Results of the mip-gene sequencing and the number of isolates identified.

mip-gene sequencing results	Isolates identified n (%)	% of match with reference strains
<i>L. anisa</i>	50 (24.8)	100%
	2 (1.0)	96.7%
	1 (0.5)	99.4%
<i>L. feeleii</i>	1(0.5)	98.4%
	9 (4.5)	98.2%
<i>L. jordanis</i>	1 (0.5)	100%
<i>L. londiniensis</i>	7 (3.5)	100%
<i>L. nautarum</i>	15 (7.4)	100%
<i>L. quateirensis</i>	4 (1.9)	98.2%
<i>L. quinvilanii</i>	1 (0.5)	96.2%
<i>L. rubrilucens</i>	30 (14.9)	100%
<i>L. species H</i>	1 (0.5)	100%
<i>L. steelei</i>	1 (0.5)	99.8%
<i>L. tauriniensis</i>	79 (39.1)	100%
TOTAL	202 (100)	

The most common species detected by the sequencing analysis were *L. taurinensis* (39.1%), followed by *L. anisa* and *L. rubrilucens* (Table 4). Among the *L. anisa* subset, two isolates, identical to each other, were found to match the reference sequence strain at 96.7%, showing 20 mismatches with the main *L. anisa* group. Only one isolates generically identified as *L. spp.* In the *mip* gene database, were matched in GenBank and identified as *L. species H* [39]. The positive control was confirmed to belong to *L. pneumophila*; in contrast, the negative controls were not included in the *mip*-gene sequences analysis.

Comparison between the MALDI Biotyper system and mip-gene sequencing results

Among the species included in the MALDI Biotyper system database, the concordance between this system and the sequencing was 91/94 (96.9%), as follows: *L. anisa* (n = 52/52), *L. rubrilucens* (n = 27/30), *L. feeleii* (n = 11/11), and *L. jordanis* (n = 1/1). Therefore, among the species included in the MALDI Biotyper system database, the overall sensitivity of this system was 96.8% (95% CI: 0.91–

0.99). In the remaining 108 isolates identified by *mip*-gene sequencing as *L. londiniensis*, *L. nautarum*, *L. quateirensis*, *L. quinlivanii*, *L. species H*, *L. steelei*, and *L. taurinensis*, whose spectra are not included in the MALDI Biotyper system database, the MALDI Biotyper system did not identify 69 isolates and led to a misidentification in 39 cases (*L. rubrilucens* instead of *L. taurinensis*). Therefore, the specificity of the MALDI Biotyper system was 69/108 = 63.9% (95% CI: 0.55–0.72). A comparison between results delivered by the three techniques evaluated in this study is shown in Table 5.

Table 5 Comparison of the results of the three techniques.

mip-gene sequencing results	Latex agglutination test results			MALDI Biotyper system results		
	positive (+), negative (-), and ambiguous (±) results (n)					
Isolate identification (n)	+	-	±	2.00-3.00 (n)	1.70-1.99 (n)	0.00-1.69 *(n)
<i>L. anisa</i> (52)	34	14	4	<i>L. anisa</i> (50)	<i>L. anisa</i> (2)	
<i>L. feeleeii</i> (11)		11		<i>L. feeleeii</i> (1)	<i>L. feeleeii</i> (8)	<i>L. feeleeii</i> (2)
<i>L. jordanis</i> (1)		1			<i>L. jordanis</i> (1)	
<i>L. londiniensis</i> (7)		5	2			Not identified (7)
<i>L. nautarum</i> (15)		14	1			Not identified (15)
<i>L. quateirensis</i> (4)		4				Not identified (4)
<i>L. quinlivanii</i> (1)		1				Not identified (1)
<i>L. rubrilucens</i> (30)		30		<i>L. rubrilucens</i> (8)	<i>L. rubrilucens</i> (17)	<i>L. rubrilucens</i> (2) Not identified (3)
<i>L. species H</i> (1)		1				Not identified (1)
<i>L. steelei</i> (1)		1				Not identified (1)
<i>L. taurinensis</i> (79)		76	3			<i>L. rubrilucens</i> (36) Not identified (40)

* Manufacturer's criteria plus sub-criteria [56-58]. In bold type the concordance between techniques.

Phylogenetic analysis

The relationship between the 202 *Legionella* isolates, positive controls, and reference strains was studied by a phylogenetic tree of the sequenced *mip* gene in order to display a graphical representation of the inter- and intraspecific genetic variability among the taxa previously identified (Figure 2).

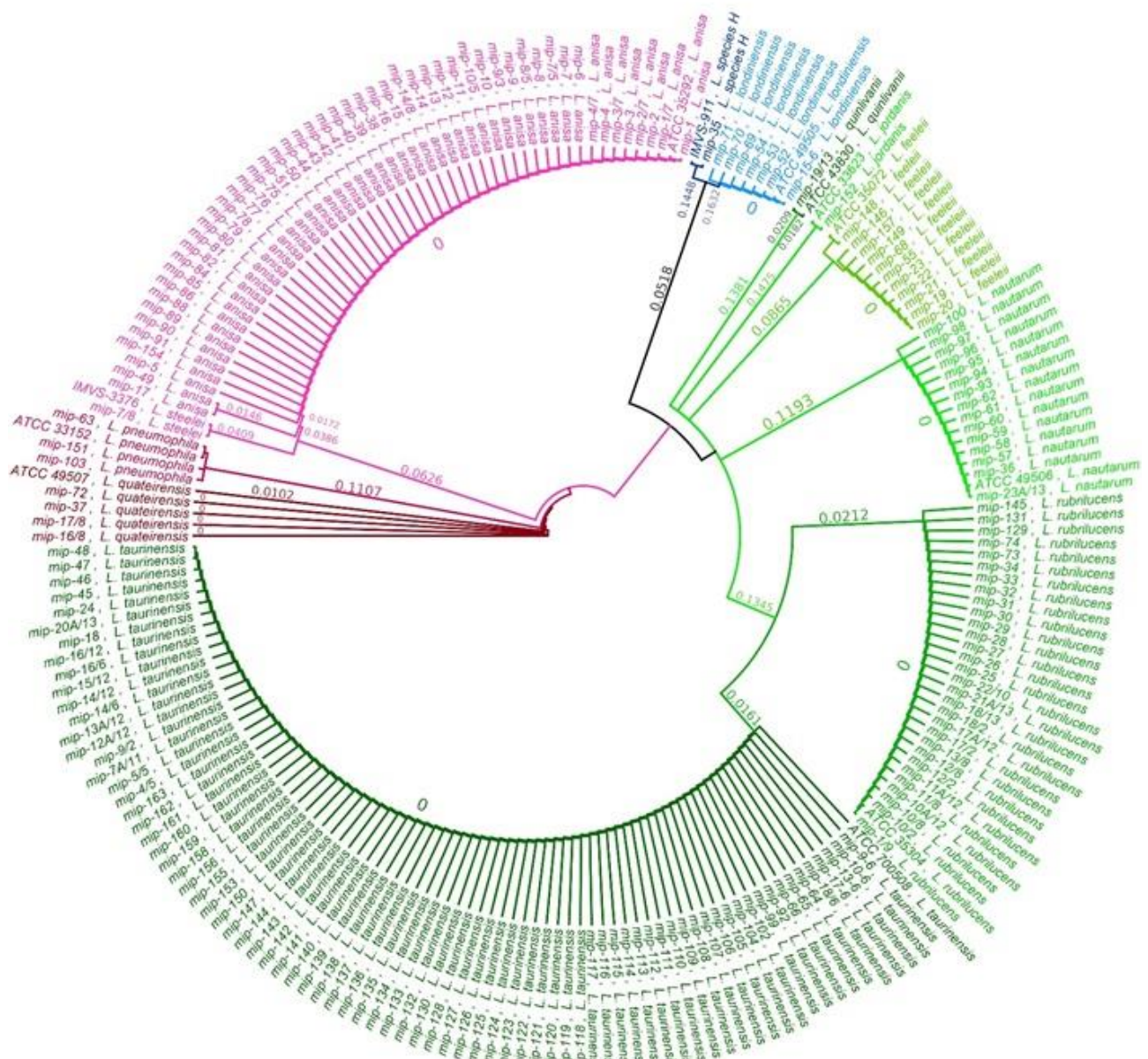


Figure 2 Phylogenetic tree obtained by the mip sequence analyses of isolates and strain types.

The phylogenetic tree shows that the clusters corresponding to the different *Legionella* species are clearly separate and distinct for each taxon, with every identified sequence falling on the same branch of the correspondent ATCC/IMVS reference mip gene. Therefore, our dataset can be separated into three major clades, in concordance with previous studies [39, 63]: One clade that includes *L. quinlivanii*, *L. jordani*, *L. feeleei*, *L. nautarum*, *L. rubrilucens*, and *L. taurinensis*; one clade represented by *L. anisa*, *L. steelei*, *L. pneumophila* (positive control), and *L. quateirensis*; and one that includes *L. londiniensis* and the so-called *L. species H*. Interestingly, the tree shows how *L.*

rubrilucens and *L. taurinensis* form a monophyletic group, revealing a high genetic similarity (96.3%). Concerning the intra-species analysis, the tree reveals the presence of a subclade inside the *L. anisa* principal clade, formed by two isolates sequences (i.e., *mip*-17 and *mip*-49), which share only 96.7% pairwise identity with the *L. anisa* reference *mip* sequence (i.e., ATCC 35392). The *mip* gene of the *L. quinlivanii* isolate shares only 96.2% sequence similarity with its correspondent reference strain (i.e., ATCC 43830).

Discussion

The role of *Legionella* in human respiratory infections emphasizes the importance of its surveillance in conventional and unconventional artificial environments that could represent a reservoir of infections [64]. One of the main concerns during environmental and clinical *Legionella* surveillance is the need of a rapid and sensitive technique that could improve therapeutic and epidemiological choices. First, it could allow the timely adoption of an appropriate antibiotic treatment, as well as the rapid identification of the source of infection by comparison between strains isolated from environmental and clinical samples, thus enabling the adoption of the correct prevention measures to control the infection. The long incubation time, the poor sensitivity of serological agglutination methods, and the long and laborious molecular techniques have long been the object of scientific discussion, suggesting the need to enhance *Legionella* identification with faster, cheaper, and more sensitive methods. MALDI–TOF MS is nowadays the universal method for microbial identification at the species level in routine microbial identification laboratories. In this study, a comparison between three different techniques (i.e., the agglutination test, *mip*-gene sequencing, and MALDI–TOF MS) was performed in order to assess whether MALDI–TOF MS technology could be a useful and valid tool to identify environmental *Legionella* species strains, reducing analytical times and costs. The strains included in this study were collected from different artificial water reservoirs, chosen considering the presence of a routine surveillance program (e.g., hospitals and healthcare facilities), and environments where a water safety plan is not required or is missed (e.g., homes, companies, and recreational communities). Starting from the cultural and diagnostic techniques commonly used on *L. pneumophila*, we evaluated the possibility of carrying out species-level identification using MALDI–TOF MS, including *Legionella* non-*pneumophila* isolates. Although some species are linked to the epidemiology of *Legionella* diseases, a standardized diagnostic approach for environmental and clinical samples has not yet been defined. These species are abundant in the environment, but are clinically less associated with human cases due to a lack of culture, a low growth rate, and the

poor sensitivity of diagnostic techniques [5, 13, 47]. All of these limits with regard to *Legionella non-pneumophila* species research led to a misidentification of strains and to an underestimation of the real risk represented by environments in which they can survive and proliferate. In our study, we reproduced routine laboratory workflows for *Legionella* surveillance, i.e., sampling, standardized culture, and biochemical–serological testing to elaborate an analytical report. Our culture results showed that most of the isolates grew after more than 10 days of incubation. Despite the reference technical guidelines provided by standard institutions, which suggest an incubation time of a minimum of 10 days, in routine laboratory practice, the time of incubation is typically no longer than 10 days, with a loss of isolate growth after this time [32, 55].

The growth of colonies on BCYE cys + is the only discriminating element for routine investigations: Isolates with positive growth are subjected to the latex agglutination test and, when agglutination fails, the results are reported in terms of counts of colonies (CFU/L) followed by a generic identification of “*Legionella non-pneumophila* species detection”. The misidentification of isolates leads to an underestimation of circulating species in the environment and, consequently, to inappropriate preventive measures to contain the risk. In the current research, the agglutination test returned positive results for only 34/202 (16.8%) isolates, failing to detect most of the *Legionella non-pneumophila* species isolates, i.e., 158/202 (78.2%). This result is related to limited equipment that does not recognize all of the *Legionella non-pneumophila* species; therefore, this assay is not suitable for discrimination within the *Legionella* genus, according to Orsini et al. [34], and highlights the need to establish, in routine analysis, a further technique able to accurately identify these *Legionella* species. The MALDI Biotyper system, following the manufacturer’s criteria, overall identified 90/202 (44.5%) *Legionella* isolates at the species level, corresponding to all of the species for which reference spectra are present in the database. A further significant number of isolates (n = 40) could have been identified by applying only the additional criteria described by Christner et al. [56]; Schubert et al. [57], and Martiny et al. [58]. The failed identification of the remaining 72 isolates is easily explained by the absence of the species in the MALDI Biotyper database [48, 49, 52]. The version of manufacturer’s database is closed and can be improved only by the manufacturer and it is covered by intellectual property rights. Therefore, the *Legionella* spectra identification produced in the study, could be represent our first “in-house database” for developing *Legionella* MALDI Biotyper application. These isolates were identified successively only by *mip*-gene sequencing as *L. taurinensis*, *L. londiniensis*, *L. nautarum*, *L. species H*, *L. quateirensis*, *L. quinlivanii*, and *L. steleei*.

Interesting results were observed among *L. tauriniensis*, as 39/79 isolates were misidentified (with a low confidence level) as *L. rubrilucens*, likely due to the close relatedness of these two species [65] [66]. In the phylogenetic tree, the two species are part of two distinct clades, although the differences between them only account for 3.7% (24 mismatches), regrouping them in the same monophyletic group. All *L. tauriniensis* isolates showed negative or ambiguous agglutination results and their identification was achieved only via the genotyping approach. The misclassification of *L. tauriniensis* with other *Legionella non-pneumophila* species (i.e., *L. rubrilucens*, *L. erythra*, and *L. spiritiensis*) occurred already in the first characterization of the species, performed on the *16srRNA* gene, that led to considering them as belonging to the same species [65, 66]. The subsequent characterization by the *mip* gene was able to distinguish between four different species. The MALDI–TOF MS approach, which establishes bacteria identification based on proteomic profiles, is likely unable to detect low differences at the genomic level, and thus returned no results. Different studies in the fields of bacteria identification suggest to combine MALDI Biotyper system analysis with Fourier-transform infrared spectroscopy (FTIR), which can be used to analyze carbohydrate and glycoproteins compounds, is able to provide a unique “fingerprint” spectrum for each species of bacteria exceeding the limit previously described [67, 68]. The three positive controls, belonging to different serogroups (i.e., 1, 3, and 6), were identified only at the species level without serogroup identification, confirming that the MALDI Biotyper system is unable to discriminate between *L. pneumophila* among serogroups. A concordance between the three methodologies included in this study was observed only for *L. anisa*. Interestingly, only two out of the 52 *L. anisa* isolates (i.e., *mip*-49 and *mip*-17) identified by the MALDI Biotyper system with a low confidence level corresponded to the sub-clade exhibiting a difference of 3.3% (20 mismatches) to the principal clade. We could speculate that these genetic differences can cause differences at the ribosomal protein level that could be underrepresented in the MALDI Biotyper system database, which includes nine *L. anisa* reference spectra. However, the differences found for the two sub-clades of *L. anisa* need further investigation. The comparison of the *mip*-gene sequencing of one isolate of *L. quinlivanii* showed a similarity of 95.91% to the reference strain ATCC 43830, with only 3.8% of differences linked to 23 mismatches. Considering the criteria adopted by Ratcliff to develop the *mip*-gene database [39], we found some discrepancies that could be improved by further investigation. Finally, we found a different isolate from all of the characterized species, matched to the *L.* species H described by Ratcliff et al. [39], close to the *L. londiniensis* clade, with differences of 34.7%. These findings will be improved by the whole-genome sequencing (WGS) approach in order to obtain more information

regarding the isolates' characteristics and their evolutionary adaptations in the environment. These findings show that the MALDI Biotyper system is a powerful tool for the identification of *Legionella* non-*pneumophila* species other than *L. pneumophila*, as demonstrated by its high sensitivity (96.8%) for the species included in the database and its low specificity (63.9%) that could be improved by amplifying the database with species missing or less represented (e.g., only one spectrum for *L. rubrilucens*). The limit of this study is linked to the need to improve the manufacturer's database and re-evaluation of the identification score criteria that could support the introduction of MALDI TOF MS into routine clinical and environmental *Legionella* surveillance. The re-evaluation of our "in-house database" with the new one from manufacturer represent the areas for future research. Moreover, the approach used in this study could develop the knowledge regarding the relationship between strains in water distribution systems, in addition to supporting the rapid identification of the source of infection, the match to clinical strains, and the adoption of corrective actions to limit the spread of bacteria and to control nosocomial and community infection.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI GenBank (accession numbers: MW021138 and MW052863–MW053066).

Author contributions

SC and MP conceived and designed the experiments and wrote the manuscript. MM and LG performed the samples collection and the experiments. SS performed the phylogenetic analysis. AG and MLS performed mip-gene sequencing. MC, PDM, and FB performed MALDI-TOF MS analysis and data interpretation. MBS and MV supply communities *Legionella* isolates. All authors contributed to the article and approved the submitted version.

Funding

This work was partially supported by the University of Bologna, Italy.

Acknowledgments

We would like to thank Graziella Ciani and Federica Marino for their support with the microbiological analyses.

References

1. Rowbotham TJ. Preliminary report on the pathogenicity of *Legionella pneumophila* for freshwater and soil amoebae. *J Clin Pathol* 1980; **33**: 1179–1183.
2. Fields BS. The molecular ecology of *legionellae*. *Trends Microbiol* 1996; **4**: 286–290.
3. Diederer BMW. *Legionella* spp. and Legionnaires' disease. *J Infect* 2008; **56**.
4. Hines SA, Chappie DJ, Lordo RA, Miller BD, Janke RJ, Lindquist HA, et al. Assessment of relative potential for *Legionella* species or surrogates inhalation exposure from common water uses. *Water Res* 2014; **56**: 203–213.
5. Mercante JW, Winchell JM. Current and emerging *Legionella* diagnostics for laboratory and outbreak investigations. *Clin Microbiol Rev* 2015; **28**: 95–133.
6. Carlson KM, Boczek LA, Chae S, Ryu H. Legionellosis and recent advances in technologies for *Legionella* control in premise plumbing systems: A review. *Water (Switzerland)* 2020; **12**: 1–22.
7. Diederer BMW. *Legionella* spp. and Legionnaires' disease. *J Infect* 2008; **56**: 1–12.
8. Arvand M, Jungkind K, Hack A. Contamination of the cold water distribution system of health care facilities by *Legionella pneumophila*: Do we know the true dimension? *Eurosurveillance* 2011; **16**: 1–6.
9. Abdel-Nour M, Duncan C, Low DE, Guyard C. Biofilms: The stronghold of *Legionella pneumophila*. *Int J Mol Sci* 2013; **14**: 21660–21675.
10. Di Pippo F, Di Gregorio L, Congestri R, Tandoi V, Rossetti S. Biofilm growth and control in cooling water industrial systems. *FEMS Microbiol Ecol* 2018; **94**: 1–13.
11. Newton HJ, Ang DKY, Van Driel IR, Hartland EL. Molecular pathogenesis of infections caused by *Legionella pneumophila*. *Clin Microbiol Rev* 2010; **23**: 274–298.
12. Cunha BA, Burillo A, Bouza E. Legionnaires' disease. *Lancet* 2016; **387**: 376–385.
13. Fields BS, Benson RF, Besser RE. *Legionella* and legionnaires' disease: 25 years of investigation. *J Clin Microbiol* 2002; **15**: 506–526.
14. Yu VL, Plouffe JF, Pastoris MC, Stout JE, Schousboe M, Widmer A, et al. Distribution of *Legionella* species and serogroups isolated by culture in patients with sporadic community-acquired legionellosis: An international collaborative survey. *J Infect Dis* 2002; **186**: 127–128.
15. Beauté J, Robesyn E. European Centre for Disease Prevention and Control - Legionnaires' disease in Europe 2015. 2017.
16. Rota, Maria Cristina, Caporali, Maria Grazia, Bella A, Scaturro, Maria, Giannitelli, Stefania, Ricci ML. Annual report of legionellosis in Italy - 2018. *Notiziario dell'Istituto Superiore di Sanità* . 2019. Rome, Italy.
17. ECDC: European Centre for Disease Prevention and Control. Legionnaires' disease. In: Annual Epidemiological Report for 2019. 2021. Stockholm.
18. LPSN Bacterio.net. List of prokaryotic names with standing in nomenclature *Legionella*.

<https://www.bacterio.net/genus/legionella>. .

19. Amemura-maekawa J, Kura F, Chida K, Ohya H, Kanatani J, Isobe J, et al. *Legionella pneumophila* and Other *Legionella* Species Isolated from Legionellosis Patients in Japan between 2008 and 2016. *Appl Environ Microbiol* 2018; **84**: 1–9.
20. Potts A, Donaghy M, Marley M, Othieno R, Stevenson J, Hyland J, et al. Cluster of Legionnaires' disease cases caused by *Legionella longbeachae* serogroup 1, Scotland, August to September 2013. *Eurosurveillance* 2013; **18**: 1–5.
21. Muder RR, Yu VL. Infection due to *Legionella* species other than *L. pneumophila*. *Clin Infect Dis* 2002; **35**: 990–998.
22. Bartram J, Chartier Y, Lee JV, Pond K, Surman-Lee S, WHO. *Legionella* and the Prevention of Legionellosis. *Emerg Infect Dis* 2007; **14**: 1006a – 1006.
23. Decker BK, Palmore TN. Hospital Water and Opportunities for Infection Prevention. *Curr Infect Dis Rep* 2014; **16**: 1–13.
24. Van Der Mee-Marquet N, Domelier AS, Arnault L, Bloc D, Laudat P, Hartemann P, et al. *Legionella anisa*, a possible indicator of water contamination by *Legionella pneumophila*. *J Clin Microbiol* 2006; **44**: 56–59.
25. Vaccaro L, Izquierdo F, Magnet A, Hurtado C, Salinas MA, Gomes TS, et al. First case of legionnaire's disease caused by *Legionella anisa* in Spain and the limitations on the diagnosis of *Legionella non-pneumophila* infections. *PLoS One* 2016; **11**: 1–12.
26. Fallon RJ, Stack BHR. Legionnaires' disease due to *Legionella anisa*. *J Infect* 1990; **20**: 227–229.
27. Fenstersheib MD, Miller M, Diggins C, Liska S, Detwiler L, Werner SB, et al. Outbreak of Pontiac fever due to *Legionella anisa*. *Lancet* 1990; **336**: 35–37.
28. Compain F, Bruneval P, Jarraud S, Perrot S, Aubert S, Napoly V, et al. Chronic endocarditis due to *Legionella anisa*: A first case difficult to diagnose. *New Microbes New Infect* 2015; **8**: 113–115.
29. Head BM, Trajtman A, Bernard K, Burdz T, Vélez L, Herrera M, et al. *Legionella* co-infection in HIV-associated pneumonia. *Diagn Microbiol Infect Dis* 2019; **95**: 71–76.
30. Steele TW, Lanser J, Sangster N. Isolation of *Legionella longbeachae* serogroup 1 from potting mixes. *Appl Environ Microbiol* 1990; **56**: 49–53.
31. Whiley H, Bentham R. *Legionella longbeachae* and legionellosis. *Emerg Infect Dis* 2011; **17**: 579–583.
32. Italian National Institute of Health. Guidelines for Prevention and Control of Legionellosis. Approvate in Conferenza Stato-Regioni Seduta Del 7 Maggio 2015. Italy, 2015. 2015.
33. Lück PC, Helbig JH, Schuppler M. Epidemiology and Laboratory Diagnosis of *Legionella* Infections. *LaboratoriumsMedizin* 2004; **26**: 174–182.
34. Orsini M, Cristino S, Grottola A, Romano-Spica V. Bacteria misagglutination in *Legionella* surveillance programmes. *J Hosp Infect* 2011; **79**: 179–180.

35. Templeton KE, Scheltinga SA, Sillekens P, Crielaard JW, Van Dam AP, Goossens H, et al. Development and clinical evaluation of an internally controlled, single-tube multiplex real-time PCR assay for detection of *Legionella pneumophila* and other *Legionella* species. *J Clin Microbiol* 2003; **41**: 4016–4021.
36. Blyth CC, Adams DN, Chen SCA. Diagnostic and typing methods for investigating *Legionella* infection. *N S W Public Health Bull* 2009; **20**: 157–161.
37. Gaia V, Fry NK, Afshar B, Luck PC, Meugnier H, Etienne J, et al. Consensus Sequence-Based Scheme for Epidemiological Typing of Clinical and Environmental Isolates of *Legionella pneumophila*. *J Clin Microbiol* 2005; **43**: 2047–2052.
38. Ratzow S, Gaia V, Helbig JH, Fry NK, Lück PC. Addition of neuA, the gene encoding N-acetylneuraminyl transferase, increases the discriminatory ability of the consensus sequence-based scheme for typing *Legionella pneumophila* serogroup 1 strains. *J Clin Microbiol* 2007; **45**: 1965–1968.
39. 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.
40. Ratcliff RM, Donnellan SC, Lanser JA, Manning PA, Heuzenroeder MW. Interspecies sequence differences in the mip protein from the genus *Legionella*: Implications for function and evolutionary relatedness. *Mol Microbiol* 1997; **25**: 1149–1158.
41. Fry NK, Afshar B, Bellamy W, Underwood AP, Ratcliff RM, Harrison TG, et al. Identification of *Legionella* spp. by 19 European reference laboratories: Results of the European Working Group for *Legionella* Infections External Quality Assessment Scheme using DNA sequencing of the macrophage infectivity potentiator gene and dedicated online tools. *Clin Microbiol Infect* 2007; **13**: 1119–1124.
42. Haroon A, Koide M, Higa F, Tateyama M, Fujita J. Identification of *Legionella pneumophila* serogroups and other *Legionella* species by mip gene sequencing. *J Infect Chemother* 2012; **18**: 276–281.
43. Lee J V., Lai S, Exner M, Lenz J, Gaia V, Casati S, et al. An international trial of quantitative PCR for monitoring *Legionella* in artificial water systems. *J Appl Microbiol* 2011; **110**: 1032–1044.
44. Maier T, Klepel S, Renner U, Kostrzewa M. Fast and reliable MALDI-TOF MS-based microorganism identification. *Nat Methods* 2006; **3**.
45. Croxatto A, Prod'homme G, Greub G. Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. *FEMS Microbiol Rev* 2012; **36**: 380–407.
46. Singhal N, Kumar M, Kanaujia PK, Viridi JS. MALDI-TOF mass spectrometry: An emerging technology for microbial identification and diagnosis. *Front Microbiol* 2015; **6**: 1–16.
47. Svarrer CW, Lück C, Elverdal PL, Uldum SA. Immunochromatic kits Xpect *Legionella* and BinaxNOW *Legionella* for detection of *Legionella pneumophila* urinary antigen have low sensitivities for the diagnosis of Legionnaires' disease. *J Med Microbiol* 2012; **61**: 213–217.
48. Moliner C, Ginevra C, Jarraud S, Flaudrops C, Bedotto M, Couderc C, et al. Rapid identification of *Legionella* species by mass spectrometry. *J Med Microbiol* 2010; **59**: 273–284.

49. Fujinami Y, Kikkawa HS, Kurosaki Y, Sakurada K, Yoshino M, Yasuda J. Rapid discrimination of *Legionella* by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Microbiol Res* 2011; **166**: 77–86.
50. Gaia V, Casati S, Tonolla M. Rapid identification of *Legionella* spp. by MALDI-TOF MS based protein mass fingerprinting. *Syst Appl Microbiol* 2011; **34**: 40–44.
51. He Y, Chang TC, Li H, Shi G, Tang YW. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry and database for identification of *legionella* species. *Can J Microbiol* 2011; **57**: 533–538.
52. Dilger T, Melzl H, Gessner A. Rapid and reliable identification of waterborne *Legionella* species by MALDI-TOF mass spectrometry. *J Microbiol Methods* 2016; **127**: 154–159.
53. Trnková K, Kotrbancová M, Špaleková M, Fulová M, Boledovičová J, Vesteg M. MALDI-TOF MS analysis as a useful tool for an identification of *Legionella pneumophila*, a facultatively pathogenic bacterium interacting with free-living amoebae: A case study from water supply system of hospitals in Bratislava (Slovakia). *Exp Parasitol* 2018; **184**: 97–102.
54. UNI EN ISO 19458:2006 - Water quality — Sampling for microbiological analysis. 2006. Geneva, Switzerland.
55. UNI EN ISO 11731:2017 Water quality - Enumeration of *Legionella*. 2017. Geneva, Switzerland.
56. Christner M, Rohde H, Wolters M, Sobottka I, Wegscheider K, Aepfelbacher M. Rapid identification of bacteria from positive blood culture bottles by use of matrix-assisted laser desorption-ionization time of flight mass spectrometry fingerprinting. *J Clin Microbiol* 2010; **48**: 1584–1591.
57. Schubert S, Weinert K, Wagner C, Gunzl B, Wieser A, Maier T, et al. Novel, improved sample preparation for rapid, direct identification from positive blood cultures using matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry. *J Mol Diagnostics* 2011; **13**: 701–706.
58. Martiny D, Dediste A, Vandenberg O, Vandenberg O. Comparison of an in-house method and the commercial Sepsityper™ kit for bacterial identification directly from positive blood culture broths by matrix-assisted laser desorption-ionisation time-of-flight mass spectrometry. *Eur J Clin Microbiol Infect Dis* 2012; **31**: 2269–2281.
59. Mentasti M, Fry NK, Afshar B, Palepou-Foxley C, Naik FC, Harrison TG. Application of *Legionella pneumophila*-specific quantitative real-time PCR combined with direct amplification and sequence-based typing in the diagnosis and epidemiological investigation of Legionnaires' disease. *Eur J Clin Microbiol Infect Dis* 2012; **31**: 2017–2028.
60. Edgar RC. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004; **32**: 1792–1797.
61. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 1993; **10**: 512–526.
62. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987; **4**: 406–425.

63. Burstein D, Amaro F, Zusman T, Lifshitz Z, Cohen O, Gilbert JA, et al. Genomic analysis of 38 *Legionella* species identifies large and diverse effector repertoires. *Nat Genet* 2016; **48**: 167–175.
64. Szewzyk U, Szewzyk R, Manz W, H. SK. MICROBIOLOGICAL SAFETY OF DRINKING WATER. *Annu Rev Microbiol* 2000; **54**: 81–127.
65. Lo Presti F, Riffard S, Meugnier H, Reyrolle M, Lasne Y, Grirnontf PAD, et al. *Legionella taurinensis* sp. nov., a new species antigenically similar to *Legionella spiritensis*. *Int J Syst Evol Microbiol* 1997; **49**: 397–403.
66. Baladron, B., Gil, V., and Pelaz C. "Serological versus sequence-based methods for Legionella identification," in *Legionella: State of the Art 30 Years After Its Recognition*, eds S. M. S. Cianciotto, Y. A. Kwaik, P. H. Edelstein, B. S. Fields, D. F. Geary, T. G. Harrison, et al. (Was. 2006).
67. Maity JP, Kar S, Lin CM, Chen CY, Chang YF, Jean JS, et al. Identification and discrimination of bacteria using Fourier transform infrared spectroscopy. *Spectrochim Acta - Part A Mol Biomol Spectrosc* 2013; **116**: 478–484.
68. Feng B, Shi H, Xu F, Hu F, He J, Yang H, et al. FTIR-assisted MALDI-TOF MS for the identification and typing of bacteria. *Anal Chim Acta* 2020; **1111**: 75–82.

Chapter 2

New Insight regarding Legionella Non-Pneumophila Species Identification: Comparison between the Traditional mip Gene Classification Scheme and a Newly Proposed Scheme Targeting the rpoB Gene

Chapter published as:

Pascale MR, Salaris S, Mazzotta M, Girolamini L, Fregni Serpini G, Manni L, Grottola A, Cristino S. New Insight regarding Legionella Non-Pneumophila Species Identification: Comparison between the Traditional mip Gene Classification Scheme and a Newly Proposed Scheme Targeting the rpoB Gene. Microbiol Spectr. 2021 Dec 22;9(3):e0116121. doi: 10.1128/Spectrum.01161-21. Epub 2021 Dec 15. PMID: 34908503; PMCID: PMC8672888.

Maria Rosaria Pascale and Silvano Salaris contributed equally to this article. Author order was determined in order of increasing seniority.

Abstract

The identification of *Legionella* non-pneumophila species (non-*Lp*) in clinical and environmental samples is based on the *mip* gene, although several studies suggest its limitations and the need to expand the classification scheme to include other genes.

In this study, the development of a new classification scheme targeting the *rpoB* gene is proposed to obtain a more reliable identification of 135 *Legionella* environmental isolates. All isolates were sequenced for the *mip* and *rpoB* genes, and the results were compared to study the discriminatory power of the proposed *rpoB* scheme. Complete concordance between the *mip* and *rpoB* results based on genomic percent identity was found for 121/135 (89.6%) isolates; in contrast, discordance was found for 14/135 (10.4%) isolates. Additionally, due to the lack of reference values for the *rpoB* gene, inter- and intraspecies variation intervals were calculated based on a pairwise identity matrix that was built using the entire *rpoB* gene (~4,107 bp) and a partial region (329 bp) to better evaluate the genomic identity obtained. The interspecies variation interval found here (4.9% to 26.7%) was then proposed as a useful sequence-based classification scheme for the identification of unknown non-*Lp* isolates. The results suggest that using both the *mip* and *rpoB* genes makes it possible to correctly discriminate between several species, allowing possible new species to be identified, as confirmed by preliminary whole-genome sequencing analyses performed on our isolates. Therefore, starting from a valid and reliable identification approach, the simultaneous use of *mip* and *rpoB* associated with other genes, as it occurs with the sequence-based typing (SBT) scheme developed for *Legionella pneumophila*, could support the development of multilocus sequence typing to improve the knowledge and discovery of *Legionella* species subtypes.

Importance

Legionella spp. are a widely spread bacteria that cause a fatal form of pneumonia. While traditional laboratory techniques have provided valuable systems for *Legionella pneumophila* identification, the amplification of the *mip* gene has been recognized as the only useful tool for *Legionella non-pneumophila* species identification both in clinical and environmental samples. Several studies focused on the *mip* gene classification scheme showed its limitations and the need to improve the classification scheme, including other genes. Our study provides significant advantages on *Legionella* identification, providing a reproducible new *rpoB* gene classification scheme that seems to be more accurate than *mip* gene sequencing, bringing out greater genetic variation on *Legionella* species. In addition, the combined use of both the *mip* and *rpoB* genes allowed us to identify presumed new *Legionella* species, improving epidemiological investigations and acquiring new understanding on *Legionella* fields.

Keywords

Legionella non-pneumophila species, *Legionella* identification, *mip* and *rpoB* gene sequencing, nucleotide identity percentage, interspecies and intraspecies variation interval, classification scheme, *Legionella* spp., whole-genome sequencing, environmental microbiology, gene classification scheme, genotyping, surveillance.

Introduction

Legionella spp. have been described as the causative agent of legionellosis in humans. The term refers to two main form of diseases, Legionnaires' disease, a severe multisystem disease involving pneumonia, and Pontiac fever, a non-pneumonic form, acute, self-limiting influenza-like illness. Additionally, extrapulmonary (e.g., sinusitis, hip wound infection, and prosthetic valve endocarditis) and asymptomatic forms are reported [1–3]. The *Legionella* genus includes Gram-negative aerobic bacteria widely found in both natural and artificial aquatic environments, where they can multiply inside free-living amoebae, protozoa, and biofilms, exploiting them as a source of nourishment and protection [4–6]. Currently, 66 *Legionella* species have been identified to date, and about half of them are linked with human infection, and some species contain more than one serogroup [7, 8]. The most studied species is *Legionella pneumophila* (*Lp*), which comprises 16 serogroups; the majority of cases, clusters, and outbreaks are attributable to serogroup 1 (Sg1). Other *Legionella* non-pneumophila species (non-*Lp*) are less studied and less commonly associated with human disease; thus, they remain undiagnosed due to limits of current diagnostic methods, which are more specific and sensitive for *Lp*. Indeed, the commonly used diagnostic method is the detection of a urinary antigen that is more sensitive for *Lp* Sg1 and does not permit the detection of *Lp* non-Sg1 or other *Legionella* species [9]. Among non-*Lp* species, *Legionella longbeachae* is the leading cause of infection in Australia and New Zealand, and potting soil mixes are considered the main source of infection [9, 10]. *Legionella anisa*, in addition to being isolated with *Lp*, is associated with several cases of legionellosis and coinfection [11–16], while *Legionella rubrilucens* was isolated from pneumonia patients coinfecting with *Lp* [17]. Considering the broad distribution of *Legionella* and the high incidence of disease, environmental *Legionella* surveillance is an important activity for preventing legionellosis [18]. Monitoring of several water sources (water distribution systems, cooling towers, fountains, spas, etc.) remains the main approach to prevent infection and to perform a fast identification of clusters and outbreaks that occur in community, hospital, and travel settings. Therefore, the possibility of rapidly identifying *Legionella* spp. with highly specific and sensitive methods represents one of the most important objectives for the control and prevention of *Legionella*. Over time, numerous methods have been developed for the detection, identification, and typing of *Legionella* spp. both in clinical and environmental samples. The culture of clinical and environmental samples is the gold standard for *Legionella* detection, and the sub-cultivation of isolated colonies on buffered charcoal yeast extract (BCYE) without L-cysteine (L-cys) is the first step

to discriminate *Legionella* from other bacteria. Serological methods, such as the agglutination test, the direct fluorescent antibody (DFA) test, and indirect immunofluorescence assay (IFA), are mostly used for discrimination between non-*Lp* species and *Lp* serogroups [19, 20]. Although these methods are commonly used, each of them shows different sensitivity and specificity and various error ranges; the culture technique is time consuming, technically difficult, and requires a long incubation time. However, serological methods lead to the occurrence of false-negative results and cross-reaction between species, limiting their specificity [1, 21]. To overcome these limitations, more rapid and precise identification of *Legionella* spp. can be provided by sequence analyses, which, as simple tools, can reduce the time needed for *Legionella* isolate identification with improved sensitivity and specificity [22, 23]. Currently, the gold standard for *Legionella* spp. typing is based on the approaches developed by the European Working Group for *Legionella* Infection (EWGLI) that are represented by a sequence-based typing (SBT) approach for clinical and environmental *Lp* strains [24, 25] and a database based on macrophage infectivity potentiator (*mip*) gene sequencing for non-*Lp* isolates [26, 27]. Currently, while for clinical and environmental *Lp* strains, a multilocus typing scheme has been developed by the EWGLI, represented by an SBT approach [24, 26], regarding the non-*Lp* isolates, identification is still based only on *mip* gene sequencing [26, 27], and no recognized and standardized typing approach was developed. Regarding the identification of *Legionella* species, several genetic markers have been proposed, including *16S rRNA*, which was subsequently replaced by the *mip* gene, as this gene can overcome the limitations of intraspecies heterogeneity in the *16S rRNA* gene [28]. However, some species and some environmental isolates could not be confidently discriminated by the *mip* scheme, such as *L. geestiana* or European wild strain LC4381 [29]. Another gene that is widely used for bacterial identification is the *rpoB* gene. This gene encodes a subunit of DNA-dependent RNA polymerase, and mutations in its sequence are known to cause rifampicin resistance. *rpoB* DNAs comprise a highly conserved region throughout bacteria that may be used for bacterial classification [30]. It can identify enteric bacteria, *Mycobacterium*, *spirochetes*, and *Legionella* species, including some causative agents of Legionnaires' disease [30, 31]. Regarding the identification of non-*Lp*, the nucleotide variation of *rpoB* is able to differentiate these species better than *16S rRNA* and *mip* in some cases [31]. The partial *rpoB* sequence (300 bp) can guarantee the genotypic classification of *Lp* and blue-white autofluorescent species [31]. This region can distinctly differentiate species that share high similarities in their *16S rRNA* gene sequences and that cannot be analyzed successfully by *mip* [26, 31]. Thus, *rpoB* analysis could clearly differentiate among *Legionella* spp. Although *rpoB*

has higher intraspecies variability, it is widely used for bacterial identification, and it is considered, in some cases, to be the best approach, such as for nontuberculous mycobacteria (NTM) and *Acinetobacter*. This marker is not sufficient for *Legionella* classification, especially for non-*Lp*, although different studies have already suggested combining *rpoB* with the *mip* gene to identify these species more accurately [32–34]. In addition, in the scientific literature, there are reference values for *mip* gene analysis that can be used to determine the inter- and intraspecies nucleotide variation; however, for *rpoB*, there are no works that establish reference intervals [26], and this limits the application of the *rpoB* gene as a marker in the classification scheme. In the present study, 135 *Legionella* spp. strains recovered from environmental communities were analyzed for *rpoB* gene sequencing, and the results obtained were compared with a *mip* gene sequencing identification scheme to study the discriminatory power of *rpoB* sequences and establish an inter- and intraspecies variation interval to improve the use of the *rpoB* gene as a target for non-*Lp* identification.

Results

All 135 isolates showed positive growth on BCYE cys+ and negative growth on BCYE cys- and tryptone soy agar (TSA) with 5% sheep blood agar. Moreover, the agglutination for *Legionella* species antisera test displayed positive results for 34/135 isolates (25.2%) and ambiguous results for 10/135 (7.4%) isolates; in contrast, most isolates (91/135 [67.4%]) showed negative results for the agglutination test. All of them were then submitted for gene amplification as previously described.

mip and *rpoB* sequencing results

All isolates (135/135 [100%]) were identified by *mip* and *rpoB* sequencing analysis at the species level as follows: *L. anisa* 51/135 (37.7%), *L. rubrilucens* 26/135 (19.2%), *L. taurinensis* 22/135 (16.3%), and *L. nautarum* 15/135 (11.1%). The remaining isolates were represented by *L. feeleii* 7/135 (5.2%), *L. londiniensis* 7/135 (5.2%), *L. quateirensis* 4/135 (3.0%), *L. quinlivanii* 1/135 (0.7%), and *L. steelei* 1/135 (0.7%). The positive control was confirmed to belong to *L. pneumophila*. Regarding the percentage of genomic identity, complete concordance between *mip* and *rpoB* results was found in 121/135 (89.6%) isolates; in contrast, discordance was returned for 14/135 (10.4%) isolates. In particular, it is possible to evaluate the number of isolates displaying concordance between the *mip* and *rpoB* results, including 49/135 belonging to *L. anisa*, 26/135 belonging to *L.*

rubrilucens, 15/135 belonging to *L. nautarum*, 7/135 belonging to *L. londiniensis*, 1/135 belonging to *L. steelei*, 22/135 belonging to *L. taurinensis*, and 1/135 belonging to *Lp*. The results obtained by *mip* and *rpoB* gene sequencing and their ranges of matches compared to the reference strains are shown in Table 1, where the 14 isolates with a discrepancy in the nucleotide identity percentage for *mip*, *rpoB*, or both genes are highlighted in bold.

Table 1 Comparison of *mip* and *rpoB* gene sequence results for match percentage (%), number of mismatches (mm), and interspecies identity and variation percentage with the respective reference strains

No. of isolates (n)	Reference strain	<i>mip</i> gene nucleotide identity (%), no. of mismatch (mm) and nucleotide variation (%)		<i>rpoB</i> gene nucleotide identity (%), no. of mismatches (mm) and nucleotide variation (%)	
		Interspecies identity interval ^a	Interspecies variation interval ^a	Interspecies identity interval ^b	Interspecies variation interval ^b
		69.5 – 96.4%	3.6 – 30.5%	73.3 – 95.1%	4.9 – 26.7%
		n = 48		n = 48	
		100.0%; 0 mm; 0%		100.0%; 0 mm; 0%	
		n = 2		n = 2	
<i>L. anisa</i> (51)	ATCC 35292	96.7%; 20 mm; 3.3%		92.4%; 25 mm; 7.6%	
		n = 1		n = 1	
		99.8%; 1 mm; 0.2%		99.4%; 2 mm; 0.6%	
<i>L. rubrilucens</i> (26)	ATCC 35304	100.0%; 0 mm; 0%		100.0%; 0 mm; 0%	
				n = 21	
<i>L. taurinensis</i> (22)	NCTC 13314	100.0%; 0 mm; 0%		100.0%; 0 mm; 0%	
				n = 1:	
				97.3%; 9 mm; 2.7%	
<i>L. nautarum</i> (15)	ATCC 49506	100.0%; 0 mm; 0%		100.0%; 0 mm; 0%	
				n. 6:	
<i>L. feeleeii</i> (7)	ATCC 35072	98.2%; 11 mm; 1.8%		95.4%; 15 mm; 4.6%	
				n. 1:	
				95.1%; 16 mm; 4.9%	
<i>L. londiniensis</i> (7)	ATCC 49505	100.0%; 0 mm; 0%		100.0%; 0 mm; 0%	
<i>L. quateirensis</i> (4)	ATCC 49507	98.2%; 11 mm; 1.8%		94.5%; 18 mm; 5.5%	
<i>L. quinlivanii</i> (1)	ATCC 43830	96.2%; 23 mm; 3.8%		95.7%; 14 mm; 4.3%	
<i>L. steelei</i> (1)	ATCC BAA2169	99.8%; 1 mm; 0.2%		100.0%; 0 mm; 0%	
POSITIVE CONTROL					
<i>L. pneumophila</i> (1)	ATCC 33152	99.0%; 6 mm; 1.0%		98.8%; 4 mm; 1.2%	

^a [26]

^b Based on the 329-bp matrix of the type strain (Fig. 2a and 2b)

Regarding *mip* gene identification, compared with the respective reference strains, our isolates showed a nucleotide identity interval of 98.2% to 100%, with the exception of two *L. anisa* isolates and one *L. quinlivanii* isolate with nucleotide identities of 96.7% and 96.2%, respectively. However, the *rpoB* gene results showed a nucleotide identity interval of 95.1% to 100%, except for two isolates of *L. anisa* and four isolates of *L. quateirensis*, which were identical to each other with nucleotide identity percentages of 92.4% and 94.5%, respectively. Moreover, for the seven isolates identified by the *mip* gene as *L. feeleii* (98.2%), a discrepancy with the *rpoB* gene identity results was found, showing a percentage of identity of 95.4% for six isolates and 95.1% for one isolate. To obtain a reliable identification scheme for the *rpoB* gene in our isolates, it was important to determine the specific intra- and interspecies variation intervals, as has been done for the *mip* sequence-based classification scheme created by Ratcliff et al. [26]. Therefore, our attention was focused mainly on the 14 isolates previously described as having higher discrepancies in nucleotide identity percentage. A pairwise identity matrix for the entire length of the *rpoB* gene based on 53 reference strains downloaded from NCBI, with a gene size from 4,101 to 4,143 bp, was built (Fig. 1a and b).

paal gear 4012 hp	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100																																																																																																																																								
1. adriatica	76.30	76.40	76.50	76.60	76.70	76.80	77.00	77.20	77.30	77.40	77.50	77.60	77.70	77.80	77.90	78.00	78.10	78.20	78.30	78.40	78.50	78.60	78.70	78.80	78.90	79.00	79.10	79.20	79.30	79.40	79.50	79.60	79.70	79.80	79.90	80.00	80.10	80.20	80.30	80.40	80.50	80.60	80.70	80.80	80.90	81.00	81.10	81.20	81.30	81.40	81.50	81.60	81.70	81.80	81.90	82.00	82.10	82.20	82.30	82.40	82.50	82.60	82.70	82.80	82.90	83.00	83.10	83.20	83.30	83.40	83.50	83.60	83.70	83.80	83.90	84.00	84.10	84.20	84.30	84.40	84.50	84.60	84.70	84.80	84.90	85.00	85.10	85.20	85.30	85.40	85.50	85.60	85.70	85.80	85.90	86.00	86.10	86.20	86.30	86.40	86.50	86.60	86.70	86.80	86.90	87.00	87.10	87.20	87.30	87.40	87.50	87.60	87.70	87.80	87.90	88.00	88.10	88.20	88.30	88.40	88.50	88.60	88.70	88.80	88.90	89.00	89.10	89.20	89.30	89.40	89.50	89.60	89.70	89.80	89.90	90.00	90.10	90.20	90.30	90.40	90.50	90.60	90.70	90.80	90.90	91.00	91.10	91.20	91.30	91.40	91.50	91.60	91.70	91.80	91.90	92.00	92.10	92.20	92.30	92.40	92.50	92.60	92.70	92.80	92.90	93.00	93.10	93.20	93.30	93.40	93.50	93.60	93.70	93.80	93.90	94.00	94.10	94.20	94.30	94.40	94.50	94.60	94.70	94.80	94.90	95.00	95.10	95.20	95.30	95.40	95.50	95.60	95.70	95.80	95.90	96.00	96.10	96.20	96.30	96.40	96.50	96.60	96.70	96.80	96.90	97.00	97.10	97.20	97.30	97.40	97.50	97.60	97.70	97.80	97.90	98.00	98.10	98.20	98.30	98.40	98.50	98.60	98.70	98.80	98.90	99.00	99.10	99.20	99.30	99.40	99.50	99.60	99.70	99.80	99.90	100.00
2. adriatica	76.30	76.40	76.50	76.60	76.70	76.80	77.00	77.20	77.30	77.40	77.50	77.60	77.70	77.80	77.90	78.00	78.10	78.20	78.30	78.40	78.50	78.60	78.70	78.80	78.90	79.00	79.10	79.20	79.30	79.40	79.50	79.60	79.70	79.80	79.90	80.00	80.10	80.20	80.30	80.40	80.50	80.60	80.70	80.80	80.90	81.00	81.10	81.20	81.30	81.40	81.50	81.60	81.70	81.80	81.90	82.00	82.10	82.20	82.30	82.40	82.50	82.60	82.70	82.80	82.90	83.00	83.10	83.20	83.30	83.40	83.50	83.60	83.70	83.80	83.90	84.00	84.10	84.20	84.30	84.40	84.50	84.60	84.70	84.80	84.90	85.00	85.10	85.20	85.30	85.40	85.50	85.60	85.70	85.80	85.90	86.00	86.10	86.20	86.30	86.40	86.50	86.60	86.70	86.80	86.90	87.00	87.10	87.20	87.30	87.40	87.50	87.60	87.70	87.80	87.90	88.00	88.10	88.20	88.30	88.40	88.50	88.60	88.70	88.80	88.90	89.00	89.10	89.20	89.30	89.40	89.50	89.60	89.70	89.80	89.90	90.00	90.10	90.20	90.30	90.40	90.50	90.60	90.70	90.80	90.90	91.00	91.10	91.20	91.30	91.40	91.50	91.60	91.70	91.80	91.90	92.00	92.10	92.20	92.30	92.40	92.50	92.60	92.70	92.80	92.90	93.00	93.10	93.20	93.30	93.40	93.50	93.60	93.70	93.80	93.90	94.00	94.10	94.20	94.30	94.40	94.50	94.60	94.70	94.80	94.90	95.00	95.10	95.20	95.30	95.40	95.50	95.60	95.70	95.80	95.90	96.00	96.10	96.20	96.30	96.40	96.50	96.60	96.70	96.80	96.90	97.00	97.10	97.20	97.30	97.40	97.50	97.60	97.70	97.80	97.90	98.00	98.10	98.20	98.30	98.40	98.50	98.60	98.70	98.80	98.90	99.00	99.10	99.20	99.30	99.40	99.50	99.60	99.70	99.80	99.90	100.00
3. adriatica	76.30	76.40	76.50	76.60	76.70	76.80	77.00	77.20	77.30	77.40	77.50	77.60	77.70	77.80	77.90	78.00	78.10	78.20	78.30	78.40	78.50	78.60	78.70	78.80	78.90	79.00	79.10	79.20	79.30	79.40	79.50	79.60	79.70	79.80	79.90	80.00	80.10	80.20	80.30	80.40	80.50	80.60	80.70	80.80	80.90	81.00	81.10	81.20	81.30	81.40	81.50	81.60	81.70	81.80	81.90	82.00	82.10	82.20	82.30	82.40	82.50	82.60	82.70	82.80	82.90	83.00	83.10	83.20	83.30	83.40	83.50	83.60	83.70	83.80	83.90	84.00	84.10	84.20	84.30	84.40	84.50	84.60	84.70	84.80	84.90	85.00	85.10	85.20	85.30	85.40	85.50	85.60	85.70	85.80	85.90	86.00	86.10	86.20	86.30	86.40	86.50	86.60	86.70	86.80	86.90	87.00	87.10	87.20	87.30	87.40	87.50	87.60	87.70	87.80	87.90	88.00	88.10	88.20	88.30	88.40	88.50	88.60	88.70	88.80	88.90	89.00	89.10	89.20	89.30	89.40	89.50	89.60	89.70	89.80	89.90	90.00	90.10	90.20	90.30	90.40	90.50	90.60	90.70	90.80	90.90	91.00	91.10	91.20	91.30	91.40	91.50	91.60	91.70	91.80	91.90	92.00	92.10	92.20	92.30	92.40	92.50	92.60	92.70	92.80	92.90	93.00	93.10	93.20	93.30	93.40	93.50	93.60	93.70	93.80	93.90	94.00	94.10	94.20	94.30	94.40	94.50	94.60	94.70	94.80	94.90	95.00	95.10	95.20	95.30	95.40	95.50	95.60	95.70	95.80	95.90	96.00	96.10	96.20	96.30	96.40	96.50	96.60	96.70	96.80	96.90	97.00	97.10	97.20	97.30	97.40	97.50	97.60	97.70	97.80	97.90	98.00	98.10	98.20	98.30	98.40	98.50	98.60	98.70	98.80	98.90	99.00	99.10	99.20	99.30	99.40	99.50	99.60	99.70	99.80	99.90	100.00
4. adriatica	76.30	76.40	76.50	76.60	76.70	76.80	77.00	77.20	77.30	77.40	77.50	77.60	77.70	77.80	77.90	78.00	78.10	78.20	78.30	78.40	78.50	78.60	78.70	78.80	78.90	79.00	79.10	79.20	79.30	79.40	79.50	79.60	79.70	79.80	79.90	80.00	80.10	80.20	80.30	80.40	80.50	80.60	80.70	80.80	80.90	81.00	81.10	81.20	81.30	81.40	81.50	81.60	81.70	81.80	81.90	82.00	82.10	82.20	82.30	82.40	82.50	82.60	82.70	82.80	82.90	83.00	83.10	83.20	83.30	83.40	83.50	83.60	83.70	83.80	83.90	84.00	84.10	84.20	84.30	84.40	84.50	84.60	84.70	84.80	84.90	85.00	85.10	85.20	85.30	85.40	85.50	85.60	85.70	85.80	85.90	86.00	86.10	86.20	86.30	86.40	86.50	86.60	86.70	86.80	86.90	87.00	87.10	87.20	87.30	87.40	87.50	87.60	87.70	87.80	87.90	88.00	88.10	88.20	88.30	88.40	88.50	88.60	88.70	88.80	88.90	89.00	89.10	89.20	89.30	89.40	89.50	89.60	89.70	89.80	89.90	90.00	90.10	90.20	90.30	90.40	90.50	90.60	90.70	90.80	90.90	91.00	91.10	91.20	91.30	91.40	91.50	91.60	91.70	91.80	91.90	92.00	92.10	92.20	92.30	92.40	92.50	92.60	92.70	92.80	92.90	93.00	93.10	93.20	93.30	93.40	93.50	93.60	93.70	93.80	93.90	94.00	94.10	94.20	94.30	94.40	94.50	94.60	94.70	94.80	94.90	95.00	95.10	95.20	95.30	95.40	95.50	95.60	95.70	95.80	95.90	96.00	96.10	96.20	96.30	96.40	96.50	96.60	96.70	96.80	96.90	97.00	97.10	97.20	97.30	97.40	97.50	97.60	97.70	97.80	97.90	98.00	98.10	98.20	98.30	98.40	98.50	98.60	98.70	98.80	98.90	99.00	99.10	99.20	99.30	99.40	99.50	99.60	99.70	99.80	99.90	100.00
5. adriatica	76.30	76.40	76.50	76.60	76.70	76.80	77.00	77.20	77.30	77.40	77.50	77.60	77.70	77.80	77.90	78.00	78.10	78.20	78.30	78.40	78.50	78.60	78.70	78.80	78.90	79.00	79.10	79.20	79.30	79.40	79.50	79.60	79.70	79.80	79.90	80.00	80.10	80.20	80.30	80.40	80.50	80.60	80.70	80.80	80.90	81.00	81.10	81.20	81.30	81.40	81.50	81.60	81.70	81.80	81.90	82.00	82.10	82.20	82.30	82.40	82.50	82.60	82.70	82.80	82.90	83.00	83.10	83.20	83.30	83.40	83.50	83.60	83.70	83.80	83.90	84.00	84.10	84.20	84.30	84.40	84.50	84.60	84.70	84.80	84.90	85.00	85.10	85.20	85.30	85.40	85.50	85.60	85.70	85.80	85.90	86.00	86.10	86.20	86.30	86.40	86.50	86.60	86.70	86.80	86.90	87.00	87.10	87.20	87.30	87.40	87.50	87.60	87.70	87.80	87.90	88.00	88.10	88.20	88.30	88.40	88.50	88.60	88.70	88.80	88.90	89.00	89.10	89.20	89.30	89.40	89.50	89.60	89.70	89.80	89.90	90.00	90.10	90.20	90.30	90.40	90.50	90.60	90.70	90.80	90.90	91.00	91.10	91.20	91.30	91.40	91.50	91.60	91.70	91.80	91.90	92.00	92.10	92.20	92.30	92.40	92.50	92.60	92.70	92.80	92.90	93.00	93.10	93.20	93.30	93.40	93.50	93.60	93.70	93.80	93.90	94.00	94.10	94.20	94.30	94.40	94.50	94.60	94.70	94.80	94.90	95.00	95.10	95.20	95.30	95.40	95.50	95.60	95.70	95.80	95.90	96.00	96.10	96.20	96.30	96.40	96.50	96.60	96.70	96.80	96.90	97.00	97.10	97.20	97.30	97.40	97.50	97.60	97.70	97.80	97.90	98.00																				

Fig 1 (a, b) Pairwise matrix developed using the entire *rpoB* gene size (4,101 to 4,143 bp) of 53 *Legionella* type strains used to determine the ranges of intra- and interspecies intervals of variation. The heatmap colors represent the range of similarity: from dark red (highest value) to dark blue (lowest value).

The matrix returned an interspecies pairwise identity interval of 72.7% to 95.0%. Therefore, the obtained interspecies variation interval was between 5.0% and 27.3%. The calculated intraspecies identity interval was 95.1% to 100%, resulting in an intraspecies variation interval between 0% and 4.9%, which permits the classification of unknown isolates as belonging to the same species. A second matrix was built considering only a 329-bp region of the *rpoB* gene (Fig. 2a and b) that was suggested by Ko et al. [31].

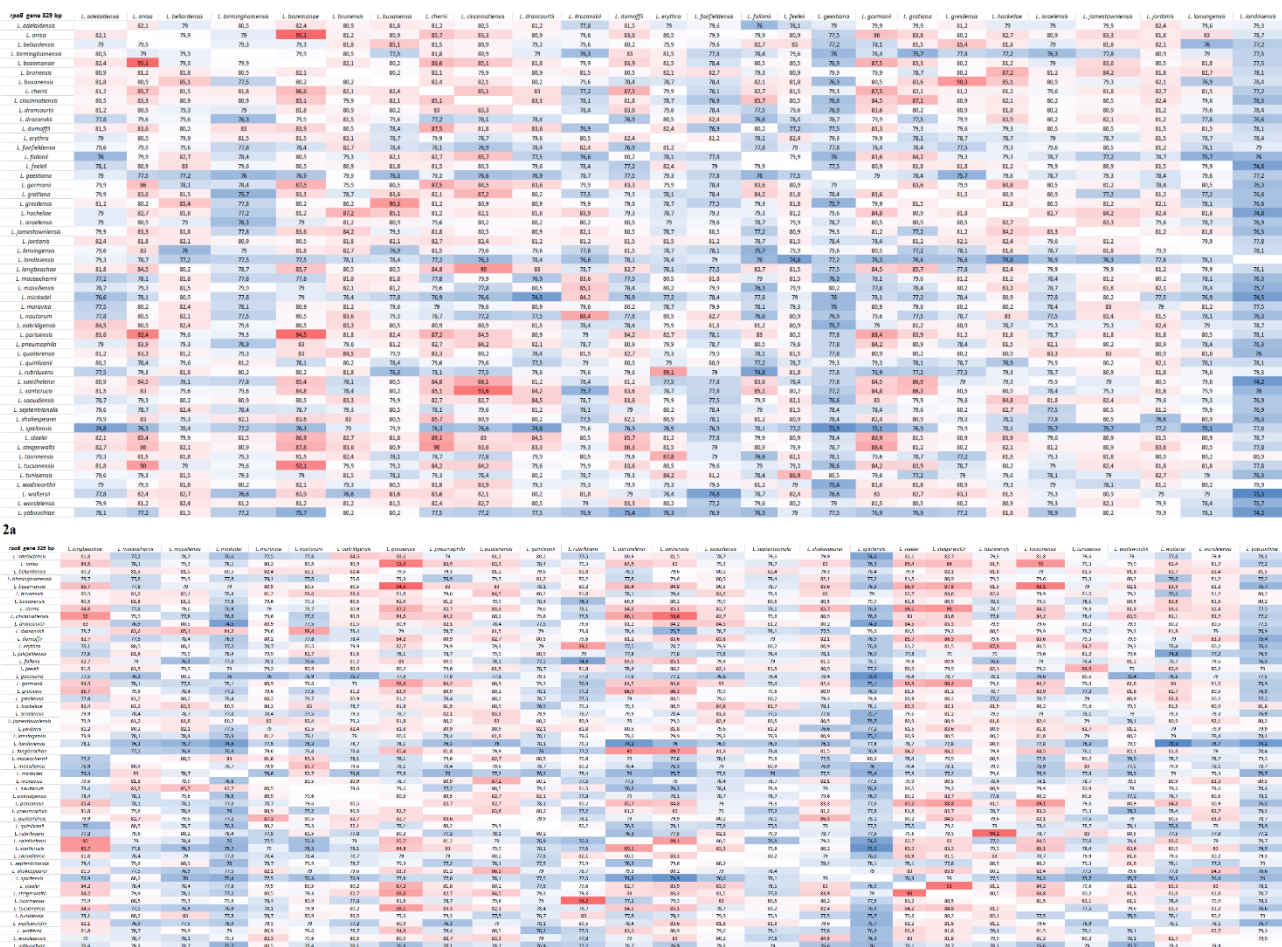


Fig 2 (a, b) Pairwise matrix developed using the selected region of the *rpoB* gene (329 bp) of 53 *Legionella* type strains used to determine the ranges of intra- and interspecies intervals of variation. The heatmap colors represent the range of similarity: from dark red (highest value) to dark blue (lowest value).

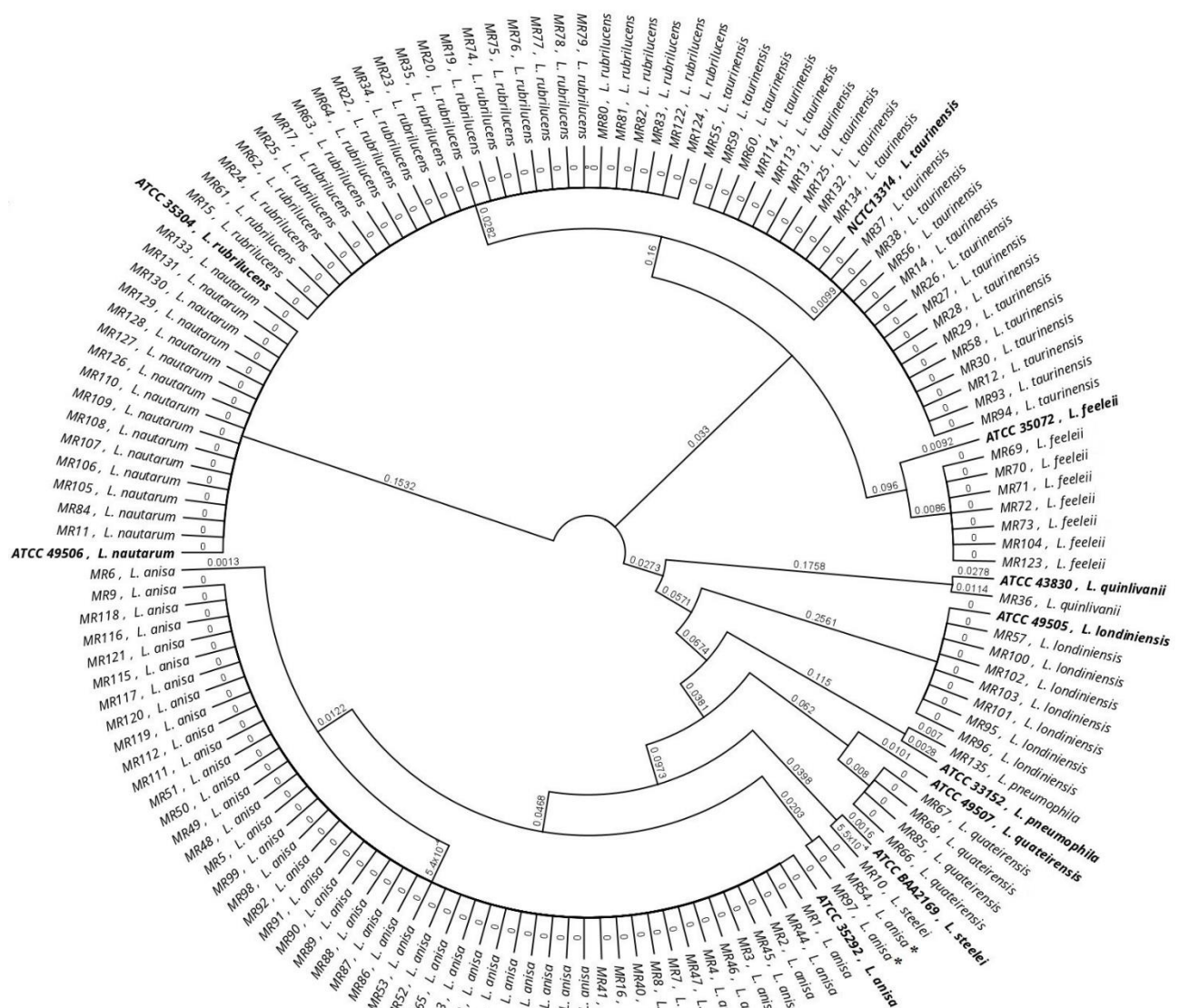
The matrix returned an interspecies pairwise identity interval of 73.3% to 95.1%. The interspecies variation interval was between 4.9% and 26.7%. The intraspecies identity interval determined was 95.2 to 100%, resulting in an intraspecies variation interval between 0% and 4.8%. As previously described, these values permit the identification of isolates as belonging to the same species. On the basis of the intra- and interspecies intervals calculated from the 329-bp *rpoB* gene region identity matrix, we analyzed the results for 14 isolates that showed discrepancies in *mip* and *rpoB* gene identification. The two *L. anisa* isolates determined according to the gold standard *mip* gene classification scheme were correctly identified; in contrast, the percentage of identity found for *rpoB* with respect to the reference strains (92.4%) did not fall within the intraspecies identity interval (lower cutoff at 95.1%), thus suggesting the possibility that the strains belong to different species. The same considerations can be applied to the four *L. quateirensis* isolates, which showed a percentage of identity for the *rpoB* gene of 94.5%. The identity values of seven strains of *L. feeleii*

and one isolate of *L. quinlivanii*, determined according to the *mip* gene classification scheme, showed borderline results with the *rpoB* classification scheme based on the observed cutoff values of 95.1% to 95.4% for the presumptive *L. feeleii* and 95.7% for *L. quinlivanii*. These findings provide further evidence of their misidentification and the necessity of further investigation. Moreover, Table 2 reports the nucleotide and amino acid differences in the wild strains with respect to the corresponding reference strains.

Table 2 *Legionella* species found during environmental surveillance with the number of nucleotide (DNA) and amino acid (AA) differences from the type strain in the *mip* and *rpoB* genes

<i>Legionella</i> species	GenBank accession no.	Culture collection and type strain	<i>Legionella</i> isolates ID	<i>mip</i> and <i>rpoB</i> sequence identification	No. of DNA and AA mismatches			
					<i>mip</i>		<i>rpoB</i>	
					DNA	AA	DNA	AA
<i>L. anisa</i>	LNXS01000032.1	ATCC 35292 and WA-316-C3	MR 1 – 5, MR 7 – 9, MR 16, MR 18, MR 21, MR 31 – 33, MR 39 – 53, MR 65, MR 86 – 92, MR 98 – 99, MR 111 – 112, MR 115–121	<i>L. anisa</i>	0	0	0	0
			MR 6	<i>L. anisa</i>	1	0	2	0
			MR 54, MR 97	<i>L. anisa</i>	20	1	25	0
<i>L. feeleeii</i>	NZ_LBHK01000054.1	ATCC 35072	MR 69 – 73, MR104	<i>L. feeleeii</i>	11	1	15	0
			MR 123	<i>L. feeleeii</i>	11	1	16	0
<i>L. londiniensis</i>	LNYS01000008.1	ATCC 49505	MR 57, MR 95 – 96, MR 100 – 103	<i>L. londiniensis</i>	0	0	0	0
<i>L. nautarum</i>	LNYS01000023.1	ATCC 49506	MR 11, MR 84, MR 105 – 110, MR 126 – 131, MR 133	<i>L. nautarum</i>	0	0	0	0
<i>L. quateirensis</i>	LNYS01000011.1	ATCC 49507	MR 66 – 68, MR 85	<i>L. quateirensis</i>	11	2	18	0
<i>L. quinlivanii</i>	LNYS01000014.1	ATCC 43830 and CDC#1442-AUS-E	MR 36	<i>L. quinlivanii</i>	23	3	14	0
<i>L. rubrilucens</i>	LNYS01000018.1	ATCC 35304 and WA-270A-C2	MR 15, MR 17, MR 19 – 20, MR 22 –25, MR 35 –35, MR 61 –64, MR 74 – 83, MR 122, MR 124	<i>L. rubrilucens</i>	0	0	0	0
<i>L. steelei</i>	LNYS01000006.1	ATCC BAA2169 and IMVS3376	MR 10	<i>L. steelei</i>	1	1	0	0
<i>L. taurinensis</i>	UGOZ01000001.1	NCTC:13314	MR 12, MR 14, MR 26 – 30, MR 37 – 38, MR 55 – 56, MR 58 – 60, MR 93 – 94, MR 113 – 114, MR 125, MR 132, MR 134	<i>L. taurinensis</i>	0	0	0	0
			MR 13	<i>L. taurinensis</i>	0	0	9	5
<i>L. pneumophila</i>	NC_002942.5	ATCC 33152 and Philadelphia 1	MR 135	<i>L. pneumophila</i>	6	0	4	0

Interestingly, it is possible to note that all the wild strains presented nucleotide differences in both genes. Despite the *rpoB* gene being characterized as having greater genetic variability (number of DNA mismatches), the deduced amino acid sequences of the *mip* gene showed a higher number of amino acid substitutions. It is important to emphasize that all 14 isolates focused on in our study showed few amino acid substitutions in the *mip* gene, from 1 to 3; in contrast, regarding the *rpoB* gene, only five amino acid substitutions were reported in *L. taurinensis*. Figures 3 and 4 display the relationship between all 135 isolates used in the study and the corresponding reference strains for the *mip* and *rpoB* genes, respectively.



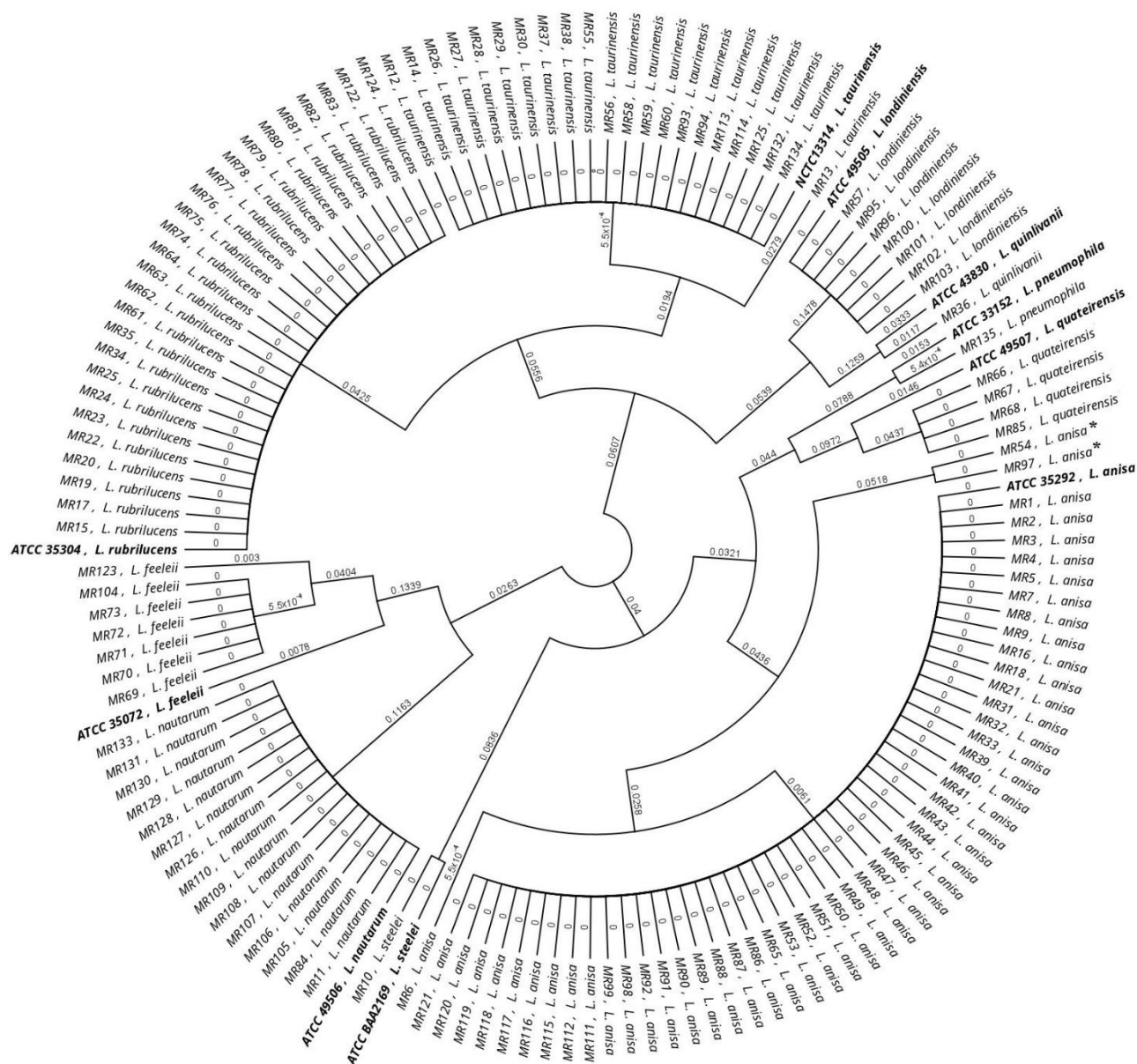


Fig 4 Analysis of the relationship among all 135 isolates and the respective type strain for the *rpoB* gene. Asterisks highlight the *L. anisa* isolates that diverge from the main branch. Reference strains are in bold.

The dendrogram built using the *mip* and *rpoB* gene sequences regrouped all isolates into 10 clades corresponding to a specific *Legionella* species. In the *mip* gene dendrogram, no relevant differences were found, with the exception of two isolates of *L. anisa* (MR 54 and MR 97) that were separated from the corresponding main branch, suggesting a possible misidentification of these isolates. In contrast, the dendrogram built using the *rpoB* gene showed the same 10 clades but with a higher genetic distance between wild types and the reference strains. Figures 5 and 6 show detailed results for *mip* and *rpoB* genetic discrepancies between the 14 wild strains and their reference strains used in this study.

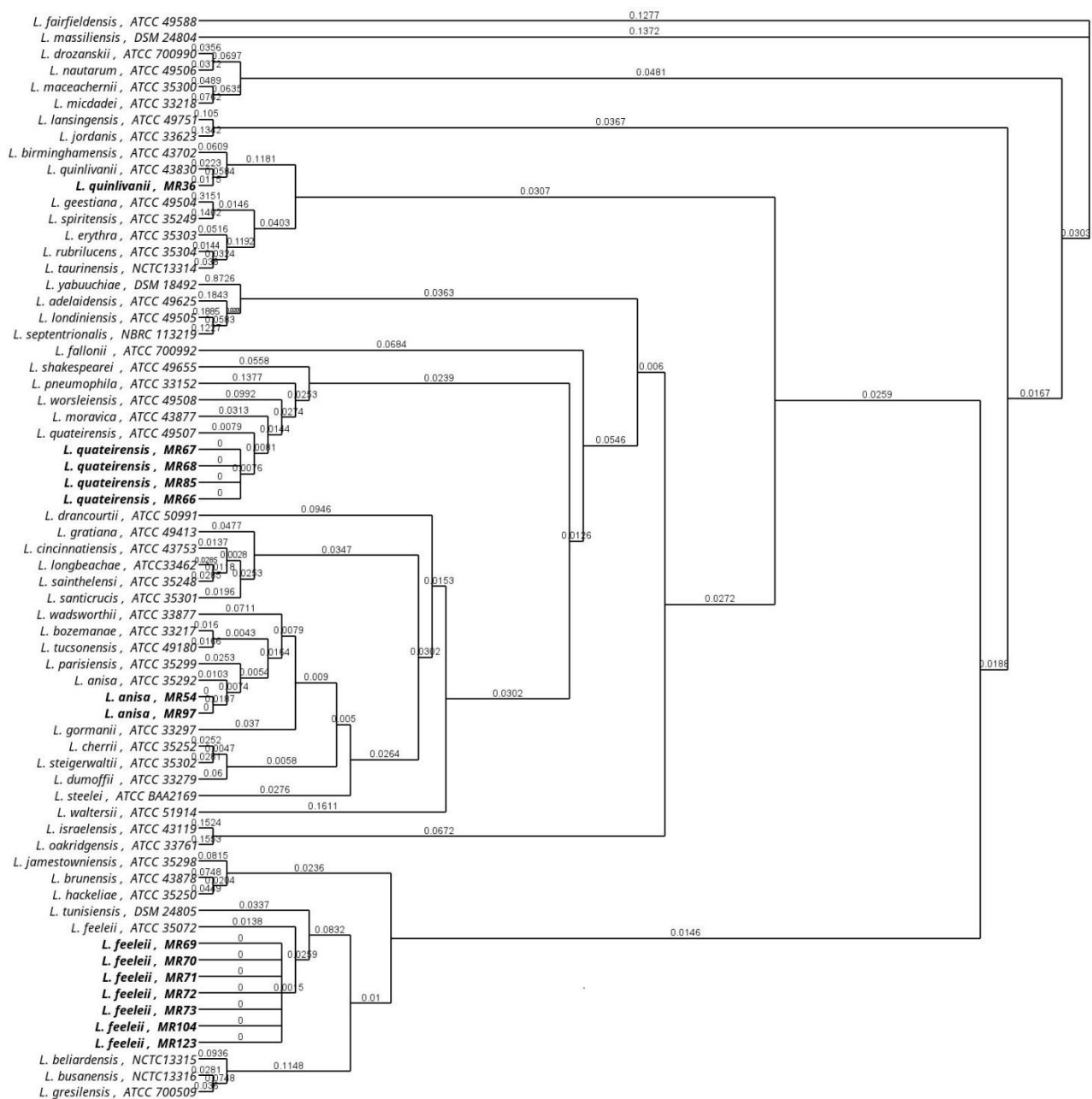


Fig 5 Phylogenetic dendrogram of *mip* *Legionella* sequences. Type strains versus wild strain isolates are in bold.

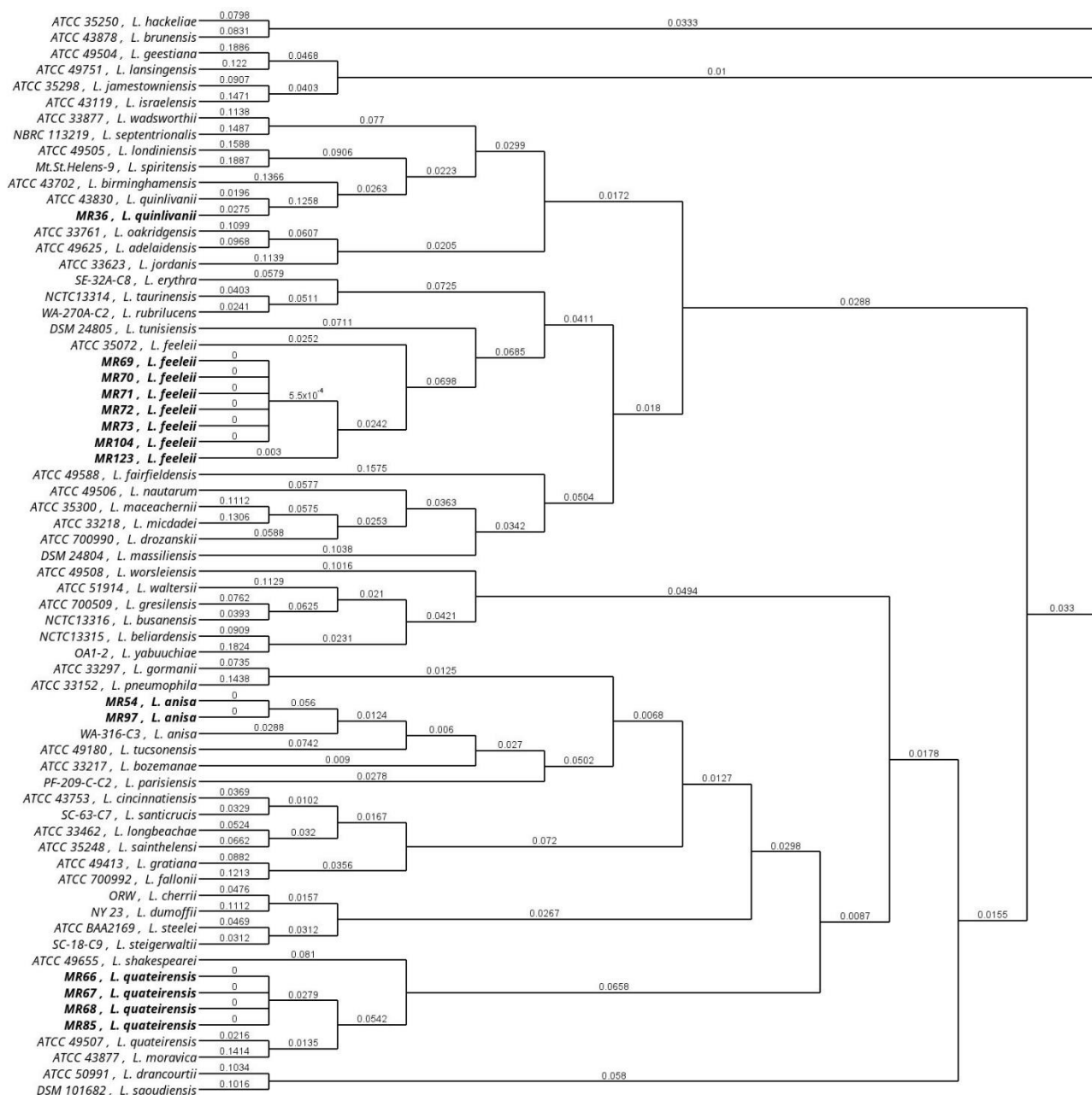


Fig 6 Phylogenetic dendrogram of *rpoB* *Legionella* sequences. Type strains versus wild strain isolates are in bold.

The *rpoB* dendrogram showed that in the main clade of *L. anisa*, two isolates (MR 54 and MR 97) were different from others based on the *mip* gene dendrogram that was previously described. Moreover, in the *L. feeleii* clade, one isolate (MR 123) is separated from the main clade, and one isolate belonging to *L. quinlivanii* (MR 36) and four isolates belonging to *L. quateirensis* (MR 66, MR 67, MR 68, and MR 85) present differences from the corresponding reference strains ATCC 43830 and ATCC 49507, respectively.

Discussion

Several studies have compared molecular methods to detect *Legionella* spp. in environmental and clinical samples, and it is well known that the amplification and sequencing of some genes for the direct detection and identification of bacteria can be simple, convenient, and specific in their differentiation of bacterial species. The use of PCR methods in *Legionella* identification and typing, thanks to their species-specific capability, has increased the power to detect and identify species, reducing the time and cost compared to culture and antibody approaches as well as improving the sensitivity and specificity of identification, especially for clinical approaches. Currently, non-*Lp* species have been mostly identified by only the *mip* gene, although several studies have shown that no single system is perfect and that other target genes need to be investigated [27]. The use of a particular region of the *rpoB* gene was already tested to determine phylogenetic relationships as well as the identification scheme for enteric bacteria, *Mycobacterium*, *Bartonella*, and other microorganisms [30, 35, 36]. Ko et al. have already shown that a partial region of *rpoB* is able to discriminate subspecies of *Lp* and several non-*Lp* species that have not been differentiated using the *mip* sequence classification scheme [37]. Many of the studies regarding the amplification of *rpoB* for *Legionella* spp. identification are exclusively focused on *Lp*, limiting the knowledge about the presence, distribution, and evolution of non-*Lp* species in the environment [37–39]. This study showed the steps needed to build a new classification scheme using the *rpoB* gene and its application to a great number of non-*Lp* isolates (n = 135) distributed in both nosocomial and community environments. The results obtained were compared with the gold standard *mip* gene classification scheme already developed by Ratcliff et al. [26] and still in use by the European Society of Clinical Microbiology and Infectious Diseases (ESCMID) Study Group for *Legionella* Infections (ESGLI). Our results confirmed, in agreement with previous studies, that both *mip* and *rpoB* are able to discriminate among *Legionella* species, considering that our isolates (89.6%) showed complete concordance between the two classification schemes. It is important to note that in some cases, there was no concordance between the *mip* and *rpoB* results, as there was a low percentage of genomic identity with respect to the reference strains for 14 isolates. In detail, our results suggest that sequencing using only *rpoB* is able to detect relevant genetic differences between the wild-type and the reference strains, which would otherwise be undetected using only the *mip* approach (e.g., *L. feeleeii*, *L. anisa*, *L. quinlivanii*, and *L. quateirensis*). This result is especially interesting given that *L. quateirensis* and *L. anisa* showed a variability percentage for the *rpoB* gene outside the intraspecies

interval of variation found here (0 to 4.8%); *L. feeleeii* and *L. quinlivanii* had values very close to the variation cutoff, suggesting that the identification scheme using only one gene limits the discovery and study of species variation and sometimes limits discrimination between different species. In line with previous results, all the dendrogram representations show that there is lower genetic diversity in the *mip* gene between and within the clades; in contrast, the diversity in *rpoB* appears to be greater, leading to the identification of several isolates that showed evident differences from their respective clade or reference strain (e.g., *L. anisa*, *L. feeleeii*, etc.). The results obtained using the *rpoB* gene seem to be useful for the identification of non-*Lp* species, and the results obtained permit the construction of the first *rpoB* gene classification scheme in the scientific literature. Thanks to the matrices described above, we built pairwise identity intervals that allowed us to classify our unknown sequences based on comparisons with reference strains. The comparison carried out using the values obtained here seems reliable, and we propose that they be used in a classification scheme. For strains whose similarity percentages are very close to the cutoff values, further in-depth analyses are recommended. Based on the intervals of variation derived from the pairwise identity matrices, the discriminatory power of the 329-bp target region for the non-*Lp* species appears to be as reliable as that of the entire gene. The comparison between the two matrices shows that the variability in the entire gene is greater than that in the selected region, suggesting that the analysis of a larger portion of the genome could increase the discriminatory power; therefore, approaches using new sequencing strategies, such as whole-genome sequencing (WGS), could contribute to better clarifying the identification of our isolates. This approach has already been applied to the four isolates of *L. quateirensis* described here. The average nucleotide identity (ANI) analysis, performed comparing their entire genome and the *L. quateirensis* type strain, showed pairwise values below the similarity threshold fixed to 95%, validating the hypothesis that the four strains belong to a presumptive new *Legionella* species [40]. In terms of the number of DNA and amino acid mismatches, most variability in the number of amino acid substitutions was observed in the *mip* gene, as all reported isolates showed discrepancies regarding the identification scheme based on the *mip* and *rpoB* genes. The role of the *mip* gene is widely documented; it is involved in the ability of *L. pneumophila* to replicate in eukaryotic cells and environmental amoebae [41]. The substitutions found could interfere with pathways influenced by *mip*, as documented for *Lp* as well as for some non-*Lp* species [42–44]. It is possible to observe that the *rpoB* gene displayed a high number of DNA mismatches with a low number of amino acid variations. This result could be

explained by the fact that *rpoB* is a housekeeping gene and that the alteration in the amino acid sequence could interfere with rifampicin resistance, as already demonstrated in other bacteria (e.g., *Mycobacterium tuberculosis*) and in a few *Legionella* species [39, 45]. Therefore, the five amino acid mismatches found in *L. taurinensis* indicate the need to study the role of these variations in terms of protein function. Further investigations on in silico protein modeling and structural prediction other than biochemical functionality studies might contribute to better clarifying the role of these amino acid alterations and their evolution in *Legionella* species. Although the non-*Lp* classification scheme using single-gene identification, such as the *mip* gene, is widely used and approved, the identification scheme for *Legionella* requires an update, such as introducing several patterns from various genes so as to increase the power of identification and improve phylogenetic studies. Especially for routine clinical and environmental laboratories where the whole-genome approach is expansive and laborious, the introduction of an easy, less expensive, and more sensitive scheme of identification could avoid errors in species characterization. Moreover, the proposed identification scheme could represent the first step toward acquiring information on different characteristics of isolates, such as changes in and development of antibiotics or disinfectant resistance, avoiding the failure of routine tests (e.g., urinary antigen test, serological and antibody-based assays), inadequate antibiotic treatments in human infection contest (e.g., rifampicin, fluoroquinolone, macrolides), and disinfection treatment. If a discrepancy is observed in this first step, then a more advanced technology, such as WGS, can be applied. This combined strategy represents an improved screening approach for *Legionella* isolate identification.

Materials and methods

The isolates involved in this study come from *Legionella* environmental surveillance programs of several facilities commonly associated with the risk of *Legionella* infections, including hospitals, companies, and communities (e.g., hotels, private apartments).

Legionella culture and isolate selection

The *Legionella* culture technique was based on ISO 11731:2017 [20]. The hot- and cold-water samples were sampled following the Italian National Unification and European Committee (UNI EN) ISO 19458:2006 [46] and Italian Guidelines [19]. Different aliquots (from 0.2 to 0.1 mL) of the untreated, filtered, heated, and acid-treated samples were seeded on plates of the selective medium glycine-vancomycin-polymyxin B-cycloheximide (GVPC) (Thermo Fisher Scientific, Diagnostic, Ltd., Basingstoke, UK) and incubated at $35 \pm 2^\circ\text{C}$ with 2.5% CO₂ for a maximum of 15 days. *Legionella* growth was evaluated every 2 or 3 days. To confirm the presence of the *Legionella* genus, suspected colonies were subcultured on buffered charcoal yeast extract (BCYE) agar with (cys) and without (cys) L-cysteine (L-cys) supplementation (Thermo Fisher Scientific, Diagnostic, Ltd., Basingstoke, UK). Moreover, as a negative control, the same isolates were spread on tryptone soy agar (TSA) with 5% sheep blood agar (Thermo Fisher Scientific, Diagnostic, Ltd., Basingstoke, UK) and incubated under the same conditions previously described, as *Legionella* is not able to grow on this medium. Only the colonies that grew on BCYE cys+ agar were considered for the next steps of the study.

Serological and biochemical typing

The predicted *Legionella* colonies were then identified using the *Legionella* latex agglutination test (Legionella latex test kit, Thermo Fisher Scientific, Diagnostic, Ltd., Basingstoke, UK), which is able to distinguish between *Lp* and non-*Lp*. In particular, among *Lp*, it is possible to identify serogroup 1 (Sg1) from Sg2 to Sg14, while among non-*Lp*, it is possible to recognize only some non-*Lp*, such as *L. anisa*, *L. bozemanii* 1 and 2, *L. gormanii*, *L. longbeachae* 1 and 2, *L. dumoffii*, and *L. jordanis*. A total of 134 strains of non-*Lp* and 1 strain of *Lp* that was previously typed by sequence-based typing (SBT) and included as a positive control were selected for the study.

Identification of *Legionella* spp. by *mip* and *rpoB* gene sequencing

The DNA of each strain was extracted using the InstaGene purification matrix (Bio-Rad, Hercules, CA), and DNA concentrations were determined using a Qubit fluorometer (Thermo Fisher Scientific, Paisley, UK). PCR analysis for all non-*Lp* isolates was performed to determine the gene sequences of *mip* and *rpoB* as described by Ratcliff et al. [26] and Ko et al. [31], respectively. *mip* gene amplification was performed using degenerate primers and modified by M13 tailing to avoid noise in the DNA sequence [47]. *mip* gene amplification was performed in a 50-μL reaction mixture containing DreamTaq Green PCR master mix 2× (Thermo Fisher Diagnostics, Basingstoke, UK) and 40 pmol of each primer; 100 ng of the DNA extracted from the presumptive colonies was added as the template. The *mip* amplicons were sequenced using tailed M13 forward and reverse primers (*mip*-595R-M13R caggaaacagctatgaccCATATGCAAGACCTGAGGGAAC and *mip*-74F-M13F tgtaaacgacggccagtGCTGCAACCG ATGCCAC) to obtain complete coverage of the region of interest [47]. Amplification was performed in a thermocycler under the following conditions: predenaturation for 3 min at 96°C, then 35 cycles consisting of 1 min at 94°C for denaturation, 2 min at 58°C for annealing, and 2 min at 72°C for extension, followed by a final extension at 72°C for 5 min. The reaction mixtures were then held at 4°C. *rpoB* gene amplification was performed as described by Ko et al. [31]. Gene amplification was performed in a 50-μL reaction volume containing 100 ng of template DNA, 40 pmol of each primer (RL1 59-GATGATATCGATCAYCTDGG-39; RL2 59-TTCVGGCGTTTCAATNGGAC-39), 1 U of Taq polymerase, and a PCR mixture consisting of PCR buffer 10×, 1.5 mM MgCl₂, and 250mM deoxynucleoside triphosphates (dNTPs). The thermal cycles consisted of 35 cycles, and each cycle consisted of 30 s at 94°C for denaturation, 30 s at 55°C for annealing, and 30 s at 72°C for extension, followed by a final extension at 72°C for 10 min. PCR products were visualized by electrophoresis on a 2% agarose gel and stained with ethidium bromide. Following purification, DNA was sequenced using BigDye chemistry and analyzed on an ABI PRISM 3100 genetic analyzer (Applied Biosystems, Foster City, CA). Raw sequencing data were assembled using CLC Main Workbench 7.6.4 software. The *mip* sequences were compared to sequences deposited in the Legionella *mip* gene sequence database using a similarity analysis tool. EWGLI has established an accessible web database (http://bioinformatics.phe.org.uk/cgi-bin/Legionella/mip/mip_id.cgi) that contains sequence data from described species and allows for the identification of non-*Lp* species. Species-level identification was performed on the basis of a similarity score of 98 to 100% compared to the sequences in the database [27] and considering the

intra- and interspecies intervals of variation previously described by Ratcliff et al. [26]. The *rpoB* sequences were compared to type strain sequences deposited in NCBI from several culture collections, including the American Type Culture Collection (ATCC), National Collection of Type Cultures, Central Public Health Laboratory (NCTC), NITE Biological Research Center, National Institute of Technology and Evaluation (NBRC), and Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSM). According to Adékambi et al. and Ko et al., the cutoff used for *rpoB* gene sequence-based identification was fixed at a 94 to 95% similarity percentage using a *rpoB* gene fragment of 300 to 600 bp [31, 48].

Elaboration of matrices for the *rpoB* gene: definition of the intra- and interspecies intervals of variation.

Legionella type strains (n = 53) retrieved from the NCBI, were used to determine the ranges of the intra- and interspecies intervals of variation for the *rpoB* gene, resulting in a pairwise identity matrix for the entire gene with a length from 4,101 to 4,143 bp (Fig. 1a and b) and for the 329-bp selected region (Fig. 2a and b), corresponding to the amplicon suggested by Ko et al. [31]. The list of type strains used in the study is reported in Table 3.

Table 3 NCBI type strains and wild strains used to build the pairwise identity matrix for intra- and interspecies interval determination in this study

<i>Legionella</i> species	GenBank accession number no.	Culture collection	Type strain
<i>L. adelaidensis</i>	LNKA01000005.1	ATCC:49625	1762-AUS-E
<i>L. anisa</i>	LNXS01000032.1	ATCC:35292	WA-316-C3
<i>L. beliardensis</i>	UGNV01000001.1	NCTC:13315	NCTC13315
<i>L. birminghamensis</i>	LNXT01000052.1	ATCC:43702	CDC#1407-AL-14
<i>L. bozemanæ</i>	NZ_LBAW01000041.1	ATCC:33217	ATCC 33217
<i>L. brunensis</i>	LN XV01000034.1	ATCC:43878	ATCC 43878
<i>L. busanensis</i>	UGOD01000001.1	NCTC:13316	NCTC13316
<i>L. cherrii</i>	LN XW01000014.1	ATCC:35252	ORW
<i>L. cincinnatiensis</i>	LN XX01000018.1	ATCC:43753	CDC#72-OH-14
<i>L. drancourtii</i>	NZ_JH413847.1	ATCC:50991	LLAP12
<i>L. drozanskii</i>	LN XY01000006.1	ATCC:700990	ATCC 700990
<i>L. dumoffii</i>	LN XZ01000001.1	ATCC:33279	NY 23
<i>L. erythra</i>	LN YA01000024.1	ATCC:35303	SE-32A-C8
<i>L. fairfieldensis</i>	NZ_JHYC01000039.1	ATCC:49588	ATCC 49588
<i>L. fallonii</i>	LN614827.1	ATCC:700992	LLAP-10
<i>L. feeleeii</i>	NZ_LBHK01000054.1	ATCC:35072	ATCC 35072
<i>L. geestiana</i>	LN YC01000041.1	ATCC:49504	ATCC 49504
<i>L. gormanii</i>	NZ_LBAY01000056.1	ATCC:33297	ATCC 33297
<i>L. gratiana</i>	LN YE01000004.1	ATCC:49413	Lyon 8420412
<i>L. gresilensis</i>	NZ_CAAAHX010000028.1	ATCC:700509	Greoux 11D13
<i>L. hackeliae</i>	NZ_LN681225.1	ATCC:35250	ATCC 35250
<i>L. israelensis</i>	CP038273.1	ATCC:43119	Bercovier 4
<i>L. jamestowniensis</i>	LN YG01000003.1	ATCC:35298	JA-26-G1-E2
<i>L. jordanis</i>	LN YJ01000005.1	ATCC:33623	BL-540
<i>L. lansingensis</i>	LN YI01000026.1	ATCC:49751	ATCC 49751
<i>L. londiniensis</i>	LN YK01000008.1	ATCC:49505	ATCC 49505
<i>L. longbeachæ</i>	CP020412.2	ATCC:33462	ATCC 33462
<i>L. maceachernii</i>	NZ_FUXJ01000030.1	ATCC:35300	ATCC 35300
<i>L. massiliensis</i>	NZ_CCVW01000002.1	DSM:24804	LegA
<i>L. micdadei</i>	NZ_LN614830.1	ATCC:33218	ATCC 33218
<i>L. moravica</i>	LN YN01000019.1	ATCC:43877	ATCC 43877
<i>L. nautarum</i>	LN YO01000023.1	ATCC:49506	ATCC 49506
<i>L. oakridgensis</i>	NZ_LCUA01000039.1	ATCC:33761	ATCC 33761
<i>L. parisiensis</i>	LN YQ01000005.1	ATCC:35299	PF-209-C-C2
<i>L. pneumophila</i>	NC_002942.5	ATCC:33152	Philadelphia 1
<i>L. quateirensis</i>	LN YR01000011.1	ATCC:49507	ATCC 49507
<i>L. quinlivanii</i>	LN YS01000014.1	ATCC:43830	CDC#1442-AUS-E
<i>L. rubrilucens</i>	LN YT01000018.1	ATCC:35304	WA-270A-C2
<i>L. sainthelensi</i>	NZ_JHXP01000047.1	ATCC:35248	ATCC 35248

TABLE 3 (Continued)

<i>Legionella</i> species	GenBank accession number no.	Culture collection	Type strain
<i>L. santicrucis</i>	LNJU01000018.1	ATCC:35301	SC-63-C7
<i>L. saoudiensis</i>	NZ_LN901320.1	DSM:101682	LS-1
<i>L. septentrionalis</i>	NZ_RZGS01000010.1	NBRC:113219	km711
<i>L. shakespearei</i>	LNW01000039.1	ATCC:49655	ATCC 49655
<i>L. spiritensis</i>	LNXX01000029.1	ATCC:35249	Mt.St.Helens-9
<i>L. steelei</i>	LNYY01000006.1	ATCC:BAA2169	IMVS3376
<i>L. steigerwaltii</i>	LNZY01000025.1	ATCC:35302	SC-18-C9
<i>L. taurinensis</i>	UGOZ01000001.1	NCTC:13314	NCTC13314
<i>L. tucsonensis</i>	LNZA01000005.1	ATCC:49180	ATCC 49180
<i>L. tunisiensis</i>	NZ_CALJ01000292.1	DSM:24805	LegM
<i>L. wadsworthii</i>	NZ_JNIA01000004.1	ATCC:33877	ATCC 33877
<i>L. waltersii</i>	LNZB01000016.1	ATCC:51914	ATCC 51914
<i>L. worsleiensis</i>	LNZC01000014.1	ATCC:49508	ATCC 49508
<i>L. yabuuchiae</i>	NZ_CAAAIW010000035.1	DSM:18492	OA1-2

* in bold type are reported *Legionella* species found during environmental surveillance.

The matrices were built using the multiple sequence comparison by log-expectation (MUSCLE) program [49] in Geneious Prime 2021.1.1 (<https://www.geneious.com>), retaining the default settings. The developed matrices permitted us to obtain a minimum and a maximum value of variation to establish intra- and interspecies intervals of divergence for the identification of the environmental isolates used in the present study. In detail, our 135 isolates were considered wild strains, and an in-house numbering scheme was used to label them (MR 1 to MR 135) (Table 4).

Table 4 GenBank accession numbers and ID labels of the *Legionella* wild strains used in this study

ID <i>Legionella</i> isolates	<i>Legionella</i> isolates	GenBank accession no.	
		<i>mip</i>	<i>rpoB</i>
MR1	<i>L. anisa</i>	MW021138	MZ367042
MR2	<i>L. anisa</i>	MW052865	MZ367043
MR3	<i>L. anisa</i>	MW052867	MZ367044
MR4	<i>L. anisa</i>	MW052869	MZ367045
MR5	<i>L. anisa</i>	MW052981	MZ367046
MR6	<i>L. anisa</i>	MW052872	MZ367047
MR7	<i>L. anisa</i>	MW052874	MZ367048
MR8	<i>L. anisa</i>	MW052875	MZ367049
MR9	<i>L. anisa</i>	MW052995	MZ367050
MR10	<i>L. steelei</i>	MW052877	MZ367051
MR11	<i>L. nautarum</i>	MW052931	MZ367052
MR12	<i>L. taurinensis</i>	MW052925	MZ367053
MR13	<i>L. taurinensis</i>	MW052973	MZ367054
MR14	<i>L. taurinensis</i>	MW052882	MZ367055
MR15	<i>L. rubrilucens</i>	MW052886	MZ367056
MR16	<i>L. anisa</i>	MW052879	MZ367057
MR17	<i>L. rubrilucens</i>	MW052895	MZ367058
MR18	<i>L. anisa</i>	MW052881	MZ367059
MR19	<i>L. rubrilucens</i>	MW052929	MZ367060
MR20	<i>L. rubrilucens</i>	MW052927	MZ367061
MR21	<i>L. anisa</i>	MW052885	MZ367062
MR22	<i>L. rubrilucens</i>	MW052914	MZ367063
MR23	<i>L. rubrilucens</i>	MW052919	MZ367064
MR24	<i>L. rubrilucens</i>	MW052890	MZ367065
MR25	<i>L. rubrilucens</i>	MW052893	MZ367066
MR26	<i>L. taurinensis</i>	MW052897	MZ367067
MR27	<i>L. taurinensis</i>	MW052901	MZ367068
MR28	<i>L. taurinensis</i>	MW052905	MZ367069
MR29	<i>L. taurinensis</i>	MW052908	MZ367070
MR30	<i>L. taurinensis</i>	MW052912	MZ367071
MR31	<i>L. anisa</i>	MW052891	MZ367072
MR32	<i>L. anisa</i>	MW052883	MZ367073
MR33	<i>L. anisa</i>	MW052894	MZ367074
MR34	<i>L. rubrilucens</i>	MW052917	MZ367075
MR35	<i>L. rubrilucens</i>	MW052921	MZ367076
MR36	<i>L. quinlivanii</i>	MW052923	MZ367077
MR37	<i>L. taurinensis</i>	MW052870	MZ367078
MR38	<i>L. taurinensis</i>	MW052873	MZ367079
MR39	<i>L. anisa</i>	MW052898	MZ367080
MR40	<i>L. anisa</i>	MW052876	MZ367081
MR41	<i>L. anisa</i>	MW052880	MZ367082
MR42	<i>L. anisa</i>	MW052887	MZ367083
MR43	<i>L. anisa</i>	MW052902	MZ367084
MR44	<i>L. anisa</i>	MW052863	MZ367085

(Continued on next page)

TABLE 4 (Continued)

ID <i>Legionella</i> isolates	<i>Legionella</i> isolates	GenBank accession no.	
		<i>mip</i>	<i>rpoB</i>
MR45	<i>L. anisa</i>	MW052866	MZ367086
MR46	<i>L. anisa</i>	MW052868	MZ367087
MR47	<i>L. anisa</i>	MW052871	MZ367088
MR48	<i>L. anisa</i>	MW052982	MZ367089
MR49	<i>L. anisa</i>	MW052983	MZ367090
MR50	<i>L. anisa</i>	MW052984	MZ367091
MR51	<i>L. anisa</i>	MW052985	MZ367092
MR52	<i>L. anisa</i>	MW052906	MZ367093
MR53	<i>L. anisa</i>	MW052909	MZ367094
MR54	<i>L. anisa</i>	MW052913	MZ367095
MR55	<i>L. taurinensis</i>	MW052888	MZ367096
MR56	<i>L. taurinensis</i>	MW052878	MZ367097
MR57	<i>L. londiniensis</i>	MW052907	MZ367098
MR58	<i>L. taurinensis</i>	MW052910	MZ367099
MR59	<i>L. taurinensis</i>	MW052915	MZ367100
MR60	<i>L. taurinensis</i>	MW052920	MZ367101
MR61	<i>L. rubrilucens</i>	MW052889	MZ367102
MR62	<i>L. rubrilucens</i>	MW052892	MZ367103
MR63	<i>L. rubrilucens</i>	MW052896	MZ367104
MR64	<i>L. rubrilucens</i>	MW052900	MZ367105
MR65	<i>L. anisa</i>	MW052904	MZ367106
MR66	<i>L. quateirensis</i>	MW052978	MZ367107
MR67	<i>L. quateirensis</i>	MW052911	MZ367108
MR68	<i>L. quateirensis</i>	MW052916	MZ367109
MR69	<i>L. feeleii</i>	MW052922	MZ367110
MR70	<i>L. feeleii</i>	MW052924	MZ367111
MR71	<i>L. feeleii</i>	MW052926	MZ367112
MR72	<i>L. feeleii</i>	MW052928	MZ367113
MR73	<i>L. feeleii</i>	MW052930	MZ367114
MR74	<i>L. rubrilucens</i>	MW052933	MZ367115
MR75	<i>L. rubrilucens</i>	MW052934	MZ367116
MR76	<i>L. rubrilucens</i>	MW052935	MZ367117
MR77	<i>L. rubrilucens</i>	MW052936	MZ367118
MR78	<i>L. rubrilucens</i>	MW052937	MZ367119
MR79	<i>L. rubrilucens</i>	MW052938	MZ367120
MR80	<i>L. rubrilucens</i>	MW052939	MZ367121
MR81	<i>L. rubrilucens</i>	MW052940	MZ367122
MR82	<i>L. rubrilucens</i>	MW052941	MZ367123
MR83	<i>L. rubrilucens</i>	MW052942	MZ367124
MR84	<i>L. nautarum</i>	MW052944	MZ367125
MR85	<i>L. quateirensis</i>	MW052945	MZ367126
MR86	<i>L. anisa</i>	MW052946	MZ367127
MR87	<i>L. anisa</i>	MW052947	MZ367128
MR88	<i>L. anisa</i>	MW052948	MZ367129

TABLE 4 (Continued)

ID <i>Legionella</i> isolates	<i>Legionella</i> isolates	GenBank accession no.	
		<i>mip</i>	<i>rpoB</i>
MR89	<i>L. anisa</i>	MW052949	MZ367130
MR90	<i>L. anisa</i>	MW052950	MZ367131
MR91	<i>L. anisa</i>	MW052951	MZ367132
MR92	<i>L. anisa</i>	MW052952	MZ367133
MR93	<i>L. taurinensis</i>	MW052954	MZ367134
MR94	<i>L. taurinensis</i>	MW052955	MZ367135
MR95	<i>L. londiniensis</i>	MW052976	MZ367136
MR96	<i>L. londiniensis</i>	MW052977	MZ367137
MR97	<i>L. anisa</i>	MW052957	MZ367138
MR98	<i>L. anisa</i>	MW052958	MZ367139
MR99	<i>L. anisa</i>	MW052959	MZ367140
MR100	<i>L. londiniensis</i>	MW052960	MZ367141
MR101	<i>L. londiniensis</i>	MW052975	MZ367142
MR102	<i>L. londiniensis</i>	MW052961	MZ367143
MR103	<i>L. londiniensis</i>	MW052962	MZ367144
MR104	<i>L. feeleii</i>	MW052963	MZ367145
MR105	<i>L. nautarum</i>	MW052964	MZ367146
MR106	<i>L. nautarum</i>	MW052965	MZ367147
MR107	<i>L. nautarum</i>	MW052966	MZ367148
MR108	<i>L. nautarum</i>	MW052967	MZ367149
MR109	<i>L. nautarum</i>	MW052968	MZ367150
MR110	<i>L. nautarum</i>	MW052969	MZ367151
MR111	<i>L. anisa</i>	MW052986	MZ367152
MR112	<i>L. anisa</i>	MW052987	MZ367153
MR113	<i>L. taurinensis</i>	MW052972	MZ367154
MR114	<i>L. taurinensis</i>	MW052971	MZ367155
MR115	<i>L. anisa</i>	MW052991	MZ367156
MR116	<i>L. anisa</i>	MW052993	MZ367157
MR117	<i>L. anisa</i>	MW052990	MZ367158
MR118	<i>L. anisa</i>	MW052994	MZ367159
MR119	<i>L. anisa</i>	MW052988	MZ367160
MR120	<i>L. anisa</i>	MW052989	MZ367161
MR121	<i>L. anisa</i>	MW052992	MZ367162
MR122	<i>L. rubrilucens</i>	MW052979	MZ367163
MR123	<i>L. feeleii</i>	MW052974	MZ367164
MR124	<i>L. rubrilucens</i>	MW052980	MZ367165
MR125	<i>L. taurinensis</i>	MW052996	MZ367166
MR126	<i>L. nautarum</i>	MW052997	MZ367167
MR127	<i>L. nautarum</i>	MW052998	MZ367168
MR128	<i>L. nautarum</i>	MW052999	MZ367169
MR129	<i>L. nautarum</i>	MW053000	MZ367170
MR130	<i>L. nautarum</i>	MW053001	MZ367171
MR131	<i>L. nautarum</i>	MW053002	MZ367172
MR132	<i>L. taurinensis</i>	MW053003	MZ367173

TABLE 4 (Continued)

ID <i>Legionella</i> isolates	<i>Legionella</i> isolates	GenBank accession no.	
		<i>mip</i>	<i>rpoB</i>
MR133	<i>L. nautarum</i>	MW053004	MZ367174
MR134	<i>L. taurinensis</i>	MW053005	MZ367175
MR135	<i>L. pneumophila</i>	MW053006	MZ367176

Phylogenetic and allelic diversity analysis. To estimate the relationship among the *Legionella* isolates involved in the study, a multiple sequence alignment (MSA) and a phylogenetic tree were performed on the *mip* and *rpoB* gene sequences. For each taxon identified as previously described, the reference *mip* and *rpoB* gene sequences of the corresponding type strains from several culture collections were retrieved and added to the analysis (Table 3). When required, manual editing was performed on the sequences, trimming them to the same length as the reference sequence. The nucleotide sequences were aligned by the MUSCLE program. The obtained MSA was passed to FastTree [50], a tool for inferring approximate maximum likelihood phylogenetic trees. FastTree uses Jukes-Cantor as a genetic distance model and the Shimodaira-Hasegawa test to estimate the reliability of each split in the tree [51]. Branch lengths were transformed to be equal, as in a cladogram. Branch labels display the substitutions per site. Both MUSCLE and FastTree were performed in Geneious Prime 2021.1.1 (<https://www.geneious.com>), retaining the default settings.

Data availability

The GenBank accession numbers of sequences generated during this study are listed in Table 4.

Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

S.C., S.S., and M.R.P. conceived and designed the experiments and wrote the paper. M.M. and L.G. performed sample collection and the experiments. S.S. performed the phylogenetic analysis. G.F.S., L.M., and A.G. performed the mip and rpoB gene sequencing.

We declare no conflicts of interest.

Reference

1. Mercante JW, Winchell JM. Current and emerging *Legionella* diagnostics for laboratory and outbreak investigations. *Clin Microbiol Rev* 2015; **28**: 95–133.
2. Cunha BA, Burillo A, Bouza E. Legionnaires' disease. *Lancet* 2016; **387**: 376–385.
3. Mondino S, Schmidt S, Rolando M, Escoll P, Gomez-Valero L, Buchrieser C. Legionnaires' Disease: State of the Art Knowledge of Pathogenesis Mechanisms of *Legionella*. *Annu Rev Pathol Mech Dis* 2020; **15**: 439–466.
4. Diederer BMW. *Legionella* spp. and Legionnaires' disease. *J Infect* 2008; **56**: 1–12.
5. Rowbotham TJ. Preliminary report on the pathogenicity of *Legionella pneumophila* for freshwater and soil amoebae. *J Clin Pathol* 1980; **33**: 1179–1183.
6. Fields BS. The molecular ecology of *legionellae*. *Trends Microbiol* 1996; **4**: 286–290.
7. Parte AC, Carbasse JS, Meier-Kolthoff JP, Reimer LC, Göker M. List of prokaryotic names with standing in nomenclature (LPSN) moves to the DSMZ. *Int J Syst Evol Microbiol* 2020; **70**: 5607–5612.
8. Muder RR, Yu VL. Infection due to *Legionella* species other than *L. pneumophila*. *Clin Infect Dis* 2002; **35**: 990–998.
9. Pierre DM, Baron J, Yu VL, Stout JE. Diagnostic testing for Legionnaires' disease. *Ann Clin Microbiol Antimicrob* 2017; **16**: 1–4.
10. Whiley H, Bentham R. *Legionella longbeachae* and legionellosis. *Emerg Infect Dis* 2011; **17**: 579–583.
11. Vaccaro L, Izquierdo F, Magnet A, Hurtado C, Salinas MA, Gomes TS, et al. First case of legionnaire's disease caused by *Legionella anisa* in Spain and the limitations on the diagnosis of *Legionella non-pneumophila* infections. *PLoS One* 2016; **11**: 1–12.
12. Van Der Mee-Marquet N, Domelier AS, Arnault L, Bloc D, Laudat P, Hartemann P, et al. *Legionella anisa*, a possible indicator of water contamination by *Legionella pneumophila*. *J Clin Microbiol* 2006; **44**: 56–59.

13. Compain F, Bruneval P, Jarraud S, Perrot S, Aubert S, Napoly V, et al. Chronic endocarditis due to *Legionella anisa*: A first case difficult to diagnose. *New Microbes New Infect* 2015; **8**: 113–115.
14. Head BM, Trajtman A, Bernard K, Burdz T, Vélez L, Herrera M, et al. *Legionella* co-infection in HIV-associated pneumonia. *Diagn Microbiol Infect Dis* 2019; **95**: 71–76.
15. Fallon RJ, Stack BHR. Legionnaires' disease due to *Legionella anisa*. *J Infect* 1990; **20**: 227–229.
16. Fenstersheib MD, Miller M, Diggins C, Liska S, Detwiler L, Werner SB, et al. Outbreak of Pontiac fever due to *Legionella anisa*. *Lancet* 1990; **336**: 35–37.
17. Matsui M, Fujii SI, Shirowa R, Amemura-Maekawa J, Chang B, Kura F, et al. Isolation of *Legionella rubrilucens* from a pneumonia patient co-infected with *Legionella pneumophila*. *J Med Microbiol* 2010; **59**: 1242–1246.
18. Rota, Maria Cristina, Caporali, Maria Grazia, Bella A, Scaturro, Maria, Giannitelli, Stefania, Ricci ML. Annual report of legionellosis in Italy - 2018. *Notiziario dell'Istituto Superiore di Sanità* . 2019. Rome, Italy.
19. Italian National Institute of Health. Guidelines for Prevention and Control of Legionellosis. Approvate in Conferenza Stato-Regioni Seduta Del 7 Maggio 2015. Italy, 2015. 2015.
20. UNI EN ISO 11731:2017 Water quality - Enumeration of *Legionella*. 2017. Geneva, Switzerland.
21. Delgado-Viscogliosi P, Simonart T, Parent V, Marchand G, Dobbelaere M, Pierlot E, et al. Rapid method for enumeration of viable *Legionella pneumophila* and other *Legionella* spp. in water. *Appl Environ Microbiol* 2005; **71**: 4086–4096.
22. Ratcliff RM. Sequence-based identification of *Legionella*. *Methods Mol Biol* 2013; **954**: 57–72.
23. Waterer GW, Baselski VS, Wunderink RG. *Legionella* and community-acquired pneumonia: A review of current diagnostic tests from a clinician's viewpoint. *Am J Med* 2001; **110**: 41–48.
24. Gaia V, Fry NK, Afshar B, Luck PC, Meugnier H, Etienne J, et al. Consensus Sequence-Based Scheme for Epidemiological Typing of Clinical and Environmental Isolates of *Legionella*

pneumophila. *J Clin Microbiol* 2005; **43**: 2047–2052.

25. Ratzow S, Gaia V, Helbig JH, Fry NK, Lück PC. Addition of neuA, the gene encoding N-acetylneuraminate cytidyl transferase, increases the discriminatory ability of the consensus sequence-based scheme for typing *Legionella pneumophila* serogroup 1 strains. *J Clin Microbiol* 2007; **45**: 1965–1968.
26. 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.
27. Fry NK, Afshar B, Bellamy W, Underwood AP, Ratcliff RM, Harrison TG, et al. Identification of *Legionella* spp. by 19 European reference laboratories: Results of the European Working Group for Legionella Infections External Quality Assessment Scheme using DNA sequencing of the macrophage infectivity potentiator gene and dedicated online tools. *Clin Microbiol Infect* 2007; **13**: 1119–1124.
28. Dahllof I, Baillie H, Kjelleberg S. *rpoB*-based microbial community analysis avoids limitations inherent in 16S rRNA gene intraspecies heterogeneity. *Appl Environ Microbiol* 2000; **66**: 3376–3380.
29. Guan W, Xu Y, Chen D li, Xu J nan, Tian Y, Chen J ping. Application of multilocus sequence analysis (MLSA) for accurate identification of *Legionella* spp. Isolated from municipal fountains in Chengdu, China, based on 16S rRNA, mip, and rpoB genes. *J Microbiol* 2012; **50**: 127–136.
30. Kim BJ, Lee SH, Lyu MA, Kim SJ, Bai GH, Kim SJ, et al. Identification of mycobacterial species by comparative sequence analysis of the RNA polymerase gene (*rpoB*). *Int J Lepr Other Mycobact Dis* 1999; **67**: 514.
31. Ko KS, Lee HK, Park MY, Lee KH, Yun YJ, Woo SY, et al. Application of RNA polymerase β -subunit gene (*rpoB*) sequences for the molecular differentiation of *Legionella* species. *J Clin Microbiol* 2002; **40**: 2653–2658.
32. Liu W, Li L, Khan MA, Zhu F. Popular molecular markers in bacteria. *Mol Genet Microbiol Virol* 2012; **27**: 103–107.

33. De Zwaan R, Van Ingen J, Van Soolingen D. Utility of *rpoB* gene sequencing for identification of nontuberculous mycobacteria in the Netherlands. *J Clin Microbiol* 2014; **52**: 2544–2551.
34. Scola B La, Gundi V a KB, Khamis A, Raoult D. Sequencing of the *rpoB* gene and flanking spacers for molecular identification of *Acinetobacter* species. *Society* 2006; **44**: 827–832.
35. Renesto P, Gouvernet J, Drancourt M, Roux V, Raoult D. Use of *rpoB* gene analysis for detection and identification of *Bartonella* species. *J Clin Microbiol* 2001; **39**: 430–437.
36. Mollet C, Drancourt M, Raoult D. *rpoB* sequence analysis as a novel basis for bacterial identification. *Mol Microbiol* 1997; **26**: 1005–1011.
37. Ko KS, Lee HK, Park MY, Park MS, Lee KH, Woo SY, et al. Population genetic structure of *Legionella pneumophila* inferred from RNA polymerase gene (*rpoB*) and DotA gene (*dotA*) sequences. *J Bacteriol* 2002; **184**: 2123–2130.
38. Ko KS, Hong SK, Lee KH, Lee HK, Park MY, Miyamoto H, et al. Detection and identification of *Legionella pneumophila* by PCR-restriction fragment length polymorphism analysis of the RNA polymerase gene (*rpoB*). *J Microbiol Methods* 2003; **54**: 325–337.
39. Nielsen K, Hindersson P, Hoiby N, Bangsbo JM. Sequencing of the *rpoB* gene in *Legionella pneumophila* and characterization of mutations associated with rifampin resistance in the *Legionellaceae*. *Antimicrob Agents Chemother* 2000; **44**: 2679–2683.
40. Girolamini L, Salaris S, Orsini M, Pascale MR, Mazzotta M, Grottola A, 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.
41. Shevchuk O, Jäger J, Steinert M. Virulence properties of the *Legionella pneumophila* cell envelope. *Front Microbiol* 2011; **2**: 1–12.
42. Cianciotto NP, Eisenstein BI, Mody CH, Engleberg NC. A mutation in the *mip* gene results in an attenuation of *Legionella pneumophila* virulence. *J Infect Dis* 1990; **162**: 121–126.
43. J. Rasch, C. M Ünal, A. Klages, Ü. Karsli, N. Heinshohn, R. M. H. J. Brouwer, M. Richter, A. Dellermann MS. Peptidyl-Prolyl-*cis/trans*-Isomerases *Mip* and *PpiB* of *Legionella pneumophila* Contribute to Surface Translocation, Growth at Suboptimal Temperature, and Infection. 2019; 1–12.

44. Debroy S, Dao J, Söderberg M, Rossier O, Cianciotto NP. *Legionella pneumophila* type II secretome reveals unique exoproteins and a chitinase that promotes bacterial persistence in the lung. *Proc Natl Acad Sci U S A* 2006; **103**: 19146–19151.
45. Cummings MP, Segal MR. Few amino acid positions in *rpoB* are associated with most of the rifampin resistance in *Mycobacterium tuberculosis*. *BMC Bioinformatics* 2004; **5**: 1–7.
46. UNI EN ISO 19458:2006 - Water quality — Sampling for microbiological analysis. 2006. Geneva, Switzerland.
47. Mentasti M, Fry NK, Afshar B, Palepou-Foxley C, Naik FC, Harrison TG. Application of *Legionella pneumophila*-specific quantitative real-time PCR combined with direct amplification and sequence-based typing in the diagnosis and epidemiological investigation of Legionnaires' disease. *Eur J Clin Microbiol Infect Dis* 2012; **31**: 2017–2028.
48. Adékambi T, Drancourt M, Raoult D. The *rpoB* gene as a tool for clinical microbiologists. *Trends Microbiol* 2009; **17**: 37–45.
49. Edgar RC. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004; **32**: 1792–1797.
50. Price MN, Dehal PS, Arkin AP. Fasttree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 2009; **26**: 1641–1650.
51. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst Biol* 2010; **59**: 307–321.

Chapter 3

Draft Genome Sequences of Legionella Presumptive Novel Species Isolated during Environmental Surveillance in Artificial Water Systems

Chapter published as:

Girolamini L, Salaris S, Orsini M, Pascale MR, Mazzotta M, Grottola A, Cristino S. 2021. Draft genome sequences of Legionella presumptive novel species isolated during environmental surveillance in artificial water systems. Microbiol Resour Announc 10:e00307-21. <https://doi.org/10.1128/MRA.00307-21>.

Luna Girolamini and Silvano Salaris contributed equally to this work. Author order was determined by drawing straws.

ABSTRACT

We present the draft genome sequences of three *Legionella* strains that were isolated from a hotel water distribution system. *Legionella* species identification was performed by macrophage infectivity potentiator (*mip*) and RNA polymerase b subunit (*rpoB*) gene sequencing. Whole-genome sequencing and average nucleotide identity results supported the hypothesis of new *Legionella* species isolation.

The *Legionella* genus contains pathogenic Gram-negative bacteria that are ubiquitous in soil and water environments. It consists of more than 60 species, all of them potentially able to cause Legionnaires' disease, a severe form of pneumonia [1]. The *Legionella* sp. strains 27fs60 (S60), 30fs61 (S61), and 30cs62 (S62) were isolated from three different samples from a hotel's hot water distribution system in the Emilia-Romagna region (Italy) during a routine *Legionella* surveillance program. Water sampling and *Legionella* isolation were performed according to ISO 19458:2006 and ISO 11731:2017, respectively [2, 3]. Samples were seeded onto selective medium with glycine-vancomycin-polymyxin B-cycloheximide (GVPC) and were incubated for 15 days at 35±2°C in 2.5% CO₂. Suspected colonies were subcultured on buffered charcoal yeast extract (BCYE) without L-cysteine (Thermo Fisher Scientific, Basingstoke, UK). The DNA was extracted with InstaGene matrix (Bio-Rad, Hercules, CA, USA), and identification of isolates was performed by macrophage infectivity potentiator (*mip*) and RNA polymerase β subunit (*rpoB*) gene sequencing [4, 5]. Amplicons were sequenced using BigDye chemistry and analyzed on an ABI Prism 3100 genetic analyzer (Applied Biosystems, Foster City, CA, USA). The *mip* sequences were compared with the European Working Group for *Legionella* Infections (EWGLI) database. A BLAST search of the NCBI database was carried out for both *mip* and *rpoB* gene sequences. The best match returned was *Legionella quateirensis* reference strain ATCC 49507 (GenBank accession number [GCA_001467955.1](#)), with similarities of 98.45% and 94.8% for *mip* and *rpoB*, respectively. One hundred nanograms of genomic DNA was used for next-generation sequencing (NGS) library preparation using the Illumina Nextera XT DNA library preparation kit (New England Biolabs, Ipswich, MA, USA). Sequencing was performed on the Illumina NextSeq 500 platform (2 × 150 bp paired-end reads). Raw reads were used as input data for TORMES v.1.2.0 [6], an automated pipeline for analysis of whole bacterial genomes. TORMES includes sequence quality filtering (PRINSEQ v.0.20.4) [7] and de novo genome assembly (SPAdes v.13.4.1) [8], as well as other downstream analyses not used for our purpose. Scaffolding was performed using TORMES contigs as input for CSAR v.1.1.1 [9] with an evolutionarily related reference genome, i.e., *Legionella fallonii* (GenBank accession number [NZ_LN614827.1](#)) The final assemblies were further improved using Geneious Prime v.2020.2.4 software [10] and were submitted to GenBank with annotation by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v.4.3 [11]. Default parameters were used for all software tools unless otherwise noted. Table 1 summarizes results from assembly and annotation by the PGAP and the completeness of genome assembly determined by Benchmarking Universal Single-Copy Orthologs (BUSCO) v.5.0.0 [12].

Table 1 Genome statistics from NCBI and BUSCO quality analyses

Attribute	Data for strain:		
	27fs60 (S60)	30fs61 (S61)	30cs62 (S62)
No. of raw reads	4,184,062	3,851,726	3,626,424
Avg read length (bp)	149	149	149
Coverage (x)	142	131	124
Total Length (bp)	4,211,919	3,709,497	4,136,543
No. of contigs	23	37	32
GC content (%)	39.00	39.10	39.00
N ₅₀ (bp)	312,097	166,809	176,017
No. of coding sequences	3,542	3,155	3,491
No. of rRNA	3	3	3
No. of tRNA	41	37	39
BUSCO results			
(%[no. of genes])			
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 FastANI tool [13] was used to compare the average nucleotide identity (ANI) of the three strains against 1,009 *Legionella* sequences that had been downloaded from the NCBI database using the ncbi-genome-download tool (<https://github.com/kblin/ncbi-genome-download>). FastANI identified the closest relative of strain S60 to be *L. quateirensis* NCTC 12376 (GenBank accession number [GCA_900452695.1](https://www.ncbi.nlm.nih.gov/nuccore/GCA_900452695.1)) (91.31%) and the closest relative of strains S61 and S62 to be *L. quateirensis* ATCC 49507 (91.45% and 91.44%, respectively). Since the assumption is that two strains showing pairwise ANI values below a given threshold (95% or 96 %) belong to different species [14], our results led us to consider these strains new species. Studying the whole genome allows investigators to better identify already known species and to discover new ones, improving the knowledge of the ecological, virulence, and resistance characteristics of *Legionella*.

Data availability

The draft genome assemblies are available in the GenBank database and can be accessed with SRA and assembly accession numbers [SRP292355](#) and [JADOBG000000000](#) (S60), [SRP295125](#) and [JADWVM000000000](#) (S61), and [SRP295130](#) and [JADWVN000000000](#) (S62).

Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

L.G., S.S., and S.C. conceived and designed the experiments and wrote the paper. M.R.P. and M.M. performed sample collection and culture experiments. M.O. performed the whole-genome sequencing. A.G. performed gene sequencing. L.G., S.S., and M.O. performed the bioinformatics analysis.

References

1. Jomehzadeh N, Moosavian M, Saki M, Rashno M. *Legionella* and legionnaires' disease: An overview. *J Acute Dis* 2019; **8**: 221–232.
2. UNI EN ISO 19458:2006 - Water quality — Sampling for microbiological analysis. 2006. Geneva, Switzerland.
3. UNI EN ISO 11731:2017 Water quality - Enumeration of *Legionella*. 2017. Geneva, Switzerland.
4. 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.
5. Ko KS, Lee HK, Park MY, Lee KH, Yun YJ, Woo SY, et al. Application of RNA polymerase β -subunit gene (rpoB) sequences for the molecular differentiation of *Legionella* species. *J Clin Microbiol* 2002; **40**: 2653–2658.
6. Quijada NM, Rodríguez-Lázaro D, Eiros JM, Hernández M. TORMES: an automated pipeline for whole bacterial genome analysis. *Bioinformatics* 2019; **35**: 4207–4212.
7. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 2011; **27**: 863–864.
8. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012; **19**: 455–477.
9. Chen KT, Liu CL, Huang SH, Shen HT, Shieh YK, Chiu HT, et al. CSAR: A contig scaffolding tool using algebraic rearrangements. *Bioinformatics* 2018; **34**: 109–111.
10. Kears M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 2012; **28**: 1647–1649.
11. Tatusova T, Dicuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, et al. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 2016; **44**: 6614–6624.
12. Seppey M, Manni M ZE. BUSCO: assessing genome assembly and annotation completeness. *Methods Mol Biol* 2019; **1962**: 227–245.
13. 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.
14. 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.

Chapter 4

Draft Genome Sequence of Legionella Species Isolated from Drinking Water in an Italian Industry

Chapter published as:

*Girolamini L, Salaris S, Orsini M, Pascale MR,
Mazzotta M, Grottola A, Cristino S. 2021.
Draft genome sequences of Legionella species
isolated from drinking water in an Italian industry.
Microbiol Resour Announc 11:e01152-21.
doi: 10.1128/mra.01152-21.*

Luna Girolamini and Silvano Salaris contributed equally to this work. Author order was determined by drawing straws.

Abstract

We report the draft genome sequences of an environmental *Legionella* strain isolated from an industrial water distribution system in Italy. Macrophage infectivity potentiator (*mip*) and b-subunit of RNA polymerase (*rpoB*) genes were used to perform the species identification. Whole-genome sequencing (WGS) and average nucleotide identity (ANI) identified the isolate as belonging to a presumptive novel *Legionella* species, with a genome length of 3,281,851 bp.

Legionella spp. are pathogenic Gram-negative bacteria that are ubiquitous in water and soil. *Legionella* includes more than 66 species and some of them are potentially able to cause a severe form of pneumonia, called Legionnaires' disease [1]. The *Legionella* sp. strain 31f133 was isolated from a drinking water in a company located in the Emilia-Romagna region (Italy) during a routine *Legionella* surveillance program.

Water samples and *Legionella* isolation were performed according to ISO 19458:2006 and ISO 11731:2017, respectively [2, 3]. Samples were seeded onto selective medium with glycine-vancomycin-polymyxin B-cycloheximide (GVPC) (Thermo Fisher Diagnostics, Basingstoke, UK) and incubated until 15 days at $35\pm 2^{\circ}\text{C}$ in 2.5% CO₂. Suspected colonies were subcultured on buffered charcoal yeast extract (BCYE) with and without L-cysteine (Thermo Fisher). DNA isolation was performed by InstaGene Matrix (Bio-Rad, Hercules, CA, USA) and the identification of isolate was performed by macrophage infectivity potentiator (*mip*) and RNA polymerase β subunit (*rpoB*) genes sequencing [4, 5]. A BigDye kit was used for the sequencing reaction and DNA sequences were analyzed on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). Sequences obtained were analyzed on the BLAST platform by the National Center for Biotechnology Information (NCBI) and European Working Group for Legionella Infections (EWGLI) databases. The best match returned was *L. feeleii*, reference strain ATCC 35072 (GenBank accession no. [GCA_001648615.1](#)), with similarities of 98.2% and 95.1% for *mip* and *rpoB*, respectively. An Illumina Nextera XT DNA Library Preparation kit (Illumina, New England Biolabs, Ipswich, MA, USA) was used to perform next-generation sequencing (NGS) library preparation using 100 ng of DNA. Subsequently, the Illumina NextSeq 500 platform (2×150 bp paired-end reads) was used for the sequencing. The bioinformatics workflow for the whole-genome sequencing (WGS) analysis included the following steps: through TORMES v.1.2.0 [6], an automated pipeline for analysis of the whole bacterial genome, raw reads were subjected to sequence quality filtering (PRINSEQ v.0.20.4) [7] and de novo genome assembly (SPAdes v.13.4.1) [8]. The generated contigs were passed to CSAR v.1.1.1 [9] in order to build the scaffolds. Scaffolding was performed using the genomes of different evolutionarily related organisms based on taxonomic identification: *Legionella hackeliae* strain ATCC 35250 (GenBank accession no. [LN681225.1](#)) by using Kraken2 v.2.0.9 [10], and *Legionella feeleii* strain ATCC 35072 based on *mip* and *rpoB* identification. The best scaffolding result was obtained for *L. hackeliae*.

To close or reduce the gaps contained in the CSAR output, a remapping of the reads using the scaffolds as a reference sequence was performed with Geneious Prime v.2021.2.2 software (<http://www.geneious.com>) [11]. The obtained draft genome was submitted to GenBank requiring the annotation, performed by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP v.4.3) [12] with the following accession numbers: [SRR16560654](#) and [JAJHHJ000000000](#). The results, summarized in Table 1, represent the assembling and annotation by the PGAP and the completeness of the genome assembly determined by Benchmarking Universal Single-Copy Orthologs (BUSCO) v.5.0.0 [13].

Table 1 Genome statistics data obtained from NCBI and BUSCO quality analyses

Attribute	Data for strain 31f133
No. of raw reads	1,321,792
Avg read length (bp)	253
Coverage (x)	99
Total Length (bp)	3,281,851
No. of contigs	3
GC content (%)	41.3
N ₅₀ (bp)	1,369,339
No. of coding sequences	2,968
No. of rRNA	1
No. of tRNA	41
BUSCO results (% [no. of genes])	
Complete	100 (124)
Single-copy complete	100 (124)
Duplicated complete	0.0 (0)
Fragmented	0.0 (0)
Missing	0 (0)
Total no. of BUSCO genes	124

A DFAST_QC [14] analysis was carried out using FastANI [15] to calculate average nucleotide identity (ANI) for a taxonomic identity of the genome by querying against 13,000 reference genomes from NCBI type strains. FastANI identified the *L. feeleii* WO-44C (ATCC 35072) (GenBank accession no. [GCA_900639755.1](#)) as the closest relative strain for our isolate 31f133, with a similarity of 93.99%. Therefore, we can consider this strain as a new *Legionella* species due to the assumption that two strains belonging to different species show pairwise ANI values below a 95% identity threshold [16].

Data availability

The draft genome assembly is available in the GenBank database and can be accessed with SRA and assembly accession numbers [SRR16560654](#) and [JAJHHJ0000000000](#).

Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or non-profit sectors.

S.C. is the corresponding author. L.G., S.S., and S.C. conceived and designed the experiments and wrote the paper. L.G., M.R.P., and M.M. performed sample collection and culture experiments. M.O. performed the whole-genome sequencing. A.G. performed gene sequencing. L.G., S.S., and M.O. performed the bioinformatics analysis.

References

1. Jomehzadeh N, Moosavian M, Saki M, Rashno M. *Legionella* and legionnaires' disease: An overview. *J Acute Dis* 2019; **8**: 221–232.
2. UNI EN ISO 19458:2006 - Water quality — Sampling for microbiological analysis. 2006. Geneva, Switzerland.
3. UNI EN ISO 11731:2017 Water quality - Enumeration of *Legionella*. 2017. Geneva, Switzerland.
4. 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.
5. Ko KS, Lee HK, Park MY, Lee KH, Yun YJ, Woo SY, et al. Application of RNA polymerase β -subunit gene (rpoB) sequences for the molecular differentiation of *Legionella* species. *J Clin Microbiol* 2002; **40**: 2653–2658.
6. Quijada NM, Rodríguez-Lázaro D, Eiros JM, Hernández M. TORMES: an automated pipeline for whole bacterial genome analysis. *Bioinformatics* 2019; **35**: 4207–4212.
7. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 2011; **27**: 863–864.
8. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012; **19**: 455–477.
9. Chen KT, Liu CL, Huang SH, Shen HT, Shieh YK, Chiu HT, et al. CSAR: A contig scaffolding tool using algebraic rearrangements. *Bioinformatics* 2018; **34**: 109–111.
10. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *bioRxiv* 2019; 1–13.
11. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 2012; **28**: 1647–1649.
12. Tatusova T, Dicuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, et al. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 2016; **44**: 6614–6624.
13. Seppey M, Manni M ZE. BUSCO: assessing genome assembly and annotation completeness. *Methods Mol Biol* 2019; **1962**: 227–245.
14. Tanizawa Y, Fujisawa T, Nakamura Y. DFAST: A flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* 2018; **34**: 1037–1039.
15. 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.
16. 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.

Chapter 5

Combining Traditional and Molecular Techniques Supports the Discovery of a Novel Legionella Species During Environmental Surveillance in a Healthcare Facility

Chapter published as:

Girolamini L, Pascale MR, Mazzotta M, Spiteri S, Marino F, Salaris S, Grottola A, Orsini M and Cristino S (2022).

Combining Traditional and Molecular Techniques Supports the Discovery of a Novel Legionella Species During Environmental Surveillance in a Healthcare Facility.

Front. Microbiol. 13:900936. doi: 10.3389/fmicb.2022.900936

Luna Girolamini and Maria Rosaria Pascale have contributed equally to this work.

Abstract

Legionella surveillance plays a significant role not only to prevent the risk of infection but also to study the ecology of isolates, their characteristics, and how their prevalence changes in the environment. The difficulty in *Legionella* isolation, identification, and typing results in a low notification rate; therefore, human infection is still underestimated. In addition, during *Legionella* surveillance, the special attention given to *Legionella pneumophila* leads to an underestimation of the prevalence and risk of infection for other species. This study describes the workflow performed during environmental *Legionella* surveillance that resulted in the isolation of two strains, named 8cVS16 and 9fVS26, associated with the genus *Legionella*. Traditional and novel approaches such as standard culture technique, MALDI-TOF MS, gene sequencing, and whole-genome sequencing (WGS) analysis were combined to demonstrate that isolates belong to a novel species. The strain characteristics, the differences between macrophage infectivity potential (*mip*), RNA polymerase β subunit (*rpoB*), and reference gene sequences, the average nucleotide identity (ANI) of 90.4%, and the DNA–DNA digital hybridization (dDDH) analysis of 43% demonstrate that these isolates belong to a new *Legionella* species. The finding suggests that, during the culture technique, special attention should be paid to the characteristics of the isolates that are less associated with the *Legionella* genus in order to investigate the differences found using more sensitive methods. The characterization of the two newly discovered isolates based on morphological, biochemical, and microscopic characteristics is currently underway and will be described in another future study.

Keywords

mip sequencing, *rpoB* sequencing, MALDI-TOF MS, water distribution system, whole-genome sequencing (WGS)

Introduction

Legionella spp. are gram-negative intracellular pathogenic bacteria, are ubiquitous in water and soil, and are represented by more than 66 species, some of which are potentially capable to cause a severe form of pneumonia, called Legionnaires' disease [1, 2]. *Legionella pneumophila* (*Lp*) is the most common infectious agent involved in Legionnaires' disease and consists of 15 different serogroups. *Lp* serogroup 1 (*Lp1*), according to epidemiological data, is mostly associated with human infections [3, 4]. Nevertheless, other *Legionella non-pneumophila* species (non-*Lp* species) (i.e., *L. anisa*, *L. rubrilucens*, or *L. longbeachae*) are responsible for human infections that are typically acquired through inhalation of contamination aerosol [3–7]. Artificial water distribution systems (WDSs) are considered the sites most associated with *Legionella* spp. proliferation and spread [8]. The presence of biofilm, water stagnation due to low water flow, dead branches, or old pipelines are factors that make the WDS's facilities (e.g., large public or private buildings, companies, hospitals, or health facilities) a potential risk to the dissemination of *Legionella* spp. [9–12]. The main attention is given to the WDS of hospitals and healthcare facilities (HCF), where *Legionella* spp. contamination is considered a high-risk factor due to the presence of elderly or immunocompromised patients who are more susceptible to the infection [13]. Legionnaires' disease in Italy, in 2020, amounted to 34.3 cases per million inhabitants; therefore, *Legionella* colonization of hospitals or communities WDS represents one of the main public health concerns [4, 14, 15]. Starting from the epidemiological data, different countries have developed specific guidelines and promoted environmental monitoring programs in order to prevent and control *Legionella* infections, following the Water Safety Plan (WSP) approach, introduced by the World Health Organization (WHO) [16]. Through this approach, an ongoing *Legionella* surveillance program is suggested, also in the absence of cases, to undertake the most appropriate control and prevention measures. The recent European Union (EU) Directive 2020/2184 suggests the prevention approach. This directive will be transposed in all the EU countries by January 2023 [17]. In this directive, *Legionella* has been introduced as a new microbiological parameter for the evaluation of the drinking water quality, highlighting the importance of the environmental monitoring of the WDS, starting from the water supply until the consumers' outlets. Moreover, to achieve control of *Legionella* proliferation, the national and international directives point out several disinfection strategies, based on chemical disinfectants (i.e., chlorine dioxide, monochloramines, and hydrogen peroxide) and physical treatments [i.e., ultraviolet (UV) light and hot temperature treatments]. All

of them have shown some advantages and disadvantages [18–22]. However, some authors have shown how these treatments, especially the use of chemical disinfectants, could select resistant strains and introduce some changes in the bacterial genome. Therefore, the use of diagnostic methods able to follow the ecological and adaptive isolates changes should be the main objective to arise also during *Legionella* environmental surveillance [23, 24]. Certainly, this aspect is strictly correlated with the sensitivity and specificity of detection and identification methods already available and in use. In recent times, the most common methods for *Legionella* isolation and identification in routinely environmental surveillance remain the culture technique and the latex agglutination test, despite the molecular approach as sequence-based typing (SBT) and subgrouping scheme based on monoclonal antibodies (MABs) for *Lp* [25–27] and macrophage infectivity potentiator (*mip*) for non-*Lp* species [28–30] are consolidated. Despite several approaches that have been developed through time, most of them are applied only when clusters or outbreaks occur and are carried out only by specialized or national reference laboratories. The limits of all the methods, previously cited, are summarized as follows: (i) long incubation time of culture technique (at least 10 days); (ii) the agglutination test that is not able to recognize all *Legionella* species, showing also mis-identification or false-negative results; (iii) MABs, that can type only *Lp* and in particular for subtyping of *Lp*1, available to the national reference laboratories; (iv) SBT technique and gene sequencing that permits compare only some gene or short fragments of gene size [26, 31–34]. Consequently, only in recent years with the data returned by whole-genome sequencing (WGS), it has been possible to obtain the most complete genome information, improving isolate typing as well as functional and drug susceptibility response. Moreover, the WGS analysis has opened new scenarios for the reconstruction of infection spread, establishing a correct relationship between environmental and clinical strains during the epidemiological investigation [35]. However, even with its great usefulness, it is not routinely used to support environmental surveillance. Despite the high operational costs, the WGS pipelines could potentially reduce overall costs for the hospitals as well as in all facilities' practices through savings of indirect costs [35]. In the context of *Legionella* environmental monitoring, the primary goal is to improve WSPs, quantify the risk level, and identify isolates. Moreover, also during routine environmental surveillance, it is possible to isolate and characterize novel new *Legionella* species, suggesting how the water microflora is subject to continuous changes during the time and in response to the treatment (physical, chemical, and functional) undertaken [36]. In this study, we described the workflow applied during *Legionella*

environmental surveillance conducted for 10 years in an HCF that led to the isolation of a novel *Legionella* species. The knowledge acquired during the surveillance period regarding the WDS, the disinfection treatment, and the level of contamination were associated with traditional techniques (i.e., culture and latex agglutination test) as well as the innovative matrix-assisted laser desorption ionization–time-of-flight mass spectrometry (MALDI-TOF MS) technique, gene sequencing, and WGS analysis. All results obtained confirm the presence in HCF of two isolates not previously detected.

Materials and methods

Characteristics of healthcare facility's water distribution system

The HCF involved in this study is a long-term care facility, built in 2011, and made up of 78 dislocated inpatient rooms on three floors for a total of 120 beds. The *Legionella* environmental surveillance program was started in 2012 with an elaboration of a *Legionella* risk assessment plan according to the Italian and Regional Guidelines [37, 38]. Briefly, the *Legionella* monitoring was carried out two times a year, usually in the spring/summer and fall/winter. The *Legionella* concentration above the risk level as suggested by guidelines requires further sampling of positive samples [37, 38]. The WDS consists of cold water derived from municipal water, heated by two water heaters in parallel, one of them connected to the solar panel. The temperature of hot water at the supply outlets was about 50°C. The hot water system is treated with hydrogen peroxide and silver salt-based disinfectant (H₂O₂/Ag⁺). The continuous disinfection treatment provides a residual concentration to the distal outlets around 20 mg/L. The WDS schematic (**Figure 1**) was developed with Solid Edge 2022 V222.00.02.03 (Siemens Digital Industry Software Inc.).

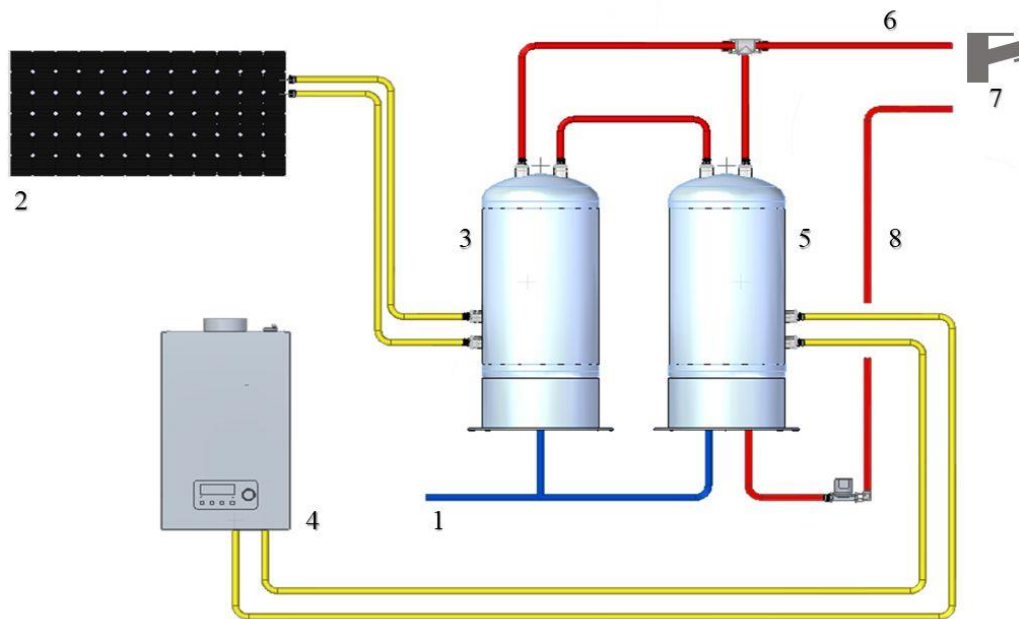


Figure 1 Representation of the HCF water distribution system: 1: cold-water supply; 2: solar panel; 3: solar heater water tank; 4: gas water heater; 5: heater water tank; 6: hot water output; 7: water distal outlets; 8: hot water return line.

Sample collection

According to the *Legionella* risk assessment plan, hot- and cold-water samples were collected every 6 months. The sampling points were chosen in accordance with Italian Guidelines [37], considering the building size, the number of inpatient rooms, the risk level of patients, and workers' exposure to bacteria, other than the facility's epidemiological data. All of these data are reported in the *Legionella* WSP developed since 2012. Starting from the technical room, and following the hot and cold WDS, for each floor, four samples were chosen as follows: three hot water samples in the vicinity of, mid-way to, and away from the technical room, and one cold-water sample away from the water supply point. A total of 30 sampling points were identified between the technical room (water supply point, hot water output, hot water return line, and hot water storage tanks) and HCF inpatient rooms (including showers, sinks, and toilet showerheads) and staff facilities. The criterion of rotation between inpatient rooms was applied. Two liters of hot or cold water for each sample was collected in post-flushing modality (2min), following the Italian National Unification and European Committee (UNI EN) International Standard Organization (ISO) 19458:2006 [39]. During sampling, temperature and disinfectant residue values were measured and reported in distal outlets. The analysis was carried out on the same day of sampling.

Microbiological analysis and isolates characterization

The *Legionella* isolation was performed using a standard culture technique in accordance with ISO 11731:2017 [40]. Briefly, for the enumeration of *Legionella*, different aliquots of the sample (from 200 to 100 µl), which comes from filtered water (untreated), heat- and acid-treated, were seeded on glycine–polymyxin B–vancomycin–cycloheximide (GVPC) selective agar (Thermo Fisher Scientific, Diagnostic, Ltd., Basingstoke, UK) and incubated at $35 \pm 2^\circ\text{C}$ with 2.5% of CO₂. The culture required a minimum of 10 for up to 15 days. Every 2 days, the plates were examined and the presumptive colonies were enumerated and sub-cultured on buffered charcoal yeast extract (BCYE) agar with and without L-cysteine (Cys+) and L-cysteine (Cys–) (Thermo Fisher Scientific, Diagnostic, Ltd., Basingstoke, UK). The *Legionella* colonies, growth only on BCYE Cys+, were identified using the Legionella latex agglutination test kit differentiating between *Lp*1, and *Lp* serogroups 2–14 (*Lp*2-14) and seven species of non-*Lp* species (Thermo Fisher Scientific, Ltd. Basingstoke, UK), based on manufacturing instructions.

MALDI-TOF MS Analysis

The isolates grown on BCYE Cys+, which returned positive or negative results for the Legionella agglutination test, were also analyzed by the MALDI Biotyper system (Bruker Daltonik GmbH, Bremen, Germany) as previously described [41]. Briefly, a fresh colony (24–48 h of incubation) was directly spotted in duplicate onto a MALDI Biotyper target plate, overlaid with 1 µl of the MALDI Biotyper matrix solution, and left to air dry before the next step. Spectra acquisition and processing were performed using the Microflex LT mass spectrometer (2,000–20,000 Da, linear positive mode) and the MALDI Biotyper Compass 4.1 software, whose library (version BDAL 7854) included the spectra of 39 *Legionella* strains. The data were interpreted in accordance with the manufacturer's instructions. Briefly, when the instrument returned a log score ≥ 2.0 ("high confidence level"), the isolates were identified at the species level, while the genus was assigned for scores between 1.7 and 1.99 ("low confidence level"). In the presence of a score between 0.00 and 1.69, the isolates were considered as "not identified." A dendrogram based on hierarchical clustering analysis (HCA) of MALDI Biotyper spectra was developed using the MALDI Biotyper Compass Explorer software to generate a treelike structure able to link the *Legionella* strains to each other using a linkage algorithm.

Sequencing of *mip* and *rpoB* genes for *Legionella* identification

InstaGene Matrix (Bio-Rad, Hercules, CA, USA) was used for DNA extraction that was quantified by Qubit fluorometer (Thermo Fisher Scientific, Paisley, UK). *mip* and *rpoB* genes were used to perform the identification of the isolates according to Ratcliff et al. [28], Ko et al. [29], and Pascale et al. [30]. PCR products were visualized by electrophoresis on 2% agarose gel. BigDye kit was used for the sequencing reaction, and the sequences were analyzed on ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). Raw sequencing data were assembled using CLC Main Workbench 7.6.4 software (QIAGEN, Redwood City, CA, USA). Furthermore, the *mip* and *rpoB* gene sequence comparison was performed using the Basic Local Alignment Search Tool (BLAST) search on the database of the National Center for Biotechnology Information (NCBI) and the database developed by the European Working Group for *Legionella* Infections [(EWGLI), renamed in ESCMID Study Group for *Legionella* Infections (ESGLI) from 2011]. The identification at the species level, considering the *mip* gene, was performed on the basis of an identity score between 98 and 100% compared to the sequences in the database [42] and considering the intervals variation of interspecies and intraspecies previously described by Ratcliff et al. [28]. As reported in Ko et al. [29], Adékambi et al. [43], and Pascale et al. [30] regarding the *rpoB* gene sequence, the threshold used for the identification was fixed at a 95% similarity percentage.

Phylogenetic and allelic diversity analysis

To estimate the relationship between the isolates found and the strictly related *Legionella* species, a multiple sequence alignment (MSA) and a concatenated phylogenetic tree were developed on the *mip* and *rpoB* gene sequences. Manual editing was performed on the sequences, if required, trimming them to the same length as the reference sequence. The tree was built using software implemented in Geneious Prime's genome browser (Geneious Prime 2022.0.2; <http://www.geneious.com>) maintaining the default settings [44]. Through the MUSCLE algorithm, the nucleotide sequences were aligned [45]. FastTree [46], a tool to deduce the approximate maximum likelihood of phylogenetic trees, was used to pass the resulting MSA. Jukes–Cantor was used, by FastTree, as a genetic distance model, and the Shimodaira–Hasegawa test was utilized to estimate the reliability of each split in the tree (default parameters) [47]. As in a cladogram, the lengths of the branches have been converted to be equal. Branch labels display the substitutions by the site.

Identification of *Legionella* by Whole-Genome Sequencing (WGS)

One hundred nanograms of DNA was used to prepare the library for the next-generation sequencing (NGS) carried out by Illumina Nextera XT DNA Library Preparation kit (Illumina, New England Biolabs, Ipswich, MA, USA). The Illumina NextSeq 500 platform (2 × 250 bp paired-end reads) was used to perform the sequencing. Raw reads were assembled using TORMES v.1.2.0 [48], an automated pipeline for whole bacterial genome analysis, using the default parameters. TORMES performed a sequence quality filtering (PRINSEQ v. 0.20.4) and a de novo genome assembly (SPAdes v. 13.4.1) [49]. The generated contigs were passed to CSAR v1.1.1 [50], 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. *Legionella* sp. PC1000 (NZ_CP059400.1) was selected as a reference sequence, based on the output of KmerFinder [51–53] that predicts prokaryotic species based on the number of overlapping k-mers, i.e., 16-mers, between the query genome and genomes in a reference database (NCBI). A further refinement was carried out by remapping the reads on the CSAR scaffolds, using the Geneious Prime 2022.0.2 software (<http://www.geneious.com>) [44]. Benchmarking Universal Single-Copy Orthologs (BUSCO) v.5.0.0 [54] was performed to evaluate the completeness of the two genome assemblies. The final draft genomes were submitted to the GenBank requiring the annotation by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP v.4.3) [55]. The OrthoANI package [56] was used to measure the intra- and inter-species genome similarities by average nucleotide identity (ANI) among the assembled draft genomes. Further, FastANI [57] through DFAST [58] was performed against 13000 prokaryotic reference genomes from NCBI to assess the taxonomic identity. It was also measured based on BLAST+ (ANIm) and MUMmer (ANIm) using JSpeciesWS [59]. The relatedness of our strains to *Legionella*-known type strains was further analyzed by applying a digital DNA–DNA hybridization (dDDH) to the closest related strain based on the previous ANI outcome. 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+ [60] as a local alignment tool. The GGDC uses a Genome Blast Distance Phylogeny (GBDP) to infer genome-to-genome distances between pairs of entirely or partially sequenced genomes. Below the similarity threshold of the in silico DDH (70%) and the ANI analysis (95%), two strains are considered distinct species [61, 62].

RESULTS

Legionella spp. isolation and features

Starting from 2012, the *Legionella* environmental monitoring returned mostly negative results (<50 colony formant unit (cfu)/L). Only two samples were positive for *Legionella* in March 2015 and 2016 with a concentration of 200 and 250 cfu/L, respectively. The samples were collected from the shower and toilet showerhead in two different inpatient rooms, both located on the ground floor. The water temperature measured was 45.8 and 44.0°C, with a disinfectant residue of 20 mg/L. Regarding the colonies' morphology, they are small (about 3mm), gray-white, with a round shape, and well compact. Concerning the identification, both isolates grew well on BCYE Cys+ and showed a quick positive reaction to the latex agglutination test for non-*Lp* species. In addition, when exposed under a Woods lamp (long wavelength UV light at 365 nm), they showed blue-white autofluorescence (Figure 2).

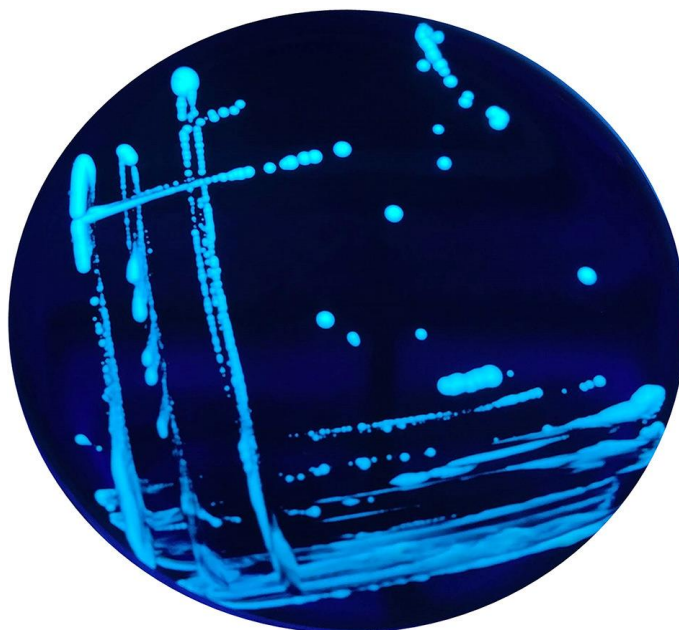


Figure 2 Blue-white autofluorescence of the non-*Lp* isolates grown on BCYE Cys+, found in 2015 and 2016, under Woods lamp (365 nm).

In particular, for further analysis, the strain isolated in 2015 was called 8cVS16, and the second one isolated in 2016 was named 9fVS26. Considering that no other samples evaluated in the same year and during the subsequent monitoring yielded positive results for *Legionella*, the contamination found may be considered point-source contamination. Instead, the two strains were no longer

detected in the water distribution system until December 2021, which was when the last sampling occurred.

MALDI-TOF MS results

The MALDI Biotyper system identified both strains with a low confidence score (yellow color) as *Legionella anisa*. In particular, the score returned for 8cVS16 was 1.76 and for 9fVS26 was 1.78. The dendrogram based on HCA, elaborated on the most closed *Legionella* species present in the instrument database, displays the 8cVS16 and 9fVS26 strains closed to each other and regrouped in one clade. This clade was well-separated by the two strictly related clades, represented by *L. anisa* and *Legionella bozemanii* reference strains (Figure 3).

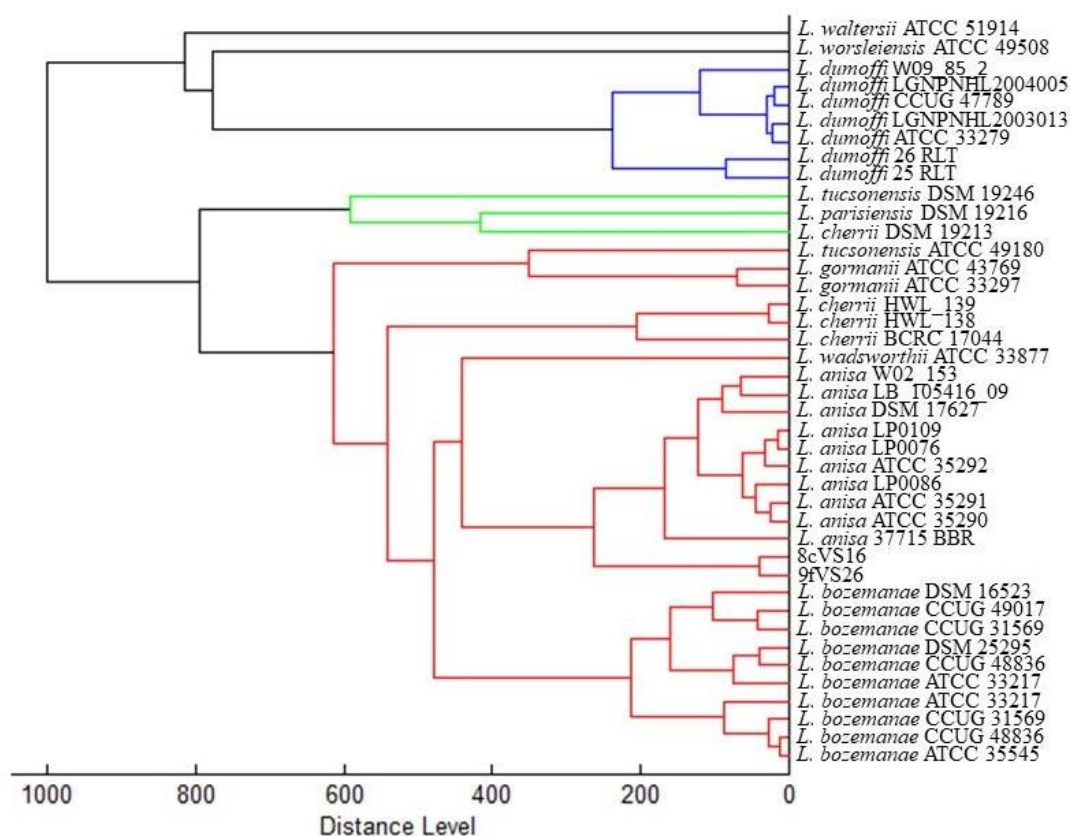


Figure 3 Dendrogram developed by HCA for 8cVS16 and 9fVS26 and related *Legionella* reference strains.

Therefore, the strains appeared very close to *L. anisa*, confirming the results returned by MALDI Biotyper software.

mip and *rpoB* results

BLAST research on NCBI and the ESGI database returned the best match for both strains *L. anisa*, reference strain ATCC 35292 (GenBank accession number GCA_900639785.1), with similarities of 96.7 and 92.4% for *mip* and *rpoB*, respectively.

Phylogenetic analysis results

The concatenated tree elaborated from *mip* and *rpoB* sequence genes (**Figure 4**) revealed the presence of two main clades, each containing four different subclades and each representing several *Legionella* species.

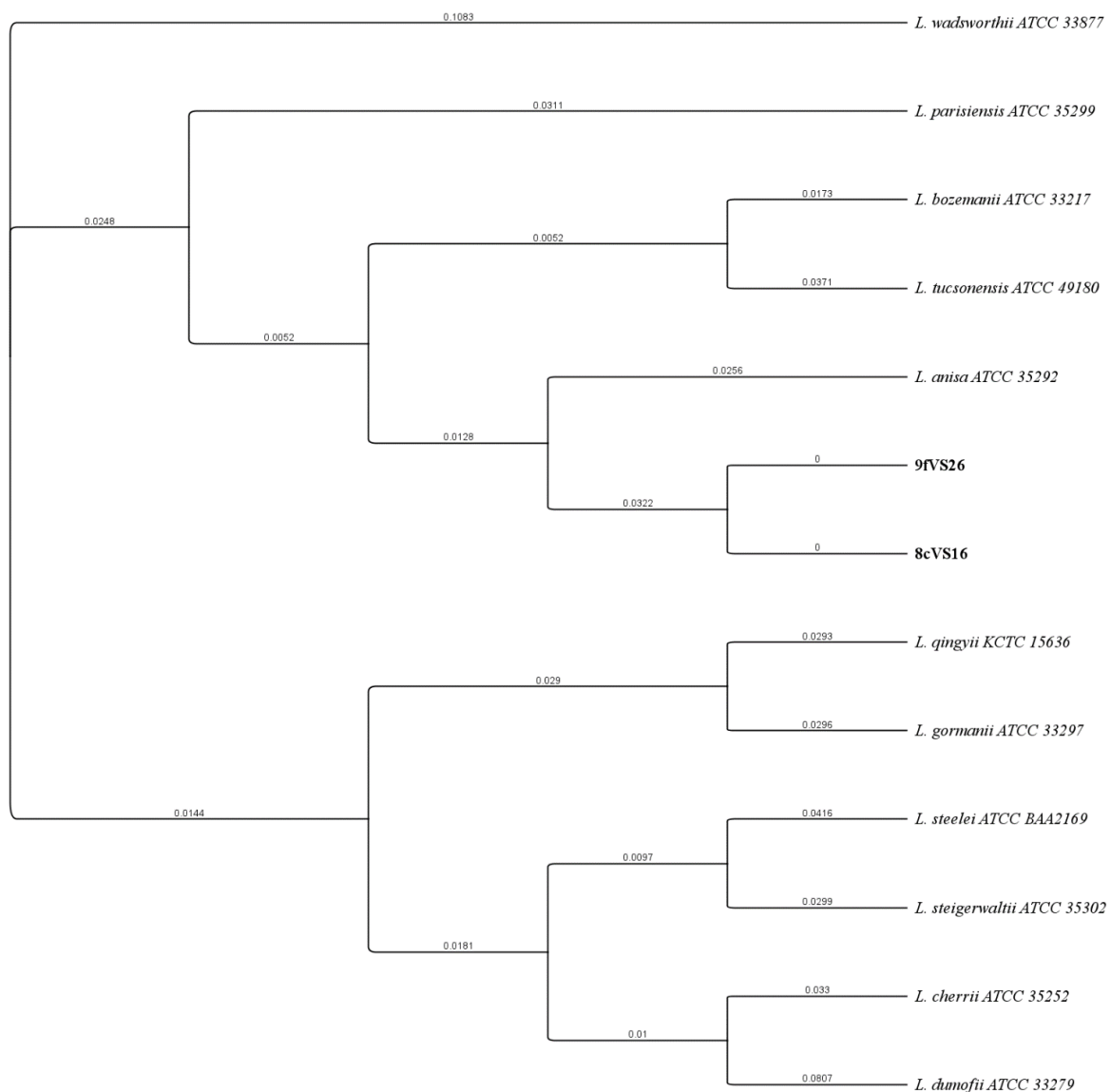


FIGURE 4 | Phylogenetic tree of the two strains (8cVS16 and 9fVS26) and closely related *Legionella* species based on concatenation of two genes (*mip* and *rpoB*). Branch lengths are transformed to be equal like a cladogram. Branch labels display the substitutions per site.

Interestingly, the 8cVS16 and 9fVS26 strains were collocated into a monophyletic group comprising *L. anisa*, *L. tucsonensis*, and *L. bozemanii* strains. Moreover, it was particularly evident that they were strictly related to each other, forming a short branch representing a sister clade with the *L. anisa* ATCC 35292.

WGS analysis results

All the results obtained by the WGS analysis are summarized in Table 1.

Table 1 Genome statistics data from NCBI and BUSCO quality analysis.

Attribute	Data for strain:	
	8cVS16	9cVS26
No. of raw reads	1,787,078	1,952,986
Avg read length (bp)	259	259
Coverage (×)	115	127
Total Length (bp)	3,906,083	3,906,100
No. of contigs	7	10
GC content (%)	38.2	38.2
N ₅₀ (bp)	855,940	858,038
No. of coding sequences	3,362	3,360
No. of rRNA	6	6
No. of tRNA	42	42
BUSCO results (%[no. of genes])		
Complete	95.2 (118)	95.2 (118)
Single-copy complete	95.2 (118)	95.2 (118)
Duplicated complete	0.0 (0)	0.0 (0)
Fragmented	0.8 (1)	0.8 (1)
Missing	4.0 (5)	4.0 (5)
Total no. of BUSCO genes	124	124

Briefly, the overall lengths of the genomes for 8cVS16 and 9fVS26, respectively, were 3,906,083 and 3,906,1003 bp, with a GC content of 38.2% for both. In particular, BUSCO analysis, performed to evaluate the completeness of the two genome assemblies, indicates that the two genomes are near-complete, with a percentage of 95.2%. Through the NCBI PGAP annotation required by GenBank to submit the draft genome, the following accession numbers were given: SRR17223244 and JAJTND0000000000 for 8cVS16 and SRR17223245 and JAJSPM0000000000 for 9fVS26. Comparison, obtained from the OrthoANI package used to measure the intra- and inter-species genome similarities, returned the following values: 99.98% between 8cVS16 and 9fVS26, confirming that the

two strains belong to the same species and are identical to each other. The taxonomic identity, obtained by FastANI, returned an identity percentage of the closest strain for both 8cVS16 and 9fVS26. This strain was *L. anisa* strain WA-316-C3 (ATCC 35292) (GCA_900639785.1) with 90.74% ANI. The analysis, performed by JSpeciesWS, confirmed the FastANI results with the following values: both 8cVS16 and 9fVS26 ANIb 90.08% compared with *L. anisa* WA-316-C3 (ATCC 35292), and ANIm 91.55% with *L. anisa* WA-316-C3 (ATCC 35292) for both 8cVS16 and 9fVS26 strains. The analysis of our strains' relatedness to known types of strains using a digital DNA-DNA hybridization (dDDH) yielded the following results: a DDH value (generalized linear model (GLM)-based) of 43% for both 8cVS16 and 9fVS26 and a probability that DDH > 70% (i.e., same species) of 5.44% (via logistic regression). Our results showed an ANI value of 90.74% and a dDDH value of 43%.

Discussion

The discovery and identification of novel bacterial species are events of great relevance, especially when they occur during routine environmental surveillance programs. The identification of novel species is much more frequent in the clinical setting, thanks to the most advanced phenotypic and genotypic technologies. In contrast, laboratories involved in environmental surveillance routinely use only the standard culture method, so detection of a potentially novel species relies heavily on the researcher's and technician's experience. The results obtained in this study prove that it is possible to detect isolates, which are difficult to identify or could be misclassified with standard methods, by applying more sensitive techniques such as genotyping and WGS, even during routine environmental monitoring, carried out in self surveillance. Commonly, also in the widely studied and known WDS, the selective stress induced by changes in water characteristics (e.g., pressure, temperature, inorganic, and organic compounds), as well as the continuous disinfection treatments, may occur during the selection of a novel species, which is never identified before [18][63, 64]. The HCF monitored, from 2012, did not show *Legionella* contamination, except for 2 years (2015 and 2016). During our surveillance, only one sample per year was found positive, with non-*Lp* species concentration over the level of risk fixed to <100 cfu/L. A continuous disinfection treatment (hydrogen peroxide and silver salt) was installed at the time of facility opening, considering the size of the facility and the characteristics of the patients (e.g., elderly, with chronic and immunocompromised diseases). According to other studies, the continuous dosage of this disinfectant can control *Legionella* proliferation, although, from HCS opening, *Legionella* was never

detected [22]. Therefore, we cannot know whether *Legionella* was always absent due to the presence of disinfectant, or whether the WDS characteristics and the fact that the structure was newly built, contributed to creating a “*Legionella* free” environment. The presence of punctual contamination in two outlets, during 10 years of monitoring, both on flexible shower tubes, which are poorly used, due to the presence of non-self-sufficient patients, suggests that probably these pipelines constituted a niche with a biofilm, were able to protect, and would promote *Legionella* survival [65–67]. Moreover, in these types of outlets, as in the showerhead tubes, it is difficult to undertake cleaning and maintenance practices due to the small size of pipelines. The substitution of the showerhead flexible tubes, performed after the communication of *Legionella* presence, could explain the absence of *Legionella* detection to date. It is not a matter to be neglected, as already demonstrated, that the continuous supply of chemical substances in WDS could promote disinfection resistance of bacterial species and the selection of novel ones [18, 63, 64]. We have already demonstrated in other facilities how the disinfectant based on peroxide hydrogen and silver salt in the same environment produces a different action on *Legionella*, promoting *L. anisa*’s resistance despite the high efficiency on *Lp1* [20, 24, 68]. The culture technique, which is essential for the isolation, is not effective in correctly recognizing these novel isolates when combined with the common and rapid identification methods (e.g., agglutination test). Our results clearly showed that the two isolates presented features similar to well-known non-*Lp* species (e.g., *L. anisa*, *L. bozemanii*, and *L. gormanii*) and showed a positive agglutination reaction to multiple species latex reagents. Moreover, considering also the MALDI-TOF MS results, they could be misclassified as *L. anisa*. In particular, the agglutination test has shown a positive result for non-*Lp* species antibodies and the phenotypic-proteomics analysis, provided by MALDI-TOF Biotyper, has identified the two strains as *L. anisa* with a low-confidence interval (yellow score), although the software can properly identify this species by assigning a green score, since the presence of its spectra in the instrument database. The results obtained were in accordance with the previous study [41, 69–71], confirming the ability of MALDI-TOF MS to classify all the *Legionella* species contained in the database. The low confidence score returned for our isolates could be explained by the differences that 8cVS16 and 9fVS26 spectra showed with the *L. anisa* reference spectra contained in the instrument database. So, we can suppose that the differences at the genomic level were also reflected at the ribosomal protein level, which is the target of MALDI-TOF MS technique. These differences found at the proteomics level appear more evident in the dendrogram returned by HCA, where it is possible to

observe how 8cVS16 and 9fVS26 form a separate clade with respect to the two main clades represented by *L. anisa* and *L. bozemanii*. Starting from these considerations, the sequencing of *Legionella* characteristic genes is essential to obtain discrimination at interspecies and intraspecies levels. Regarding the genomic analysis, *mip* and *rpoB* gene sequencing confirmed the previous results. The two isolates were associated with *L. anisa*, showing an identity percentage of 96.7% and 92.4%, for *mip* and *rpoB*, respectively, below the established threshold [28–30, 43]. The phylogenetic tree well-points out these differences, showing how the two isolates, despite their proximity to the *L. anisa* clade, represent a single subclade that needs to be further investigated. The discrepancies found in sequence identity percentage have led us to investigate in more depth by applying the WGS technique. The genomes of 8cVS16 and 9fVS26 were then compared with the closest genome returned by FASTANI belonging to *L. anisa* (ATCC 35292). Although the whole genome sequences for both isolates were almost completed, with a coverage of 95.2%, the ANI analysis returned a result of 90.74% and the dDDH analysis reported a value of 43% compared to the *L. anisa* ATCC 35292. Considering that the percentage obtained was under the established similarity thresholds (95 and 70%, respectively) [61, 62], we can consider the two isolates found belonging to a novel *Legionella* species. More detailed studies on (i) morphology, including flagellar structures, (ii) *Legionella* growth tests at various temperatures and on culture media, (iii) biochemical, and (iv) antibiotic susceptibility tests will be the next steps in obtaining isolates characterization, resulting in the deposit of the type strain in culture collections and the most complete description of the novel species. This study can be considered as a support tool for all laboratories that, during the phases of environmental surveillance, may find isolates with similar characteristics to the best-known species but with discrepancies in results. These laboratories are encouraged not to stop at the common identification tests but to continue the investigations in order to discover novel isolates and understand the dynamics that led to their development. The results of this study should not surprise us should we consider that the environment is undergoing profound changes (e.g., global warming), which can lead to an increase in temperature in both natural and artificial reservoirs. Moreover, the increasingly widespread use of disinfectants as a preventive strategy, other than changes in water characteristics and pipeline materials, leads to a strong impact on the selection of bacteria and the resistance development. Therefore, acquiring extensive knowledge of these reservoirs and the events that can promote changes in the ecological niche of bacteria can help to prevent their spread in the man-made environment and contain the

occurrence of cases, clusters, or outbreaks. In this context, environmental surveillance is critical, as is a proper risk assessment plan that takes into account changes in *Legionella* contamination over time. The latter will primarily focus on correct isolate identification and characterization, as well as novel approaches able to recognize non-*Lp* species with unknown rates of infection and pathogenicity other than antibiotic resistance.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

Author contributions

LG, MRP, SSa, and SC conceived and designed the experiments and wrote the paper. LG, MRP, MM, FM, and SSp performed sample collection and culture experiments. MO performed the whole-genome sequencing. AG performed gene sequencing. LG, SSa, and MO performed the bioinformatics analysis.

The manuscript has been read and approved by all named authors. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors thank Dr. Miriam Cordovana for her support and suggestions for the MALDI-TOF MS analysis.

References

1. Jomehzadeh N, Moosavian M, Saki M, Rashno M. *Legionella* and legionnaires' disease: An overview. *J Acute Dis* 2019; **8**: 221–232.
2. Parte AC, Carbasse JS, Meier-Kolthoff JP, Reimer LC, Göker M. List of prokaryotic names with standing in nomenclature (LPSN) moves to the DSMZ. *Int J Syst Evol Microbiol* 2020; **70**: 5607–5612.
3. ECDC: European Centre for Disease Prevention and Control. Legionnaires' disease. In: Annual Epidemiological Report for 2019. 2021. Stockholm.
4. Rota, M. C., Caporali, M. G., Bella, A., Scaturro, M., Giannitelli, S., and Ricci ML. Il sistema di sorveglianza della Legionellosi in Italia: i risultati del 2020. *Boll Epidemiol Naz* 2021; **1**: 32–38.
5. Muder RR, Yu VL. Infection due to *Legionella* species other than *L. pneumophila*. *Clin Infect Dis* 2002; **35**: 990–998.
6. Matsui M, Fujii SI, Shirowa R, Amemura-Maekawa J, Chang B, Kura F, et al. Isolation of *Legionella rubrilucens* from a pneumonia patient co-infected with *Legionella pneumophila*. *J Med Microbiol* 2010; **59**: 1242–1246.
7. Cunha BA, Burillo A, Bouza E. Legionnaires' disease. *Lancet* 2016; **387**: 376–385.
8. Mercante JW, Winchell JM. Current and emerging *Legionella* diagnostics for laboratory and outbreak investigations. *Clin Microbiol Rev* 2015; **28**: 95–133.
9. Di Pippo F, Di Gregorio L, Congestri R, Tandoi V, Rossetti S. Biofilm growth and control in cooling water industrial systems. *FEMS Microbiol Ecol* 2018; **94**: 1–13.
10. Totaro M, Costa AL, Frendo L, Profeti S, Casini B, Gallo A, et al. Evaluation of *Legionella* spp. Colonization in residential buildings having solar thermal system for hot water production. *Int J Environ Res Public Health* 2020; **17**: 1–8.
11. Nisar MA, Ross KE, Brown MH, Bentham R, Whiley H. *Legionella pneumophila* and protozoan hosts: Implications for the control of hospital and potable water systems. *Pathogens* 2020; **9**: 20–30.
12. Nisar MA, Ross KE, Brown MH, Bentham R, Whiley H. Water Stagnation and Flow Obstruction Reduces the Quality of Potable Water and Increases the Risk of Legionelloses. *Front Environ Sci* 2020; **8**: 1–13.
13. Spagnolo AM, Cristina ML, Casini B, Perdelli F. *Legionella pneumophila* in healthcare facilities. *Rev Res Med Microbiol* 2013; **24**: 70–80.
14. Kyritsi MA, Mouchtouri VA, Katsioulis A, Kostara E, Nakoulas V, Hatzinikou M, et al. *Legionella* colonization of hotel water systems in touristic places of Greece: association with system characteristics and physicochemical parameters. *Int J Environ Res Public Health* 2018; **15**.
15. Brunello A, Civilini M, De Martin S, Felice A, Franchi M, Iacumin L, et al. Machine learning-assisted environmental surveillance of *Legionella*: A retrospective observational study in Friuli-Venezia Giulia region of Italy in the period 2002–2019. *Informatics Med Unlocked* 2022; **28**: 100803.

16. Bartram J, Chartier Y, Lee JV, Pond K, Surman-Lee S, WHO. *Legionella* and the Prevention of Legionellosis. *Emerg Infect Dis* 2007; **14**: 1006a – 1006.
17. European Parliament the Council of the European Union. Directive (EU) 2020/2184 of the European Parliament and of the Council of 16 December 2020 on the quality of water intended for human consumption. *Official Journal of the European Union* . 2020.
18. McDonnell G, Russell D. Antiseptics and Disinfectants: Activity, Action, and Resistance. *Am Soc Microbiol* 1999; **12**: 147–179.
19. Mancini B, Scurti M, Dormi A, Grottola A, Zanotti A, Cristino S. Effect of monochloramine treatment on colonization of a hospital water distribution system by *Legionella* spp.: A 1 year experience study. *Environ Sci Technol* 2015; **49**: 4551–4558.
20. Girolamini L, Dormi A, Pellati T, Somaroli P, Montanari D, Costa A, et al. Advances in *Legionella* control by a new formulation of hydrogen peroxide and silver salts in a hospital hot water network. *Pathogens* 2019; **8**: 1–21.
21. Richardson SD, Plewa MJ, Wagner ED, Schoeny R, DeMarini DM. Occurrence, genotoxicity, and carcinogenicity of regulated and emerging disinfection by-products in drinking water: A review and roadmap for research. *Mutat Res - Rev Mutat Res* 2007; **636**: 178–242.
22. Lin YE, Stout JE, Yu VL. Controlling *Legionella* in Hospital Drinking Water: An Evidence-Based Review of Disinfection Methods. *Infect Control Hosp Epidemiol* 2011; **32**: 166–173.
23. Jakubek D, Le Brun M, Leblon G, Dubow M, Binet M. The impact of monochloramine on the diversity and dynamics of *Legionella pneumophila* subpopulations in a nuclear power plant cooling circuit. *FEMS Microbiol Ecol* 2013; **85**: 302–312.
24. Girolamini L, Salaris S, Pascale MR, Mazzotta M, Cristino S. Dynamics of *Legionella* Community Interactions in Response to Temperature and Disinfection Treatment: 7 Years of Investigation. *Microb Ecol* 2022; **83**: 353–362.
25. Helbig JH, Kurtz JB, Pastoris MC, Pelaz C, Lück PC. Antigenic lipopolysaccharide components of *Legionella pneumophila* recognized by monoclonal antibodies: Possibilities and limitations for division of the species into serogroups. *J Clin Microbiol* 1997; **35**: 2841–2845.
26. Gaia V, Fry NK, Afshar B, Luck PC, Meugnier H, Etienne J, et al. Consensus Sequence-Based Scheme for Epidemiological Typing of Clinical and Environmental Isolates of *Legionella pneumophila*. *J Clin Microbiol* 2005; **43**: 2047–2052.
27. Ratzow S, Gaia V, Helbig JH, Fry NK, Lück PC. Addition of neuA, the gene encoding N-acetylneuraminate cytidyl transferase, increases the discriminatory ability of the consensus sequence-based scheme for typing *Legionella pneumophila* serogroup 1 strains. *J Clin Microbiol* 2007; **45**: 1965–1968.
28. 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.
29. Ko KS, Lee HK, Park MY, Lee KH, Yun YJ, Woo SY, et al. Application of RNA polymerase β -subunit gene (rpoB) sequences for the molecular differentiation of *Legionella* species. *J Clin Microbiol* 2002; **40**: 2653–2658.

30. Pascale MR, Salaris S, Mazzotta M, Girolamini L, Fregni Serpini G, Manni L, et al. New Insight regarding *Legionella* Non- *Pneumophila* Species Identification: Comparison between the Traditional *mip* Gene Classification Scheme and a Newly Proposed Scheme Targeting the *rpoB* Gene. *Microbiol Spectr* 2021; **9**.
31. Helbig JH, Bernander S, Castellani Pastoris M, Etienne J, Gaia V, Lauwers S, et al. Pan-European study on culture-proven Legionnaires' disease: Distribution of *Legionella pneumophila* serogroups and monoclonal subgroups. *Eur J Clin Microbiol Infect Dis* 2002; **21**: 710–716.
32. Orsini M, Cristino S, Grottola A, Romano-Spica V. Bacteria misagglutination in *Legionella* surveillance programmes. *J Hosp Infect* 2011; **79**: 179–180.
33. Lück, C., Fry, N. K., Helbig, J. H., Jarraud, S., and Harrison TG. Typing Methods for *Legionella*. In: C. Buchrieser and H. Hilbi (Totowa NHP (ed). *Legionella: Methods and Protocols*. 2013. pp 119–148.
34. Walker JT, McDermott PJ. Confirming the Presence of *Legionella pneumophila* in Your Water System: A Review of Current *Legionella* Testing Methods. *J AOAC Int* 2021; **104**: 1135–1147.
35. Quainoo S, Coolen JPM, Sacha A. F. T. van Hijum C, Martijn A. Huynen c WJGM, Willem van Schaik E, Wertheim HFL. Whole-Genome Sequencing of Bacterial Pathogens : the Future of Nosocomial. *Clin Microbiol Rev* 2017; **30**: 1015–1064.
36. Li Q, Yu S, Li L, Liu G, Gu Z, Liu M, et al. Microbial communities shaped by treatment processes in a drinking water treatment plant and their contribution and threat to drinking water safety. *Front Microbiol* 2017; **8**: 1–16.
37. Italian National Institute of Health. Guidelines for Prevention and Control of Legionellosis. Approvate in Conferenza Stato-Regioni Seduta Del 7 Maggio 2015. Italy, 2015. 2015.
38. Emilia-Romagna Region. Regional Guidelines for Surveillance and Control of Legionellosis. Delibera Della Giunta Regionale 12 Giugno 2017, N. 828. 2017. 2017; **2012**: 1–136.
39. UNI EN ISO 19458:2006 - Water quality — Sampling for microbiological analysis. 2006. Geneva, Switzerland.
40. UNI EN ISO 11731:2017 Water quality - Enumeration of *Legionella*. 2017. Geneva, Switzerland.
41. Pascale MR, Mazzotta M, Salaris S, Girolamini L, Grottola A, Simone ML, et al. Evaluation of MALDI-TOF Mass Spectrometry in Diagnostic and Environmental Surveillance of *Legionella* Species: A Comparison With Culture and *Mip*-Gene Sequencing Technique. *Front Microbiol* 2020; **11**: 1–12.
42. Fry NK, Afshar B, Bellamy W, Underwood AP, Ratcliff RM, Harrison TG, et al. Identification of *Legionella* spp. by 19 European reference laboratories: Results of the European Working Group for *Legionella* Infections External Quality Assessment Scheme using DNA sequencing of the macrophage infectivity potentiator gene and dedicated online tools. *Clin Microbiol Infect* 2007; **13**: 1119–1124.
43. Adékambi T, Drancourt M, Raoult D. The *rpoB* gene as a tool for clinical microbiologists. *Trends Microbiol* 2009; **17**: 37–45.

44. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 2012; **28**: 1647–1649.
45. Edgar RC. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004; **32**: 1792–1797.
46. Price MN, Dehal PS, Arkin AP. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS One* 2010; **5**.
47. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst Biol* 2010; **59**: 307–321.
48. Quijada NM, Rodríguez-Lázaro D, Eiros JM, Hernández M. TORMES: an automated pipeline for whole bacterial genome analysis. *Bioinformatics* 2019; **35**: 4207–4212.
49. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012; **19**: 455–477.
50. Chen KT, Liu CL, Huang SH, Shen HT, Shieh YK, Chiu HT, et al. CSAR: A contig scaffolding tool using algebraic rearrangements. *Bioinformatics* 2018; **34**: 109–111.
51. Hasman H, Saputra D, Sicheritz-Ponten T, Lund O, Svendsen CA, Frimodt-Moller N, et al. Rapid whole-genome sequencing for detection and characterization of microorganisms directly from clinical samples. *J Clin Microbiol* 2014; **52**: 139–146.
52. Larsen M V., Cosentino S, Lukjancenko O, Saputra D, Rasmussen S, Hasman H, et al. Benchmarking of methods for genomic taxonomy. *J Clin Microbiol* 2014; **52**: 1529–1539.
53. Clausen PTLC, Aarestrup FM, Lund O. Rapid and precise alignment of raw reads against redundant databases with KMA. *BMC Bioinformatics* 2018; **19**: 1–8.
54. Seppey M, Manni M ZE. BUSCO: assessing genome assembly and annotation completeness. *Methods Mol Biol* 2019; **1962**: 227–245.
55. Tatusova T, Dicuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, et al. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 2016; **44**: 6614–6624.
56. Yoon SH, Ha S min, Lim J, Kwon S, Chun J. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek, Int J Gen Mol Microbiol* 2017; **110**: 1281–1286.
57. 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.
58. Tanizawa Y, Fujisawa T, Nakamura Y. DFAST: A flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* 2018; **34**: 1037–1039.
59. 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.

60. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: Architecture and applications. *BMC Bioinformatics* 2009; **10**: 1–9.
61. 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**.
62. 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.
63. Girolamini L, Salaris S, Orsini M, Pascale MR, Mazzotta M, Grottola A, 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.
64. Girolamini L, Salaris S, Orsini M, Pascale MR, Mazzotta M, Grottola A, et al. Draft Genome Sequence of *Legionella* Species Isolated from Drinking Water in an Italian Industry. *Microbiol Resour Announc* 2022; **11**: 20–22.
65. Storey M V, Ashbolt J, Stenström TA. Biofilms, thermophilic amoebae and *Legionella pneumophila*-a quantitative risk assessment for distributed water. *Water Sci Technol* 2004; **50**: 77–82.
66. Kaplan JB. Biofilm Dispersal: Mechanisms, Clinical Implications, and Potential Therapeutic Uses. *J Dent Res* 2010; **89**: 205–218.
67. Mahapatra A, Padhi N, Mahapatra D, Bhatt M, Sahoo D, Jena S, et al. Study of biofilm in bacteria from water pipelines. *J Clin Diagnostic Res* 2015; **9**: 9–11.
68. Farhat M, Moletta-Denat M, Frère J, Onillon S, Trouilhé MC, Robine E. Effects of disinfection on *Legionella* spp., eukarya, and biofilms in a hot water system. *Appl Environ Microbiol* 2012; **78**: 6850–6858.
69. Moliner C, Ginevra C, Jarraud S, Flaudrops C, Bedotto M, Couderc C, et al. Rapid identification of *Legionella* species by mass spectrometry. *J Med Microbiol* 2010; **59**: 273–284.
70. Gaia V, Casati S, Tonolla M. Rapid identification of *Legionella* spp. by MALDI-TOF MS based protein mass fingerprinting. *Syst Appl Microbiol* 2011; **34**: 40–44.
71. Sværre CW, Uldum SA. The occurrence of *Legionella* species other than *Legionella pneumophila* in clinical and environmental samples in Denmark identified by *mip* gene sequencing and matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clin Microbiol Infect* 2012; **18**: 1004–1009.

Chapter 6

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

Chapter published as:

Girolamini L, Pascale MR, Salaris S, Mazzotta M, Orsini M, Grottola A, Zini N, Cristino S.

Legionella bononiensis sp. nov., isolated from a hotel water distribution system in northern Italy. Int J Syst Evol Microbiol. 2022 Sep;72(9).

doi: 10.1099/ijsem.0.005512.

Luna Girolamini and Maria Rosaria Pascale contributed equally to this work.

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. The strains do not exhibit antimicrobial resistance. The *mip* and *rpoB* gene sequences of the three strains showed close matches to those of *Legionella quateirensis* ATCC 49507T 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 measure by average nucleotide identity between the three strains and their respective closest strains were: 91.32 % *L. quateirensis* NCTC 12376T, 91.45 % *L. quateirensis* ATCC 49507T and 91.45 % *L. quateirensis* ATCC 49507T, respectively. The digital DNA–DNA hybridization analysis demonstrated how the isolates were separated from the most related phylogenetic *Legionella* species (*L. quateirensis* ATCC 49507T, ≤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 49507T. The results obtained confirm the status of an independent species. The name proposed for this species is *Legionella bononiensis* sp. nov. with 30cs62T (=ATCC TSD-262^T=DSM 112526^T) as the type strain.

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 30cs62T are OK495707, OK495709 and OK495708, respectively. The GenBank accession numbers for the whole genome sequences of strains 27fs60, 30fs61 and 30cs62T are JADOBG000000000, JADWVM000000000 and JADWVN000000000, respectively.

Supplementary Material

Eight supplementary figures and four supplementary tables are available in the Appendix A

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.

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 % CO₂ 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).

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 <i>rpoB</i> gene (4107)	94.8	214	11
Partial <i>rpoB</i> (329)	94.5	18	0
Partial <i>mip</i> (611)	98.2	11	2

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 bp 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][17]. BUSCO version 5.0.0. was used 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).

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 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.

Table 3. *Genome statistics data from the NCBI*

Attribute	27fs60	30fs61	30cs62 ^T
No. of raw reads	4,184,062	3,851,726	3,626,424
Average read length (bp)	149	149	149
Coverage (×)	142	131	124
Total Length (bp)	4,211,919	3,709,497	4,136,543
No. of contigs	23	37	32
G+C Content (mol%)	39.00	39.10	39.00
N50 (bp)	312,097	166,809	176,017
No. of coding sequences	3,542	3,155	3,491
No. of rRNA gene	3	3	3
No. of tRNA gene	41	37	39

Briefly, the genomes total lengths were 4,211,919, 3,709,497 and 4,136,543 bp for 27fs60, 30fs61 and 30cs62^T, respectively, with a G+C content of approximately 39 mol%. 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^T) 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 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* and *L. shakespearei* strains. Particularly, they formed a sister group with the *L. quateirensis* ATCC 49507^T.

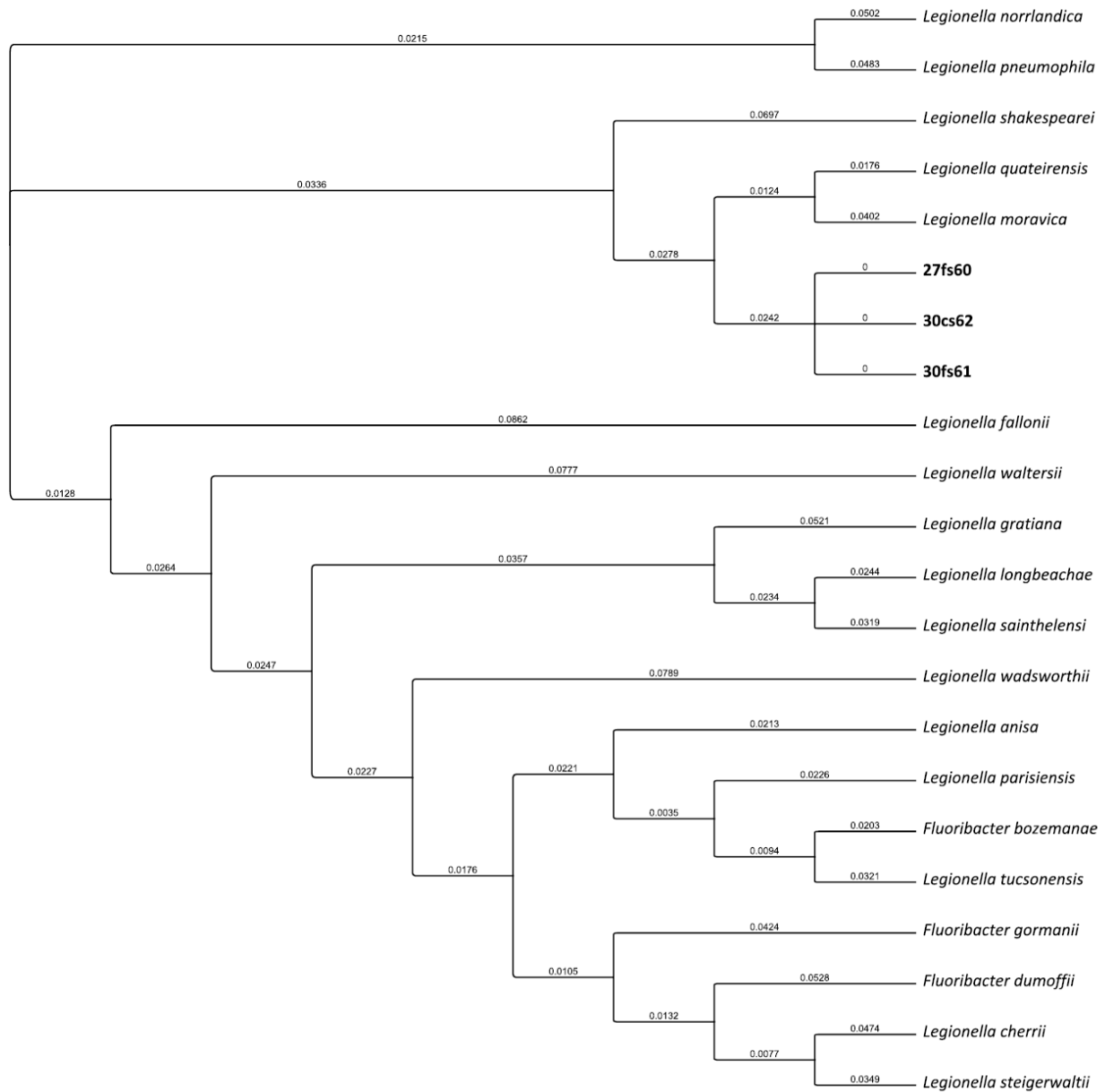


Fig. 1. Phylogenetic tree of the three strains (27fs60, 30fs61 and 30cs62^T) and closely related *Legionella* species based on concatenation of four genes (16S rRNA, *mip*, *rpoB*, *rnpB*). Branch lengths are transformed to be equal like a cladogram. Branch labels display the substitutions per site.

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 Cys + agar plates and the growth at different temperatures was evaluated: 32, 35 and 37 °C with 2.5 % of CO₂. 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).

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).



Fig. 2. Strain 30cs62T from a water sample on glycine -vancomycin -polymyxin B -cycloheximide (GVPC) agar medium.

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% CO₂. 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 CO₂, 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-wavelength 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).

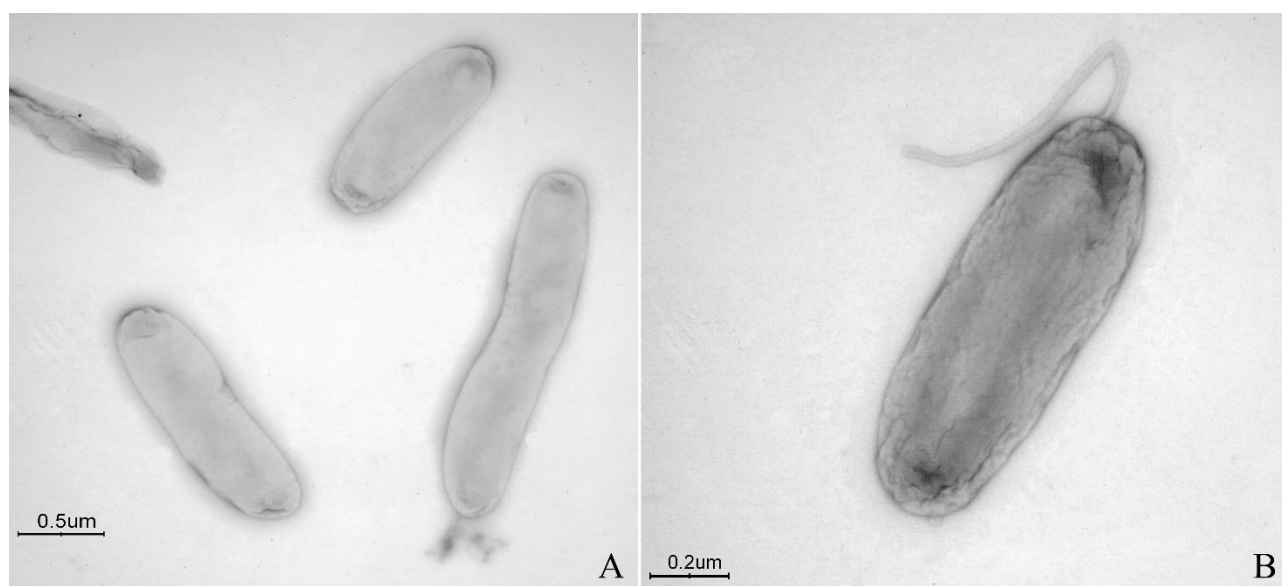


Fig. 3. Transmission electron microscopy (TEM) images of strain 30cs62T grown on BCYE Cys+ agar for 2 days at 35°C with 2.5% CO₂. (A) Aflagellate 30cs62^T form and (B) flagellate 30cs62^T form. Bar, (A) 0.5 μm and (B) 0.2 μm. Magnification: (A) ×12 000 and (B) ×20 000.

The transitory flagella presence could be 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.

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-b-naphthylamide)	-
MEL (melibiose)	-	GGT (g-Glutamyl b-naphthylamide)	+
RHA (rhamnose)	-	TRY (tryptophane b-naphthylamide)	-
SOR (sorbitol)	-	IND (tryptophane)	+
MNT (mannitol)	-	NO ₃ (sodium nitrate)	+
ADO (adonitol)	-	GLR (p-nitrophenyl β-glucuronide)	-
GAL (galactose)	-	NAG (p-Nitrophenyl-N-acetyl-β,D-glucosaminide)	-
PHO (p-nitrophenyl phosphate)	+	GGL (γ-L-glutamyl p-nitroanilide)	-
BGL (p-nitrophenyl α-β-glucoside)	-	ESC (esculin)	-
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 (p-Nitrophenyl-phosphoester)	+	BANA (N-Benzyl-arginine-b-naphthylamide)	-
αGLU (p-Nitrophenyl-α,D-glucoside)	-	EST (triglyceride)	+
βGLU (p-Nitrophenyl-β,D-glucoside)	-	INO (inositol)	-
ONPG (p-Nitrophenyl, b,D-galactoside)	-		

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.

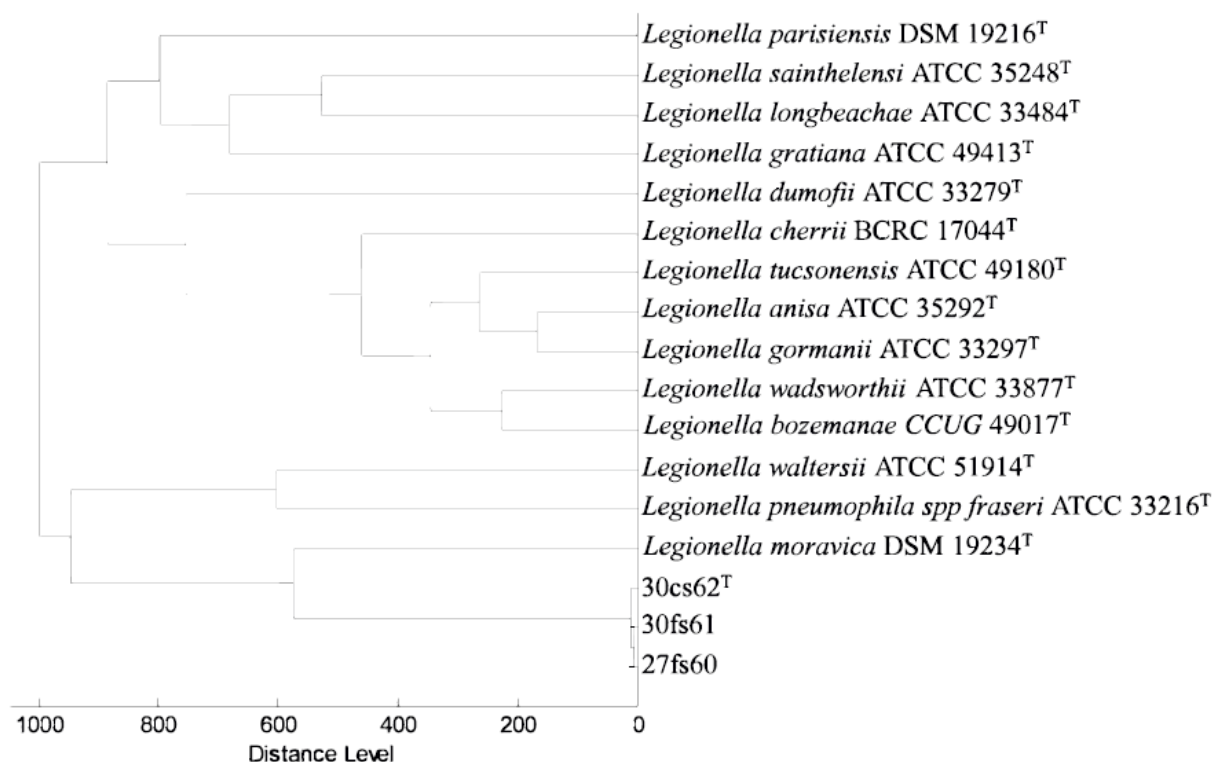


Fig. 4. Dendrogram based on whole-cell MALDI-TOF mass spectra of strains 27fs60, 30fs61 and 30cs62^T and *Legionella* reference strains present in the instrument database.

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, p-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. 2014; **42**: 613–616.
3. Fields BS, Benson RF, Besser RE. *Legionella* and legionnaires' disease: 25 years of investigation. *J Clin Microbiol* 2002; **15**: 506–526.
4. Rota MC, Caporali MG, Bella A, Scaturro M, Giannitelli S, Ricci ML. The legionellosis surveillance system in Italy: 2019 results. *Boll Epidemiol Naz* 2020; **1**: 32–38.
5. Centers for Disease Control and Prevention (CDC). *Legionella* (Legionnaires' Disease and Pontiac Fever) - Causes, How it Spreads, and People at Increased Risk. <https://www.cdc.gov/legionella/about/causes-transmission.html>. .
6. Correia AM, Ferreira JS, Borges V, Nunes A, Gomes B, Capucho R, 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, Simões MJ, 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. UNI EN ISO 19458:2006 - Water quality — Sampling for microbiological analysis. 2006. Geneva, Switzerland.
9. UNI EN ISO 11731:2017 Water quality - Enumeration of *Legionella*. 2017. Geneva, Switzerland.
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 MY, Lee KH, Yun YJ, Woo SY, et al. Application of RNA polymerase β -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, Grottola A, 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, Kulikov AS, et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012; **19**: 455–477.
15. Chen KT, Liu CL, Huang SH, Shen HT, Shieh YK, Chiu HT, 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, Sturrock S, 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, Zaslavsky L, et al. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 2016; **44**: 6614–6624.
18. Yoon SH, Ha S min, Lim J, Kwon S, Chun J. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek, Int J Gen Mol Microbiol* 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, Bealer K, 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**.
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**.

27. EUCAST. Antimicrobial Susceptibility Testing of *Legionella pneumophila*. 2021.

Chapter 7

*Summary of preliminary results obtained during the three
moths abroad in French National Reference Center of
Legionella, Hospices Civils of Lyon, France.*

Legionella are bacteria that live in fresh water where are mainly in the form of intracellular parasites of unicellular protozoan host: the free-living amoebas (about 15 different species of *Amoebas* [1]). The humans being is an accidental host of *Legionella*, which was infected by inhalation of the contaminated water droplets [1–3]. In this case, *Legionella* can reach the pulmonary alveoli where comes into contact with the macrophages. The illnesses caused by *Legionella* are called Legionellosis (LD).

The intracellular growth within *Amoebas* and macrophages was promoted by *Icm/Dot* type IV secretion system and the mechanism of infection is similar [4, 5]. The *Icm/Dot* secretion system is useful to deliver some macromolecules across kingdom barriers and is the first contact between *Legionella* and host-cells and phagosome biogenesis [6, 7].

Since the *Legionella* is an intracellular bacterium that can caused LD in human, antimicrobial therapy must be able to reach the intracellular environment with the appropriate therapeutic concentrations (e.g., erythromycin, rifampicin, tetracyclines, macrolides and fluoroquinolones). At the moment azithromycin or levofloxacin are the most used [8].

Unfortunately, the laboratories do not routinely perform antibiotic susceptibility test (AST) on *Legionella* isolates for some reasons, therefore the introduce of antibiotic often is empirical. The reasons are i) the difficulty of the *Legionella* isolation from clinical samples; ii) the absence of a specific and standardized AST methods; iii) the absence of ‘epidemiological cut-off’ (ECOFF) values for *Legionella* spp.. Regarding the EUCAST cut-off, the only value defined in *vitro* as minimum inhibitory concentration (MIC) threshold were obtained by combining the results of various studies that allows the discrimination of wild type (WT) strains compared to those with established resistance mechanisms; iv) the correlation between in *vitro* results and clinical outcome is unclear [9, 10].

Although rare cases of acquired antimicrobial resistance in *Legionella* were reported [10–14], some treatment based on macrolides and/or fluoroquinolones fail in management of patients with LD [15–17]. Antibiotic resistance in *Lp* can caused by different reasons, for example, resistance to ciprofloxacin is due to a single point mutation in *gyrA* gene [18, 19], instead the reduced susceptibility to azithromycin (MICs 0.125–2 mg/L) is due to the presence of a *LpeAB* efflux pump system [10, 20–22]. The selection of these antibiotic resistance in the bacteria is not only linked to the clinical setting but also to the natural environment. The latter has a fundamental role in the development and spreading of antibiotic resistance [23, 24].

Unfortunately, antimicrobial susceptibility of WT environmental isolates is usually not performed during the routine analysis [25]. During my abroad period of three months in France (Sep-Dic 2022), at the French National Reference Center of *Legionella*, Hospices Civils of Lyon, I performed different test to analyze some characteristics of the three novel *Legionella* species (*L. bononiensis* (30cs62^T), 8cVS16^T and 31fl33^T) discovered in Bologna area during my PhD project. The test carried out including the infectivity and antibiotic susceptibility for these strains.

In particular, the infection experiments in three different types of cells were performed: *A. castellanii*, *A. polyphaga* and macrophage U937 cell lines. Moreover, the broth microdilution (BMD)-based AST methodology on the strains was tested.

Regarding the *Legionella* infection, the amoebae plate test (APT) was performed on the Amoebae [26]. Briefly, the number of cells present on microliter in a flask of amoebae is counted and through the formula⁴

$$V = \frac{N}{C} = \frac{\text{no. of plate} * 3.75 * 10^6}{\text{no. amoebae}/\mu\text{L}}$$

the volume needed to obtain the correct number of cells to be taken to make an amoebae layer in a plate of Buffered Charcoal Yeast Extract plus L-Cysteine (BCYE Cys+) was calculated.

Thereafter those volume is centrifugated, the supernatant removed, and the pellet diluted with 1.5 mL of Page's Amoeba Saline (PAS) solution for each plate needed. The solution is placed over the entire surface of the plate. During the time that the layer dry (at least one hour), the bacterial suspension of 0.5 McFarland is prepared and diluted from 10⁰ to 10⁻⁷. A single 3μL drop of each dilution is distributed in a horizontal line in both the plate with and without (control plate) amoeba layer.

For each strain, one couple of the plate was utilized. To perform the test each strain was flanked by the respective reference strain and two control strains (*L. pneumophila* (Lp) paris serogroup (SG) 1 (CIP107629) (positive control) and *Lp* paris SG 1 $\Delta dotA$ ($\Delta dotA$) (negative control)).

In detail:

- Group 1: *L. bononiensis* (30cs62^T), *L. quateirensis* (ATCC 49507^T), *Lp* paris SG1 and $\Delta dotA$;
- Group 2: 8cVS16^T, *L. anisa* (ATCC 35292), *Lp* paris SG1 and $\Delta dotA$;

⁴ 3.75×10⁶ is the number of amoeba cells needed to create a good layer for one plate.

- Group 3: 31f133^T, *L. feeleeii* (ATCC 35072), *L. donaldsonii* (NCTC 13292), *Lp paris* SG1 and $\Delta dotA$.

The amoebae tested (*A. castellanii* and *A. polyphaga*) were infected by 8cVS16^T, 31f133^T, *L. feeleeii*, *L. donaldsonii*, and *Lp paris* SG1.

With regard to the other cell tested for *Legionella* infectivity: macrophage, the procedure performed was carried out according to Guillemot et al. [27].

TECAN Infinite M200 Pro multiplate reader was utilized to check the growth/infectivity of the strains on the macrophages. The instrument exploits the emission of fluorophores. For this reason, all the strains that were already utilized in the APT experiment were transformed by electroporation with plasmids pXDC50 (m-cherry) and pXDC31 (GFP). The macrophage utilized in this experiment was obtained by human monocyte U937 cell line. The latter was cultured in a flask with Roswell Park Memorial Institute (RPMI) medium containing 10% fetal bovine serum (FBS) at 37°C and 5% CO₂ in a humidified atmosphere. Forty-eight hours before the infection, these cells were differentiated with 50 ng/mL phorbol 12- myristate 13-acetate (PMA) in a 96-wells microplate (1.5×10⁵ cells/well). For each strain, a bacterial solution was made, and the optical density at 600nm (OD₆₀₀) was calculated. To infect the macrophages, a Multiplicity of infection (MOI) of 10, i.e., 10 times as many bacteria as the number of cells (1.5×10⁶ bacteria/well) was used. The strains are diluted in CO₂ Independent Medium to reach MOI of 10 and they were incubated two hours at 37°C to prepare the bacteria for the infection. In the mean while the macrophages were washing one time with CO₂ Independent Medium+10% FBS. The strains (200μL/well) were added to each well and the microplates were centrifuged for 5 min at 500 g to favor the interactions between the bacteria and the macrophage and incubated for 1 h at 37°C and 5% CO₂ in a humidified atmosphere. The cells were washed 2 times with CO₂ Independent Medium+10%FBS+chloramphenicol+Isopropyl β-d-1-thiogalactopyranoside (IPTG)+glutamine in order to remove the bacteria unincorporated in the cells. The growth of the strains was followed at λ= 610 nm for m-cherry and λ= 587 nm for GFP on a TECAN Infinite M200 Pro multiplate reader until the transition between the exponential and the stationary growth phases for five days. To check if the bacterial inoculums contain the vitality strains the inoculums were cultured on BCYE Cys++chloramphenicol+IPTG (7.5*10³ bacteria/100μL) by easySpiral dilute (Interscience, Paris, France) that is an automatic diluter and plater of petri dishes. The macrophages were infected by 8cVS16^T, 31f133^T, *L. feeleeii*, *L. donaldsonii*, and *Lp paris* SG1. The analysis of the results obtained are in progress.

In the wake of the innovation of my research project and based on wanting to characterize the newly discovered species as much as possible, we decided to try to understand the antibiotic sensitivity of these strains. To determine the Minimum inhibitory concentrations (MICs) the AST test was performed by BMD test using buffered yeast extract broth supplemented with 1g/L α -ketoglutarate, 0.25 g/L L-Cysteine, and 0.4 g/L ferric pyrophosphates (α BYE), without charcoal to avoid the inactivation of antibiotic [9], on 96-well microtiter plates. Briefly, the bacterial suspension (our three novel *Legionella* strains and *Lp paris* SG1 as positive control) of 0.5 McFarland (5×10^7 bacteria/mL) is prepared and diluted 1:100 in α BYE. The antibiotic tested were azithromycin, erythromycin, ciprofloxacin, levofloxacin, moxifloxacin, doxycycline, rifampicin.

A volume of 80 μ L for each antibiotic was added to the first well of the 96-well microplates and then diluted in 2-fold decreasing concentration. The final concentration of the antibiotic, before the dilution, is:

- Azithromycin: 16 mg/L;
- Erythromycin: 16 mg/L;
- Ciprofloxacin: 0.512 mg/L;
- Levofloxacin: 0.512 mg/L;
- Moxifloxacin: 0.512 mg/L;
- Doxycycline: 64 mg/L;
- Rifampicin: 0.064 mg/L.

One hundred sixty microliters of bacterial suspension were added in each well, for a total of 8×10^4 *Legionella* cells. The negative control consisting in 200 μ L of α BYE and the positive in 160 μ L of bacterial suspension and 40 μ L of α BYE. The plates were incubated at 35°C for 40h. The MICs are determined considering the first well without the presence of the “bottom pellet” in the wells, this result is calculated as MIC $\pm$ 1 well.

Each isolate was tested at least three times to verify the results.

Considering the fact that there is an absence of ECOFF values for *Legionella* spp. and usually the non-*pneumophila* *Legionella* species (n-pL) were not tested, it is difficult to compare the result obtain with the three novel species with regard to *Lp* results.

Based on this, the results found are as follows: *L. bononiensis* strain shows a high sensitivity to azithromycin (MIC $\leq$ 0.008 mg/L). To the best of our knowledge, until know, this species is the first n-pL that shows this high sensitivity to azithromycin. In literature, only other few *Lp*: *Lp* SG2-15

MIC ≤ 0.008 mg/L, *Lp* SG1 MIC 0.001 mg/L, and *Lp* MIC 0.008 mg/L, showed the high sensitivity to azithromycin [10, 25, 28, 29].

Instead, 8cVS16^T and 31fI33^T strains showed a low sensitivity to azithromycin (MIC 1 mg/L). To better explain the low sensitivity to azithromycin, the possible implication of efflux pump activity in that phenotype of 8cVS16^T and 31fI33^T strains was investigated [25].

The AST procedure previously described was performed, with the addition of reserpine/resazurin microtiter assay to observe the quantitative inhibition of the efflux pump activity. Reserpine is an efflux pump inhibitor (EPI) of resistance nodulation cell division (RND) efflux pump systems [25]. Resazurin is a pH and oxidoreduction indicator with a blue to purple color (at pH > 6.5) and it can be irreversibly reduced to the pink-colored.

Concisely, reserpine was prepared in three different concentrations: 1, 10 and 20 mg/L in dimethylsulfoxide (DMSO). Resazurin was prepared dissolved 20 mg in 100 mL of sterile distilled water (0.02 wt%).

The determination of BMD MIC for Azithromycin was carrying out in two separate rows of the 96-wells microtiter plate. The first row without, the second with reserpine.

The controls consisting of:

- Negative control: 200 µL of αBYE;
- Negative control+reserpine: 199 µL of αBYE + 1 µL of reserpine;
- Positive control: 160 µL of bacterial suspension + 40 µL of αBYE;
- Positive control+reserpine: 160 µL of bacterial suspension + 40 µL of αBYE + 1 µL of reserpine;
- Control with DMSO: 160 µL of bacterial suspension + 40 µL of αBYE + 1 µL of DMSO, to check the toxicity of the DMSO on the strains.

Each plate was incubated for 72h at 35°C in a humidified thermostat.

After this time 30 µL of resazurin solution was added to each well [30] and the plates were incubated again at 35 °C for 24 h.

The changing of the color from blue to pink indicated the growth of *Legionella*. This time the MIC was quantified as the lowest drug concentration where there was the change in the color.

Concerning the two strains under analysis (8cVS16^T and 31fI33^T), their showed a low effect with reserpine (azithromycin MIC 0.5 mg/L), suggesting a possible role of efflux pump in that phenotype.

Regarding the presence of DMSO in reserpine solution, this compound did not show any toxicity for the strains during AST test.

References

1. Boamah DK, Zhou G, Ensminger AW, O'Connor TJ. From many hosts, one accidental pathogen: The diverse protozoan hosts of *Legionella*. *Front Cell Infect Microbiol* 2017; **7**.
2. Allegra S, Leclerc L, Massard PA, Girardot F, Riffard S, Pourchez J. Characterization of aerosols containing *Legionella* generated upon nebulization. *Sci Rep* 2016; **6**.
3. Schoen ME, Ashbolt NJ. An in-premise model for *Legionella* exposure during showering events. *Water Res* 2011; **45**: 5826–5836.
4. Otto GP, Wu MY, Clarke M, Lu H, Roger Anderson O, Hilbi H, et al. Macroautophagy is dispensable for intracellular replication of *Legionella pneumophila* in *Dictyostelium discoideum*. *Mol Microbiol* 2004; **51**: 63–72.
5. Hägele S, Köhler R, Merkert H, Schleicher M, Hacker J, Steinert M. *Dictyostelium discoideum*: A new host model system for intracellular pathogens of the genus *Legionella*. *Cell Microbiol* 2000; **2**: 165–171.
6. Watarai M, Derre I, Kirby J, Gowney JD, Dietrich WF, Isberg RR. *Legionella pneumophila* is internalized by a macropinocytotic uptake pathway controlled by the *dot/icm* system and the mouse Lgn1 locus. *J Exp Med* 2001; **194**: 1081–1095.
7. Hilbi H, Segal G, Shuman HA. *Icm/Dot*-dependent upregulation of phagocytosis by *Legionella pneumophila*. *Mol Microbiol* 2001; **42**: 603–617.
8. Dedicoat M, Venkatesan P. The treatment of Legionnaires' disease. *J Antimicrob Chemother*. 1999; **43**: 747–752.
9. Portal E, Sands K, Portnojs A, Chalker VJ, Spiller OB. *Legionella* antimicrobial sensitivity testing: Comparison of microbroth dilution with BCYE and LASARUS solid media. *J Antimicrob Chemother* 2021; **76**: 1197–1204.
10. EUCAST. Antimicrobial Susceptibility Testing of *Legionella pneumophila*. 2021. Vaxjo, Sweden; 2021.
11. Vandewalle-Capo M, Massip C, Descours G, Charavit J, Chastang J, Billy PA, et al. Minimum inhibitory concentration (MIC) distribution among wild-type strains of *Legionella pneumophila* identifies a subpopulation with reduced susceptibility to macrolides owing to efflux pump genes. *Int J Antimicrob Agents* 2017; **50**: 684–689.
12. Wilson RE, Hill RLR, Chalker VJ, Mentasti M, Ready D. Antibiotic susceptibility of *Legionella pneumophila* strains isolated in England and Wales 2007–17. *J Antimicrob Chemother* 2018;

73: 2757–2761.

13. Koshkolda T LC. Antibiotic susceptibility of clinical *Legionella pneumophila* serogroup 1 strains isolated in Germany. *J Antimicrob Chemother* 2018; **73**: 541–542.
14. Natås OB, Brekken AL, Bernhoff E, Hetland MAK, Löhr IH, Lindemann PC. Susceptibility of *Legionella pneumophila* to antimicrobial agents and the presence of the efflux pump *LpeAB*. *J Antimicrob Chemother* 2019; **74**: 1545–1550.
15. Jia X, Ren H, Nie X, Li Y, Li J, Qina T. Antibiotic Resistance and Azithromycin Resistance Mechanism of *Legionella pneumophila* Serogroup 1 in China. *Antimicrob Agents Chemother* 2019; **63**: 1–11.
16. Shadoud L, Almahmoud I, Jarraud S, Etienne J, Larrat S, Schwebel C, et al. Hidden Selection of Bacterial Resistance to Fluoroquinolones In Vivo: The Case of *Legionella pneumophila* and Humans. *EBioMedicine* 2015; **2**: 1179–1185.
17. Ceccarelli, G.; Venditti, M.; Scaturro, M.; Ricci ML. Clinical symptoms and treatment of legionellosis. In: Moran-Gilad, J., Gibbs, R.E. E. (ed).2020. Caier Academic Press, Poole, UK, pp 173–186.
18. Bruin JP, Koshkolda T, IJzerman EPF, Lück C, Diederer BMW, Den Boer JW, et al. Isolation of ciprofloxacin-resistant *Legionella pneumophila* in a patient with severe pneumonia. *J Antimicrob Chemother* 2014; **69**: 2869–2871.
19. Schrock J, Hackman BA, Plouffe JF. Susceptibility of ninety-eight clinical isolates of *Legionella* to macrolides and quinolones using the E-test. *Diagn Microbiol Infect Dis* 1997; **28**: 221–223.
20. Vandewalle-Capo M, Massip C, Descours G, Charavit J, Chastang J, Billy PA, et al. Minimum inhibitory concentration (MIC) distribution among wild-type strains of *Legionella pneumophila* identifies a subpopulation with reduced susceptibility to macrolides owing to efflux pump genes. *Int J Antimicrob Agents* 2017; **50**: 684–689.
21. Massip C, Descours G, Ginevra C, Doublet P, Jarraud S, Gilbert C. Macrolide resistance in *Legionella pneumophila*: The role of *LpeAB* efflux pump. *J Antimicrob Chemother* 2017; **72**: 1327–1333.
22. Mallegol J, Fernandes P, Melano RG, Guyarda C. Antimicrobial activity of solithromycin against clinical isolates of *Legionella pneumophila* serogroup 1. *Antimicrob Agents Chemother* 2014; **58**: 909–915.
23. Ottaviani D, Leoni F, Rocchegiani E, Mioni R, Costa A, Virgilio S, et al. An extensive

- investigation into the prevalence and the genetic and serological diversity of toxigenic *Vibrio parahaemolyticus* in Italian marine coastal waters. *Environ Microbiol* 2012; **15**: 1377–1386.
24. Davies J. Origins and evolution of antibiotic resistance. *Microbiologia* 1996; **12**: 9–16.
 25. Cocuzza CE, Martinelli M, Perdoni F, Giubbi C, Vinetti MEA, Calaresu E, et al. Antibiotic susceptibility of environmental *Legionella pneumophila* strains isolated in northern Italy. *Int J Environ Res Public Health* 2021; **18**: 0–9.
 26. Albers U, Reus K, Shuman HA, Hilbi H. The amoebae plate test implicates a paralogue of IpxB in the interaction of *Legionella pneumophila* with *Acanthamoeba castellanii*. *Microbiology* 2005; **151**: 167–182.
 27. Guillemot J, Ginevra C, Allam C, Kay E, Gilbert C, Doublet P, et al. TNF- α response in macrophages depends on clinical *Legionella pneumophila* isolates genotypes. *Virulence* 2022; **13**: 160–173.
 28. Portal E, Sands K, Portnojs A, Chalker VJ, Spiller OB. *Legionella* antimicrobial sensitivity testing: Comparison of microbroth dilution with BCYE and LASARUS solid media. *J Antimicrob Chemother* 2021; **76**: 1197–1204.
 29. Miyashita N, Kobayashi I, Higa F, Aoki Y, Kikuchi T, Seki M, et al. In vitro activity of various antibiotics against clinical strains of *Legionella* species isolated in Japan. *J Infect Chemother* 2018; **24**: 325–329.
 30. Martin A, Camacho M, Portaels F, Palomino JC. Resazurin Microtiter Assay Plate Testing of *Mycobacterium tuberculosis* Susceptibilities to Second-Line Drugs: Rapid, Simple, and Inexpensive Method. *Society* 2003; **47**: 3616–3619.

4. Conclusions

The knowledge of the natural and man-made water environment other than the events that promote changes in the ecological niche where *Legionella* live, can help to prevent their spread in man-made environment and contain the occurrence of cases, clusters or outbreaks. For this reason, an optimal environmental surveillance and a proper risk assessment plan are necessary to understand the *Legionella* contamination during the time. The surveillance and risk assessment plan are not exhaustive without the correct identification and characterization of the isolates presents in the environment. In this context play a fundamental role the workflow methods presented in this dissertation that could help to characterize and recognise other non-*pneumophila* *Legionella* species, their features, rates of infection and antibiotic susceptibility. This dissertation would point out how the research and studies on environments are the best way to control and prevent the impact of *Legionella* infection on public health.

Suggestions for future work

This dissertation would represent only the first step in the knowledge of the wide world of non-*pneumophila* *Legionella* species. Other studies should be conducted on the taxonomy characteristics (e.g., morphology, including flagellar structures, growth tests at various temperatures and on culture media, biochemical, and antibiotic susceptibility tests) of other novel species found during these years.

Regarding the three novel species presented in this dissertation, *L. bononiensis* does not infect either *A. castellanii* and *A. polyphaga*, as well as macrophages U937 cell lines, for this reason, other hosts could also be investigated. In addition, due to its high sensitivity to azithromycin ($\text{MIC} \leq 0.008$ mg/L), further studies will be conducted in order to evaluate the exacted azithromycin MIC. Regarding 8cVS16^T and 31fI33^T strains, the low sensitivity to azithromycin (MIC 1 mg/L) and the low effect induce to reserpine treatment (decrease in azithromycin MIC to 0.5 mg/L), suggesting a possible role of efflux pump in that phenotype, that would be further investigated thought a knockout genes model involved on efflux pump.

Appendix A:

Supplementary material - Chapter 6

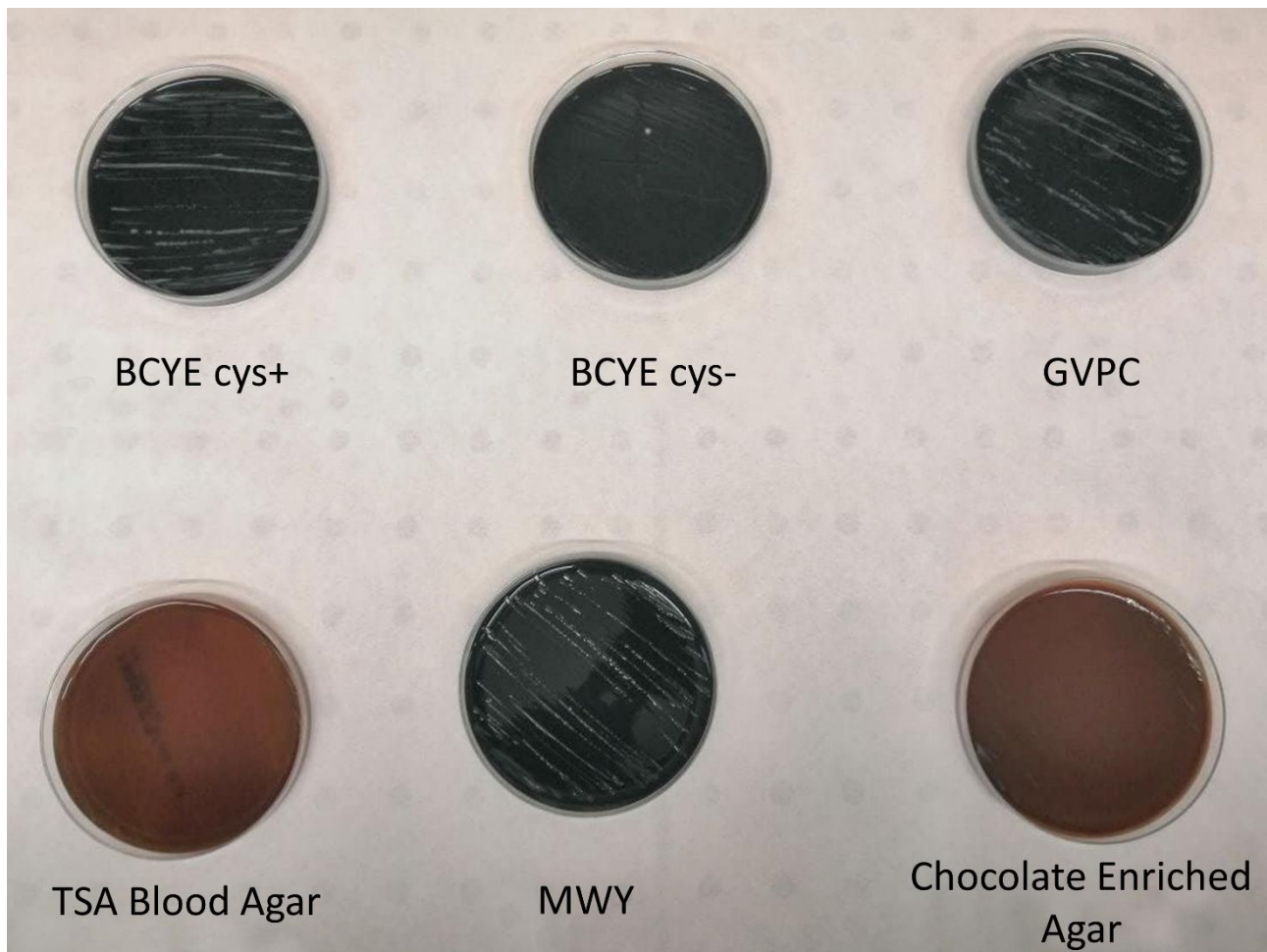


Figure S1. The 30cs62T strain growth on different medium: BCYE Cys+ and Cys-, GVPC, TSA Blood Agar, MWY and Chocolate Enriched Agar.

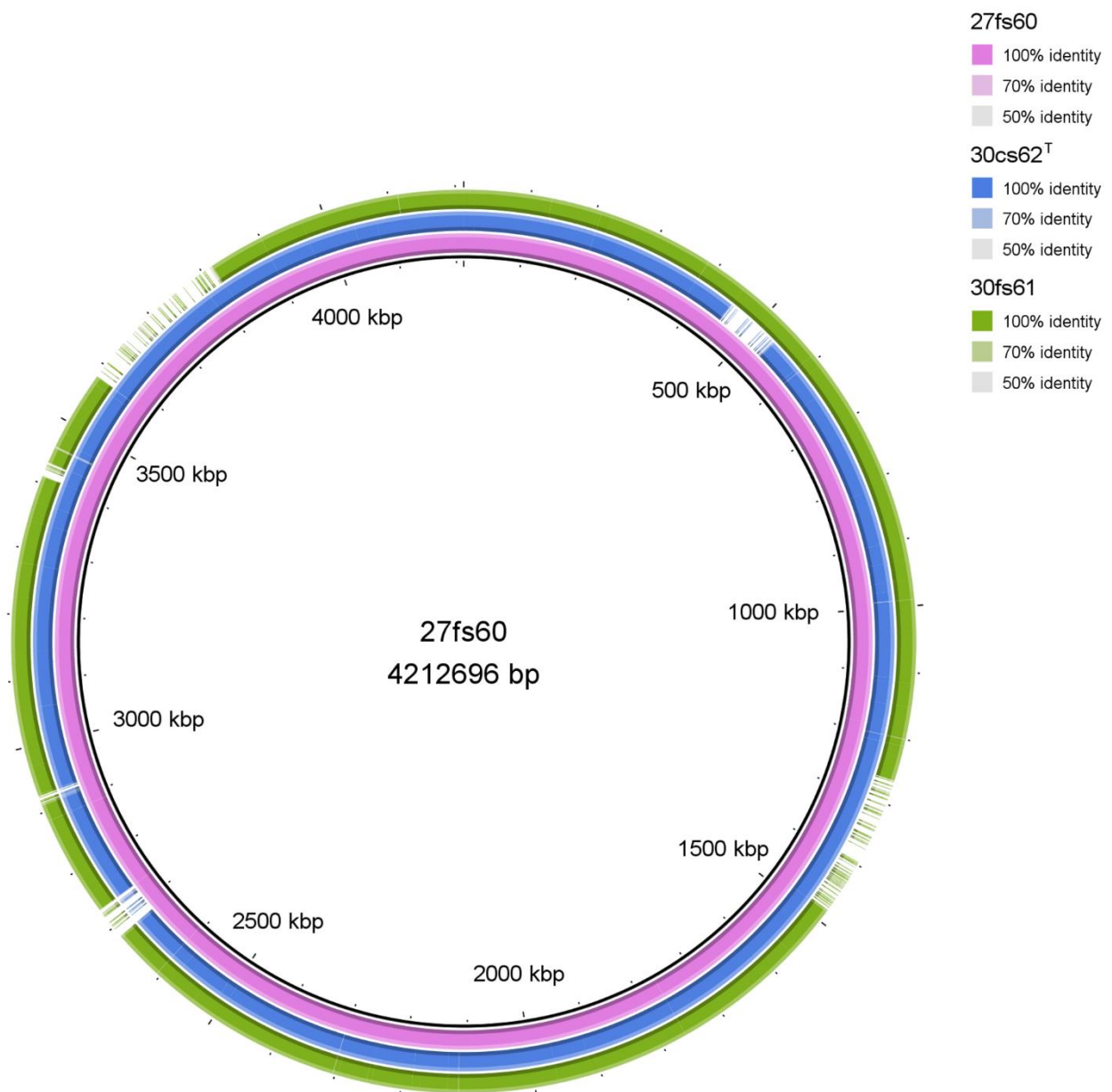


Figure S2. Genome's alignment of 27fs60, 30fs61 and 30cs62T strains, provided by BLAST Ring Image Generator (v. 0.95) software.

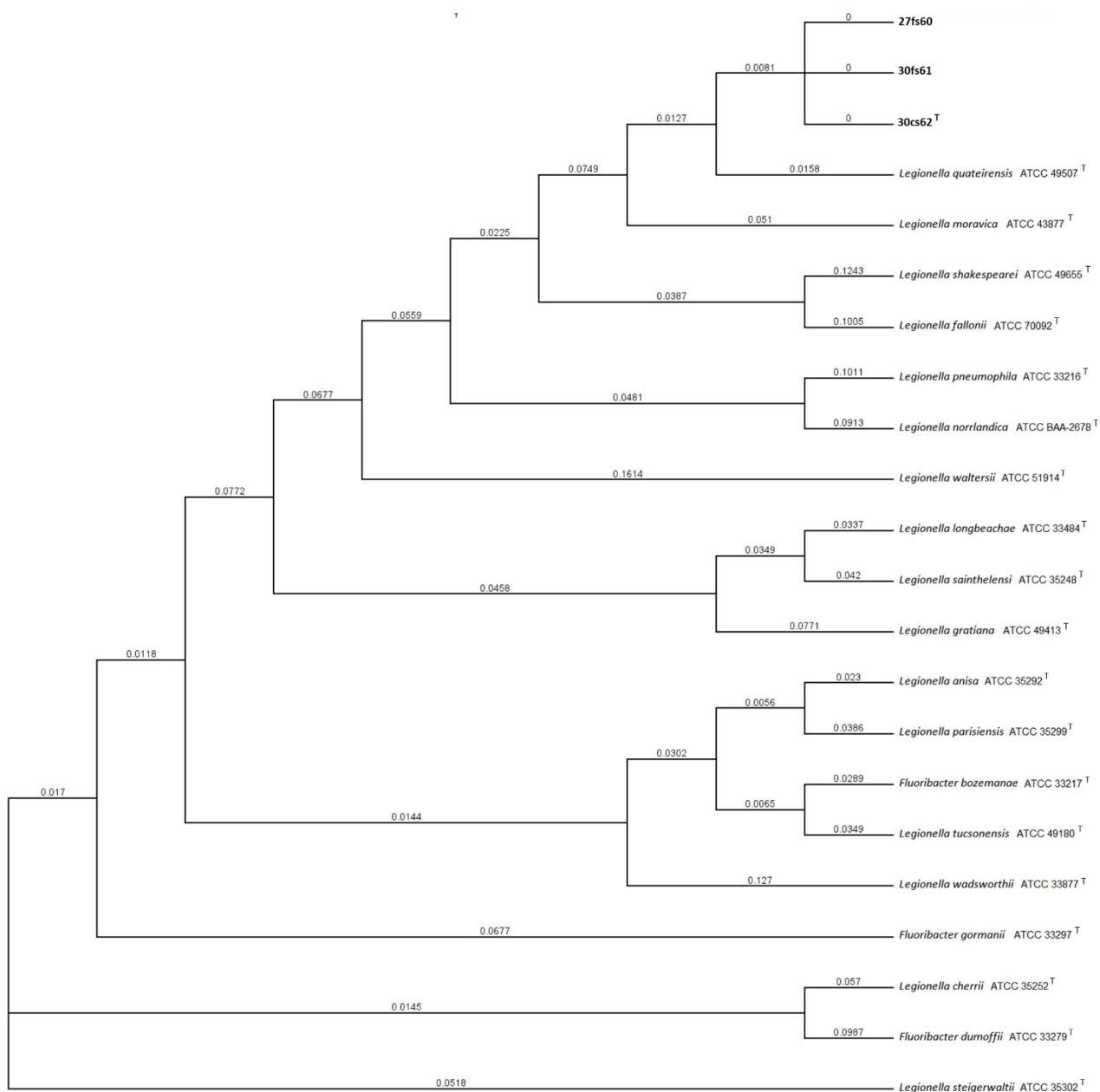


Figure S3: Phylogenetic tree based on *mip* gene of the three strains (27fs60, 30fs61 and 30cs62T) and closely related species of the genus *Legionella*. Branch lengths are transformed to be equal like a cladogram. Branch labels display the substitutions per site.

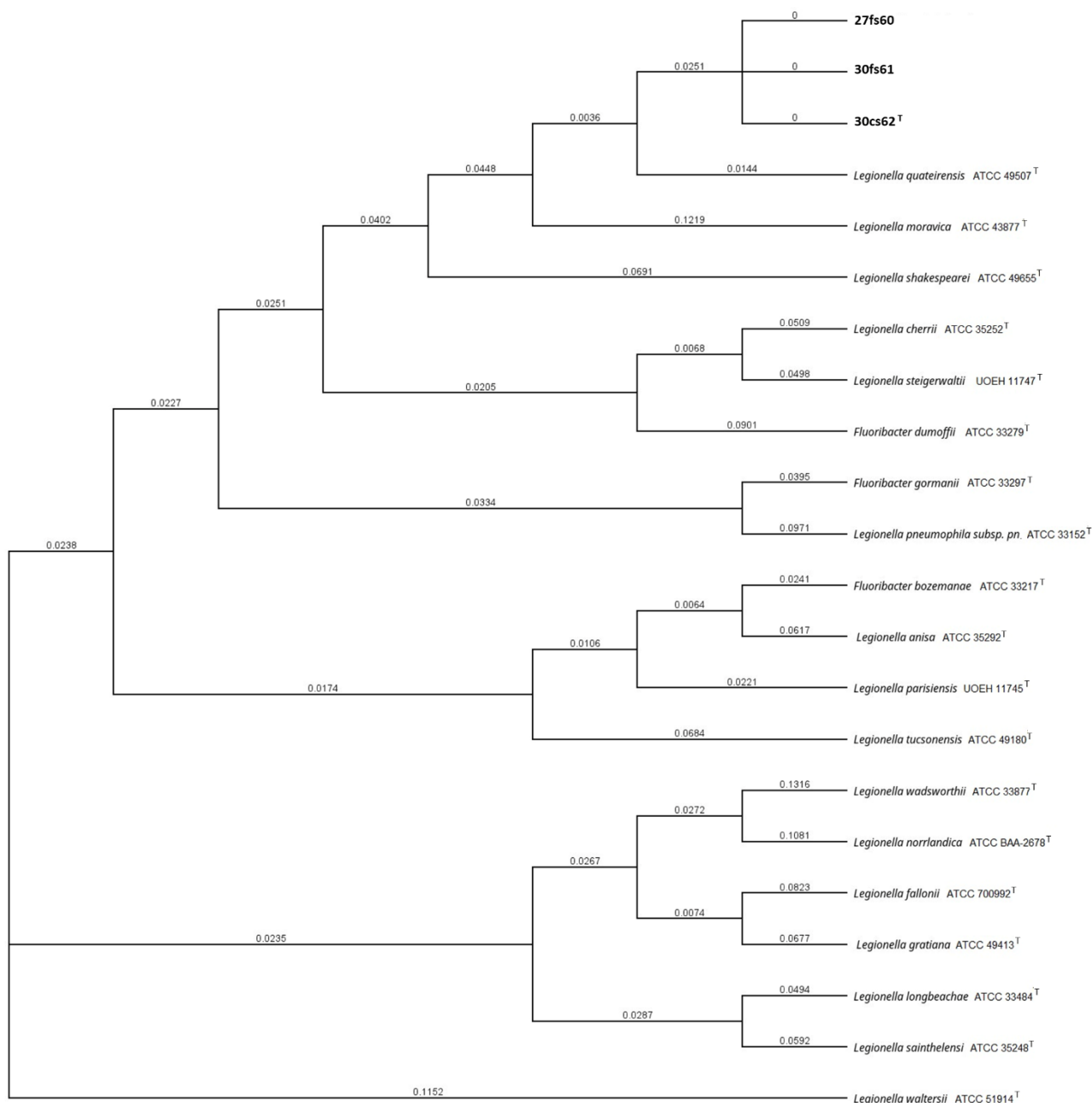


Figure S4: Phylogenetic tree based on *rpoB* gene of the three strains (27fs60, 30fs61 and 30cs62T) and closely related species of the genus *Legionella*. Branch lengths are transformed to be equal like a cladogram. Branch labels display the substitutions per site.

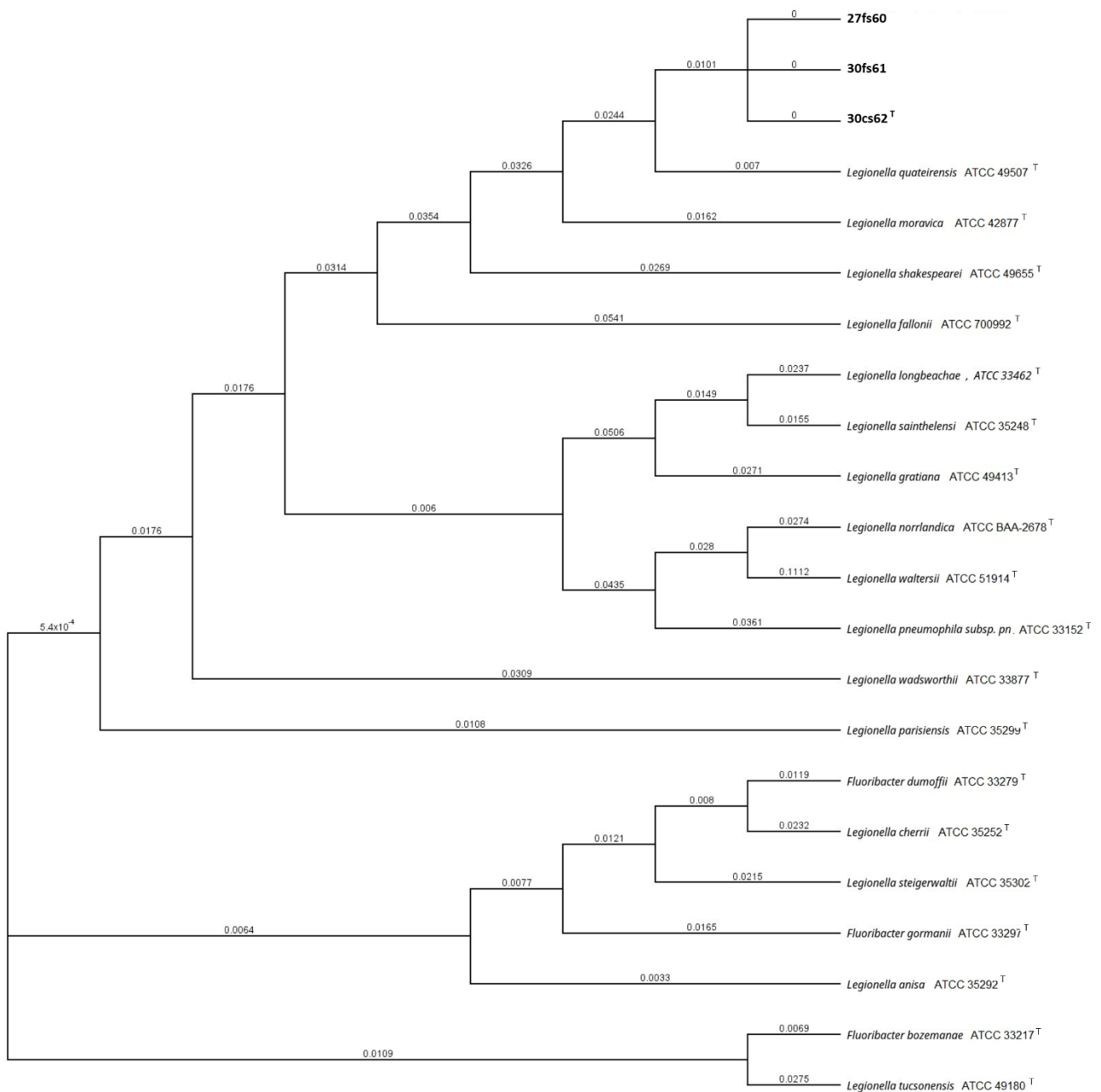


Figure S5: Phylogenetic tree based on *rnpB* gene of the three strains (27fs60, 30fs61 and 30cs62^T) and closely related species of the genus *Legionella*. Branch lengths are transformed to be equal like a cladogram. Branch labels display the substitutions per site.

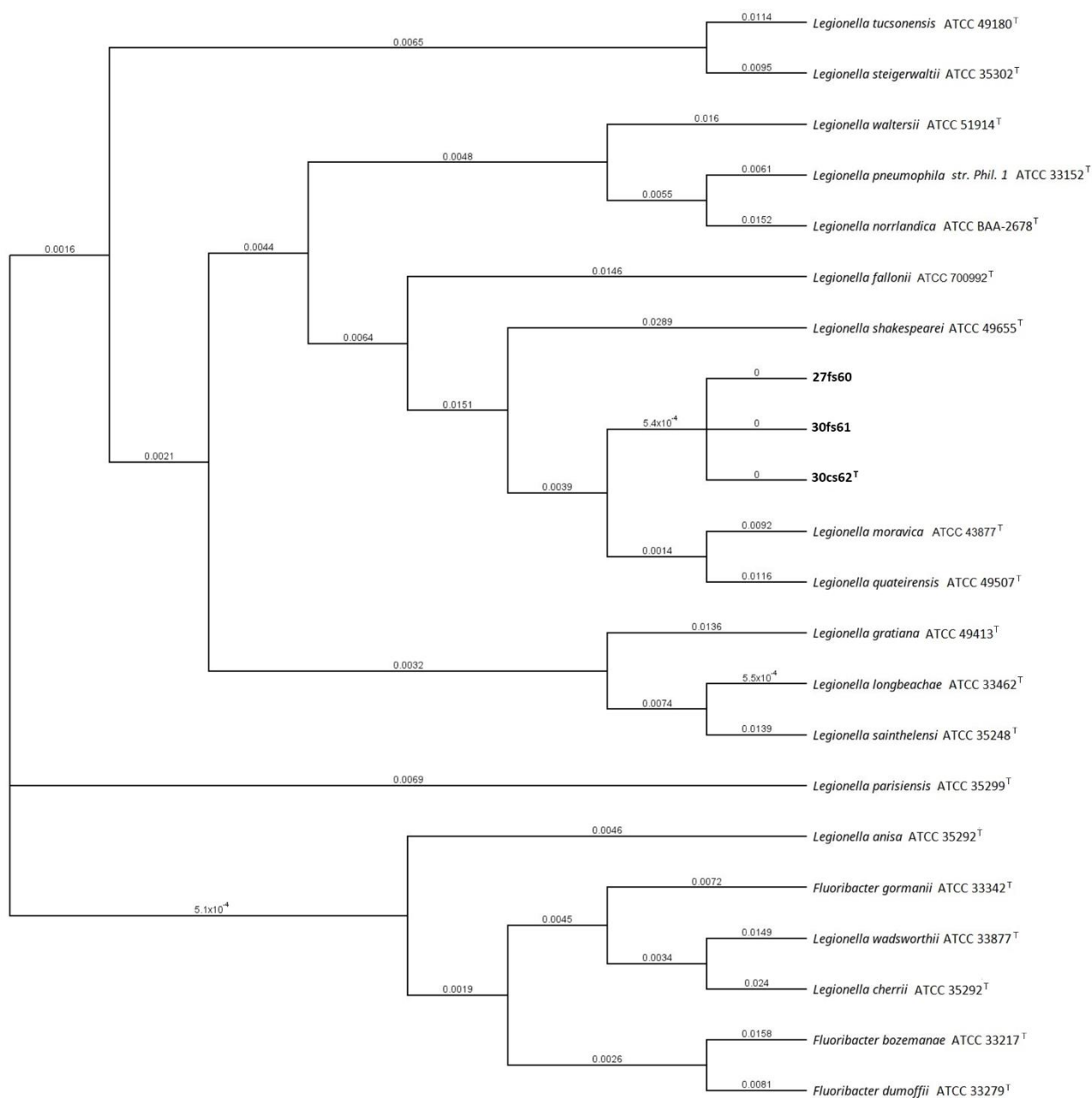


Figure S6: Phylogenetic tree based on 16S rRNA gene of the three strains (27fs60, 30fs61 and 30cs62T) and closely related species of the genus *Legionella*. Branch lengths are transformed to be equal like a cladogram. Branch labels display the substitutions per site.

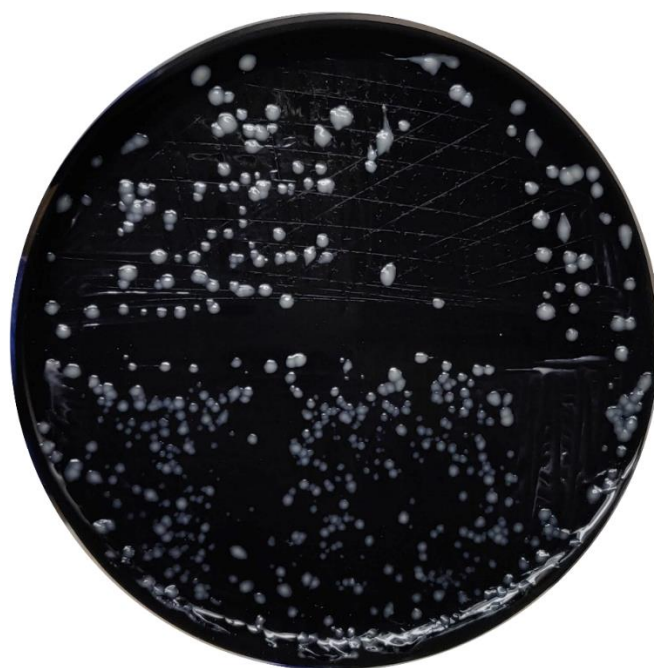
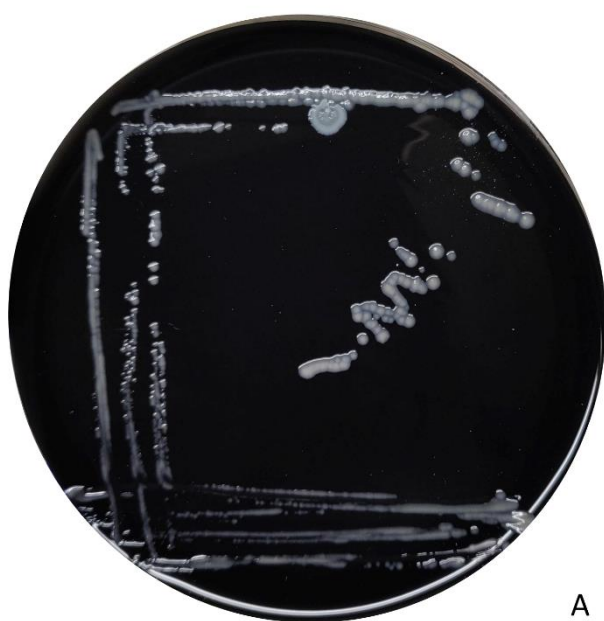
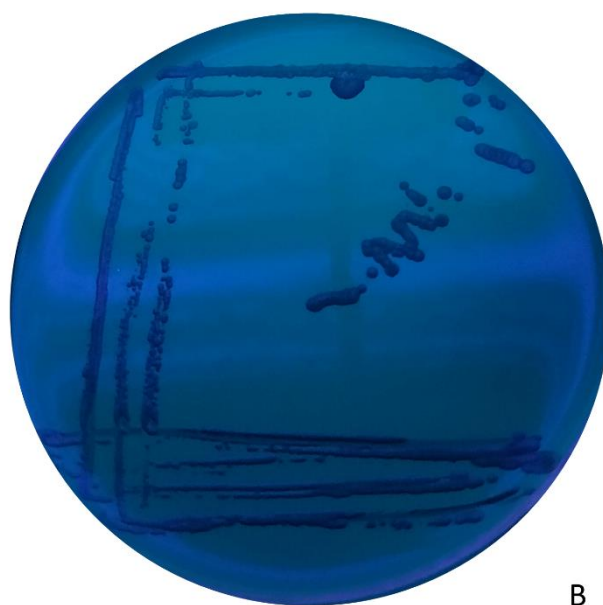


Figure S7: Strain 30cs62T on BCYE Cys+ medium.



A



B

Figure S8. Strain 30cs62T growth for 3 days at 35°C with 2.5% CO₂ on BCYE Cys+ (A) and under Woods lamp (365 nm) (B).

Table S1. Analysis of virulence genes presence in three strains: presence (+), absence (-) and sequences percentage of identity (%) respect to *L. pneumophila philadelphia* Serogroup 1 (ATCC 33152^T) and *L. quateirensis* (ATCC 49507^T)

List of virulence genes	27fs60		30fs61		30cs62 ^T	
	<i>L. pneumophila</i>	<i>L. quateirensis</i>	<i>L. pneumophila</i>	<i>L. quateirensis</i>	<i>L. pneumophila</i>	<i>L. quateirensis</i>
<i>htpB</i>	+	-	+	-	+	-
	87.81%		87.81%		87.81%	
<i>momps</i>	+	+	+	+	+	+
	68.05%	87.10%	68.05%	87.10%	68.05%	87.10%
<i>mip</i>	+	-	+	-	+	-
	83.59%		83.59%		83.59%	
<i>cpxA</i>	+	+	+	+	+	+
	72.77%	92.85%	72.77%	92.85%	72.77%	92.85%
<i>cpxR</i>	+	+	+	+	+	+
	82.41%	91.70%	82.41%	91.70%	82.41%	91.70%
<i>flaA</i>	+	+	+	+	+	+
	76.61%	88.89%	76.61%	88.89%	76.61%	88.89%
<i>lvrA1</i>	+	-	+	-	+	-
	80.12%		80.12%		80.12%	
<i>wzm/rfbA2</i>	-	-	-	-	-	-
<i>enhC</i>	-	-	-	-	-	-
<i>lvrB</i>	-	-	-	-	-	-
<i>lvrC</i>	-	-	-	-	-	-
<i>lvrE</i>	-	-	-	-	-	-
<i>pilD</i>	-	-	-	-	-	-
<i>pilE2</i>	-	-	-	-	-	-
<i>rtxA</i>	-	-	-	-	-	-

Table S2. Genome features of 27fs60,30fs61 and 30cs62T strains in comparison with *L. quateirensis* (ATCC 49507^T) and *L. pneumophila philadelphia* Serogroup 1 (ATCC 33152^T)

Strains	No. of bases (bp)	No. of coding DNA Sequence (CDS)	No. of genes	No. of rRNA	No. of tRNA	No. of tmRNA
27fs60	4,212,696	3,606	3,651	3	41	1
30fs61	3,710,901	3,204	3,245	4	37	-
30cs62 ^T	4,137,742	3,540	3,584	4	39	1
<i>L. quateirensis</i>	4,281,523	3,608	3,660	9	42	1
<i>L. pneumophila</i>	3,397,754	3,020	3,073	9	43	1

Table S3. Comparison between longer and shorter genomes and list of missing genes among 27fs60,30fs61, 30cs62^T strains and related type strains *L. quateirensis* (ATCC 49507^T) and *L. pneumophila philadelphia* Serogroup 1 (ATCC 33152^T)

Genome comparison	Total number of annotated genes	List of missing genes in shorter genome
27fs60 vs 30cs62 ^T	1661 vs 1610	<i>umuC_3, prpB, prpC, fadD, COQ3, sad, sqr, ybgC, rlmN, hisS, katG1, csrA_8, artJ_7, yedK_3, ldh_3, rplI, nudL, der, tolQ, dtpA_6, algU, lpxH, pdhA, tyrS, bepA_4, lexA_4, pdhB, mhpC, ghrB, clcA_3, add_3, ynhG, gdh, pcsA, rpsF, folK, flgE_3, ygfZ, sdhE, ndk, rpsR, spmB, exbD, spmA, ppsR, hemL, fadJ, rubA1, pdhC, rodZ, tolB, yveA_3, mepM, queF, pabB, prpD, bamB, mnmC, minC</i>
30cs62 ^T vs 30fs61	1610 vs 1501	<i>bpoC, ydfG, dsbA, dnaJ_3, arcB_6, hemE, gcvPB, ubil, czcA_5, ppk, ubiH, artQ, pepP, ssrA, queA, yajC, ttgC, gmk, ftsB, argG, glnQ, tam_4, spoT, murU, lptD, fni, zur, pksJ, bprV, rhaS_1, amgK, ftsN, mmsB, rlmB, gcvPA, hchA, rph, secD, tktA, recG, secF, queE, ibpB, bauC, artP, folB, uvrB, prlC, macA, rpiA, pykA, rhaS_2, eno, arcB_8, btuD_5, udp, uup, tal, rutD_3, mdtA_4, arcB_7, fieF, rbsC, kdpB_3, mdtC, polA, gcvH, mvaA, lpxD_4, argR, ynjF, pgk, pcpC, aaeB, surA, ycaC_3, epmB, adeP, mdIC, aatA, efp, fliY, metK, engB, argH, grpE, rpoZ, tqsA, aroH, rutD_2, pal_3, rutD_1, yhhQ, tam_5, mdly, yceF, alsD, pdxA, dfrA, argS, yabJ, cc4, gcvT, esiB_4, gap, proC, dltA_4, rnr, vldW, frcA, fabF_4, bepF_3, mdtN, expZ, smc_6, hoxN, rayT_5, iolG_4, dnaK, czcC_5, zapA, ycdF, hin_3, dapB, yffB, msrB, alsS, oprB, per1, acyP, czcA_6</i>
27fs60 vs 30fs61	1661 vs 1501	<i>bpoC, ydfG, umuC_3, dsbA, dnaJ_3, arcB_6, hemE, gcvPB, ubil, ppk, ubiH, artQ, pepP, ssrA, queA, yajC, sqr, ttgC, gmk, ftsB, argG, glnQ, katG1, tam_4, spoT, murU, lptD, fni, zur, bprV, pksJ, rhaS_1, amgK, yedK_3, ftsN, mmsB, rlmB, ldh_3, gcvPA, hchA, rph, secD, tktA, recG, secF, queE, ibpB, bauC, artP, folB, uvrB, prlC, tyrS, bepA_4, macA, lexA_4, rpiA, pykA, rhaS_2, eno, arcB_8, btuD_5, mhpC, udp, uup, tal, rutD_3, clcA_3, add_3, arcB_7, fieF, rbsC, kdpB_3, polA, gcvH, mvaA, lpxD_4, argR, ynjF, pgk, pcpC, flgE_3, aaeB, surA, ycaC_3, epmB, adeP, mdIC, aatA, efp, fliY, metK, engB, argH, grpE, rpoZ, tqsA, aroH, spmB, rutD_2, pal_3, spmA, rutD_1, yhhQ, tam_5, mdly, yceF, alsD, pdxA, dfrA, argS, yabJ, cc4, gcvT, esiB_4, yveA_3, gap, proC, dltA_4, rnr, vldW, mepM, frcA, fabF_4, bepF_3, queF, mdtN, expZ, smc_6, hoxN, rayT_5, iolG_4, dnaK, zapA, ycdF, hin_3, dapB, yffB, msrB, alsS, oprB, per1, acyP</i>
<i>L. quateirensis</i> vs 27fs60	1665 vs 1661	<i>iolS, ltxB_1, ykfB, menH, APOA1BP, erg, ahID, rcsC_4, proP_7, decR, patA_4, higA-2, cmlA, xerC_6, gdx, xerC_7, acpP_3, dmpP, cusA_3, cysM, ltxD, rhmT, srfAA, lasB, sasa_7, asnB, besA, maolI, yeiG, fmt_3, fixB, argA, dhfrIII, cadI_2, nhaP2, mcbD, hpr, pinR, srkA_6, pasI, rdoA, sbmC, yiaD, neuA_1, dgcT, tral_3, bacC, ybcJ, pyrR, yccU, copA_3, yciV, pheA, cadI_1, greB, mdtA_4, lysN, mhpA, pksN, kynu, rdgB, tagO, tesA, mdtC, argT, luxR, glbN_3, ohrB, dauA_3, gdhB, hcaB_3, patA_5, argP, rcp1, merT, dap_4, tehB, nudC, emrD, hmo, proP_8, ltxB_2, ilvE, merR1, ant1, bmrU, recD2, nifH, maa, apbE, ndh, patA_3, resA_3, proB, dapL, yhbO, rapA_3, yfcG, srkA_7, yveA_4, neuA_3, rayT_6, birA_3, bioD1, aaeA, mecl, czcD_3, nmoA, fabG_5, tetA, yfiC, dltA_5, srpA, ipuF, proA, yhdJ, neuA_2, oqxB26, rmlC, xerC_5, ndhJ, mmpA, yofA</i>

<i>L. quateirensis</i> vs 30fs61	1665 vs 1501	<i>iolS, ydfG, ltxB_1, ykfB, menH, dsbA, APOA1BP, dnaJ_3, hemE, ahID, gcvPB, ubil, ppk, rcsC_4, ubiH, artQ, decR, patA_4, lptD_2, higA-2, pepP, cmlA, ssrA, queA, yajC, xerC_6, ttgC, xerC_7, dmpP, gmk, ftsB, cysM, argG, glnQ, ltxD, katG1, tam_4, rhmT, spoT, srfAA, sasA_7, murU, asnB, besA, maoll, yeiG, fni, zur, fmt_3, bprV, fixB, amgK, ftsN, argA, mmsB, rlmB, gcvPA, dhfrIII, lptD_1, rph, secD, cadl_2, nhaP2, tktA, hpr, pinR, recG, srkA_6, pasl, secF, queE, rdoA, sbmC, yiaD, neuA_1, bauC, artP, folB, uvrB, prlC, tyrS, dgcT, sqr_1, tral_3, bacC, rpiA, ybcJ, pykA, yccU, eno, copA_3, yciV, mhpC, pheA, cadl_1, udp, uup, tal, rutD_3, greB, mdtA_4, lysN, mhpA, fieF, pksN, kynu, rdgB, rbsC, tagO, mdtC, argT, luxR, polA, gcvH, mvaA, rhaS, argR, ynjF, pgk, ohrB, flgE_3, aaeB, dauA_3, gdhB, hcaB_3, patA_5, argP, surA, rcp1, merT, epmB, adeP, mdIC, dap_4, nudC, aatA, efp, metK, engB, argH, grpE, expZ_1, rpoZ, hmo, proP_8, ltxB_2, tqsa, aroH, spmB, rutD_2, ilvE, spmA, rutD_1, merR1, ant1, bmrU, yhhQ, recD2, mdly, yceF, alsD, nifH, pdxA, maa, apbE, argS, sqr_2, yabJ, expZ_2, cc4, patA_3, dapL, gcvT, yhbO, rapA_3, yveA_3, gap, proC, dltA_4, yfcG, rnr, srkA_7, mepM, yveA_4, neuA_3, frcA, rayT_6, fabF_4, birA_3, bioD1, aaeA, mecl, czcD_3, nmoA, queF, mdtN, fabG_5, hoxN, tetA, yfiC, dltA_5, rayT_5, srpA, iolG_4, dnaK, zapA, ipuF, yhdJ, neuA_2, oqx26, dapB, yffB, msrB, alsS, rmlC, per1, xerC_5, acyP, ndhJ, mmpA, yofA</i>
<i>L. quateirensis</i> vs 30cs62 ^T	1665 vs 1610	<i>iolS, ltxB_1, ykfB, prpB, menH, APOA1BP, ahID, prpC, rcsC_4, fadD, decR, patA_4, higA-2, cmlA, xerC_6, sad, xerC_7, ybgC, dmpP, rlmN, hisS, cusA_3, cysM, ltxD, katG1, rhmT, srfAA, sasA_7, asnB, besA, maoll, yeiG, fmt_3, fixB, argA, rplI, dhfrIII, cadl_2, nudL, nhaP2, der, hpr, tolQ, pinR, dtpA_6, algU, srkA_6, pasl, rdoA, sbmC, yiaD, neuA_1, lpxH, pdhA, tyrS, dgcT, sqr_1, tral_3, bacC, ybcJ, pdhB, yccU, copA_3, yciV, mhpC, pheA, cadl_1, greB, ynhG, lysN, mhpA, pksN, kynu, rdgB, tagO, argT, luxR, pcsA, rpsF, aroK_3, folK, ohrB, flgE_3, dauA_3, gdhB, hcaB_3, COQ3_2, patA_5, argP, rcp1, merT, dap_4, nudC, ygfZ, sdhE, ndk, hmo, proP_8, ltxB_2, rpsR, spmB, exbD, ilvE, spmA, merR1, ant1, bmrU, ppsR, recD2, nifH, maa, COQ3_3, apbE, hemL, fadJ, sqr_2, rubA1, patA_3, pdhC, COQ3_1, dapL, rodZ, yhbO, rapA_3, tolB, yveA_3, yfcG, srkA_7, mepM, yveA_4, neuA_3, rayT_6, birA_3, bioD1, aaeA, mecl, czcD_3, nmoA, queF, fabG_5, tetA, yfiC, dltA_5, srpA, pabB, prpD, bamB, ipuF, yhdJ, neuA_2, oqx26, mnmC, minC, rmlC, xerC_5, ndhJ, mmpA, yofA</i>

27fs60 vs <i>L. pneumophila</i>	1661 vs 1509	<i>ndoR_1, gnu, tdcB_1, nmtR_2, relG, kmtR, pasT, hcaE, umuC_3, traM, hlyU, wecA, hdl IVa, epsF_3, dhaA, rayT_1, intA_2, pstB, fabG_3, dinG_2, nhaP, pao, iolX, arcB_6, yhhW, pstS, madD, ohrA, rppH_3, proP_6, hdfR_2, mexB, hybG, noc_1, sasA_4, csrA_4, cbh_2, lrp, kdpD, csrA_7, fabG_4, acr3, kdpC, cpo_1, cpo_2, tdcD, rlpA_3, sqr, emrK, cmdD, qmcA, car, yfeW, ampC, parB_5, gshAB, thrB, cmpR, cbpD, pksC_1, mmcO, norM, csrA_8, inhA, thiD, lpxP_1, yhdE, cnrA, atpD_3, bprV, srkA_3, rayT_4, artJ_7, merA, ybiT, xerD, psuK, riza, pgl_2, topB, ykfA, hchA, traK, iolG_3, yxel_1, xerC_4, proP_5, gno, cmoM_1, rocG, ndoR_2, pnuC, dtpA_6, csrA_5, bepG, pgl_1, sutR, fliN_2, traL, yxel_2, artP, kdpA, yfcJ_2, cysZ, ycil, ycgM, COQ5_3, pdxH_3, plsB1, ntpA, arfB, hcl, yfcJ_1, tenA, arcB_8, btuD_5, sasA_5, fadI, mdtE, mhpC, dinG_1, relE, tal, ybeZ, fliD, oqxB23, arcB_3, fdx4, dehH1, tcmN, ynhG, czcC_4, arcB_7, des, kdga, kat, sasA_6, kdpB_3, pduX, srkA_4, caiA, anr, farB, pksC_2, hdfR_1, ppsC, thrA, parA, lpxD_4, srkA_5, lpxP_2, ntpB, nqo2, pcpC, czcB, tdcB_2, csd, aaeB, moaA, atpE_1, arsC, dgt, fliN_1, wecD, intA_3, sfmM2_2, smc_5, lgrD, fadA_3, ble, yisK, sufA, mscS, traJ, sdhE, noc_2, parB_3, yafN, pnoA, doeB, clsA, mdoB, cadA_3, tqsa, gbuA, nmtR_1, fpr, oadA, apbC, pal_3, dgkA, arsA, hybD, rayT_2, desA, yafP, yscN, tagG, bar, intA_1, yacG, thiM, mdly, alsD, ydgA, yafQ, czcA_4, alkI, arcB_2, cvfB, mtfA, cbh_1, sfmM2_1, ylpA_2, gpuA, dfrA, artI, arcB_5, traL_2, rayT_3, atoC_3, xanQ, trpS2, nfdA, pilE, yveA_3, gap, srpC, sodC1, alr, cspA, prs_3, traL_1, vldW, arcB_1, btuD_4, sctC_3, exuT, bepF_3, hlyB, cmoM_2, yafV, traG, smc_6, csrA_6, hoxN, rayT_5, iolG_4, pstA, ompP1, rstA, pepN_3, dns, parB_4, cpnA, pstC, arcB_4, bioC, atpE_2, yjjG, yffB, amiC, dcyD, cysC, nphT7, alsS, oxyL, hpxO, oprB, cheB, rraA, per1, astA, doxA, fbiB, ylpA_1</i>
30cs62^T vs <i>L. pneumophila</i>	1610 vs 1509	<i>gnu, nmtR_2, relG, pasT, kmtR, hcaE, traM, hlyU, wecA, hdl IVa, epsF_3, dhaA, rayT_1, erg, intA_2, pstB, fabG_3, nhaP, dinG_2, pao, iolX, arcB_6, yhhW, pstS, madD, ohrA, rppH_3, proP_6, hdfR_2, czcA_5, mexB, hybG, noc_1, proP_7, sasA_4, csrA_4, cbh_2, lrp, kdpD, csrA_7, fabG_4, acr3, kdpC, cpo, tdcD, rlpA_3, emrK, cmdD, qmcA, car, yfeW, ampC, parB_5, gshAB, thrB, ndoR, lasB, cmpR, cbpD, pgl, pksC_1, mmcO, norM, inhA, thiD, lpxP_1, yhdE, cnrA, atpD_3, bprV, srkA_3, rayT_4, merA, ybiT, xerD, psuK, riza, topB, ykfA, hchA, traK, iolG_3, xerC_4, proP_5, gno, mcbD, cmoM_1, rocG, pnuC, csrA_5, bepG, sutR, fliN_2, traL, artP, kdpA, yfcJ_2, cysZ, tdcB, ycil, ycgM, COQ5_3, pdxH_3, plsB1, ntpA, arfB, pyrR, hcl, yfcJ_1, tenA, arcB_8, btuD_5, sasA_5, fadI, mdtE, dinG_1, relE, tal, ybeZ, fliD, oqxB23, arcB_3, fdx4, dehH1, tcmN, czcC_4, arcB_7, des, kdga, kat, sasA_6, tesA, kdpB_3, mdtC, pduX, srkA_4, caiA, anr, farB, pksC_2, hdfR_1, ppsC, parA, lpxD_4, srkA_5, lpxP_2, ntpB, nqo2, pcpC, czcB, tuf, csd, aaeB, moaA, atpE_1, arsC, dgt, fliN_1, wecD, intA_3, sfmM2_2, emrD, smc_5, lgrD, fadA_3, thrA_1, ble, yisK, sufA, mscS, traJ, noc_2, parB_3, yafN, pnoA, doeB, clsA, mdoB, cadA_3, gbuA, tqsa, nmtR_1, fpr, oadA, apbC, pal_3, dgkA, arsA, hybD, rayT_2, desA, yafP, yscN, tagG, bar, intA_1, yacG, thiM, mdly, alsD, ydgA, yafQ, yxel, czcA_4, alkI, arcB_2, cvfB, mtfA, cbh_1, sfmM2_1, ylpA_2, gpuA, dfrA, artI, arcB_5, ndh, traL_2, rayT_3, atoC_3, xanQ, trpS2, nfdA, pilE, gap, srpC, sodC1, alr, cspA, prs_3, traL_1, vldW, arcB_1, btuD_4, sctC_3, exuT, bepF_3, hlyB, cmoM_2, traG, yafV, smc_6, csrA_6, hoxN, rayT_5, iolG_4, czcC_5, pstA, ompP1, rstA, pepN_3, dns, parB_4, cpnA, pstC, arcB_4, thrA_2, bioC, atpE_2, yjjG, yffB, amiC, dcyD, nphT7, cysC, alsS, oxyL, hpxO, oprB, cheB, rraA, per1, czcA_6, astA, doxA, fbiB, ylpA_1</i>

<i>L. pneumophila</i> vs 30fs61	1509 vs 1501	<i>bphD, setA, bpoC, traV, ydfG, ydjZ, glgE, ung, dsbA, APOA1BP, cfiA, sidJ_2, aadK, oqx13, pncB2, rbsB, yegS, mdtN_1, hemE, ah1D, gcvPB, cynT, rsbV, ybgl, ubil, ppk, ubiH, artQ, alr-1, ytbE, aspC, hcaC, cynR_1, pepP, ssrA, queA, pdaC, clcB, yajC, hypC, norW, lolC, mprA, dmpP, yqaB, zntB, gmk, bcr_1, ftsB, cusS, argG, glnQ, elaA, katG1, rutD, tam_4, korB, spoT, sidD, murU, aacA4, yxaF, lptD, bamC, fni, zur, pheA_2, ycdF_2, pheA_1, thiK, pksJ, rhaS_1, amgK, ftsN, mmsB, QTRT1, rlmB, gcvPA, dhfrIII, ycgJ, rph, cypB, secD, asnO, proX, tktA, potA, merR, csrA2_2, leuO, recG, ousA, nuoE, ppc, secF, queE, ibpB, yiaD, yajO, sdeA_1, sodC, decR_2, alx, ycfS, csrA2_1, bauC, folB, lem3, ligJ, uvrB, prlC, tyrS, macA, aroH_2, lexA_4, rclA, kpsM, rpiA, pykA, rhaS_2, thiG, eno, dkgA, yciV, yajG, lubX, decR_1, empA, dcsG, ibpA_2, ttgC_1, rimP, udp, uup, yfiC_2, sdeA_2, mdtA_4, cynR_2, fliO, fadM, mhpA, fieF, rdgB, rbsC, tagO, isca, polA, sthA_2, gcvH, hipB, mvaa, udk_3, drrA, bcr_2, argR, hepA, pgk, emrA, adhA, virB9, ohrB, ravA, kynB, vsr, gdhB, neuA, argP, surA, mdtN_2, rcp1, epmB, adeP, mdlC, ynjF_1, Int-Tn, ttgC_2, crt, aatA, efp, fliY, sthA_1, metK, engB, argH, nylA, cpxA, fabH_3, grpE, acdS, menH_2, rpoZ, spmB, ycdF_1, spmA, msrB_1, sixA, aroF, galP, argI, yhhQ, potB, secE, acrC, tam_5, cfa, recD2, yceF, htrA, ddaF, yfbR, fosB, pdxA, yfiC_1, argS, pbuE, nixA, fosA, kup_3, lexA_5, pla, yabJ, nlhH, msrAB, pdc, cc4, dprA, menH_1, uhpC, gcoB, pcaF, gcvT, sidJ_1, yhbO, esiB_4, yedI, proC, yfcG, rnr, aroH_1, gapA, mepM, frca, hsrA, fabF_4, ybdK, aruG, queF, blal, pncA, expZ, yhaJ, dnaK, lrgA, aseR, ibpA_1, rfbB, potD, zapA, phaC_4, yohK, udg, dapB, ynjF_2, ackA, mmgB, spnN, intS, sidE, roxA, acyP, dhaT, pabC, msrB_2, ratA, tam_6, ydeN</i>
--	-----------------------------	---

Table S4. MICs of strain 30cs62^T and *L. pneumophila* to antimicrobial drugs

Antimicrobial	Concentration range tested (mg/L)	MIC for 30cs62 ^T	Interpretation	MIC for <i>L. pneumophila</i>	Interpretation	EUCAST Cut-off for <i>L. pneumophila</i>
Azithromycin	0.016-256	0.19	S	0.125	S	0.25
Erythromycin	0.016-256	0.38	S	0.125	S	0.5
Ciprofloxacin	0.002-32	0.5	S	0.75	S	1
Rifampicin	0.002-32	0.016	S	0.023	S	0.032
Tigecycline	0.016-256	2	S	12	S	16
Imipenem	0.002-32	0.12	Not interpretable	0.03	Not interpretable	/

R: resistant; S: sensible according to EUCAST guidance document (27).

Acknowledgements

First of all, I would like to thank my tutor and mentor, Prof. Sandra Cristino for giving me the opportunity to work with her and grow both professionally and culturally. Thank you for always having a thought for me and for including me in all the thousands of super innovative projects. Thank you so much for everything.

A special thanks go to my research group and friends “*Legionella* Team”: Maria Rosaria, Federica, Marta, Simona and Carlo, because without you it wouldn't have been the same, and because I love you like a sibling, dear “stellini”. Additionally, also to Jessica, because even though she is no longer physically part of the group, she will forever remain so.

I want to also thank the trainees, because in their own small way, they too have contributed to my personal and professional growth.

In addition, I'd like to thank you to my friend, my pole, my goldfish and my favorite bioinformatic, Silvano, thanks for all the knowledge, support and patience.

Merci aussi à ma super-famille française qui a su égayer mon séjour à l'étranger en me faisant sentir que je faisais partie d'eux. Merci Silvia S., Silvia M., Elena, Pavo, Andrea, Marco, Francesca, Filippo, Giulio.

Moreover, I also wish to express my sincere appreciation to the ESGI group, for all the love that they have and show for the working with *Legionella*.

Of course, a special mentioning goes to my family, because if I am what I am, it is also thanks to you.

In particular, a special and huge thank you goes to Michele for all the support, patience and because he is a saint with me.

Last but not least, super special thanks to all my friends for the lightheartedness they give me.