INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY

RESEARCH ARTICLE

Crespi *et al.*, *Int. J. Syst. Evol. Microbiol.* 2023;73:005686

DOI 10.1099/ijsem.0.005686

MICROBIOLOGY SOCIETY

Legionella maioricensis sp. nov., a new species isolated from the hot water distribution systems of a hospital and a shopping center during routine sampling

Sebastian Crespi[1], Vladimír Drašar[2], Francisco Salvà-Serra[3,4,5,6], Daniel Jaén-Luchoro[3,4], Beatriz Piñeiro-Iglesias[3,4], Paul Christoffer Lindemann[7], Francisco Aliaga-Lozano[8], Victor Fernández-Juárez[9], Guillem Coll-García[6], Edward R. B. Moore[3,4,5] and Antoni Bennasar-Figueras[6,*]

Abstract

Two *Legionella*-like strains isolated from hot water distribution systems in 2012 have been characterized phenotypically, biochemically and genomically in terms of DNA relatedness. Both strains, HCPI-6^T and EUR-108, exhibited biochemical phenotypic profiles typical of *Legionella* species. Cells were Gram-negative motile rods which grew on BCYEα agar but not on blood agar and displayed phenotypic characteristics typical of the family *Legionellaceae*, including a requirement for L-cysteine and testing catalase positive. Both strains were negative for oxidase, urease, nitrate reduction and hippurate negative, and non-fermentative. The major ubiquinone was Q12 (59.4% HCPI-6^T) and the dominant fatty acids were $C_{16:1}\,\omega 7c$ (28.4% HCPI-6^T, ≈16% EUR-108), $C_{16:0}$ iso (≈22.5% and ≈13%) and $C_{15:0}$ anteiso (19.5% and ≈23.5%, respectively). The percent G+C content of genomic DNA was determined to be 39.3 mol%. The 16S rRNA gene, *mip* sequence and comparative genome sequence-based analyses (average nucleotide identity, ANI; digital DNA–DNA hybridization, dDDH; and phylogenomic treeing) demonstrated that the strains represent a new species of the genus *Legionella*. The analysis based on the 16S rRNA gene sequences showed that the sequence similarities for both strains ranged from 98.8–90.1% to other members of the genus. The core genome-based phylogenomic tree (protein-concatemer tree based on concatenation of 418 proteins present in single copy) revealed that these two strains clearly form a separate cluster within the genus *Legionella*. ANI and dDDH values confirmed the distinctiveness of the strains. Based on the genomic, genotypic and phenotypic findings from a polyphasic study, the isolates are considered to represent a single novel species, for which the name *Legionella maioricensis* sp. nov. is proposed. The type strain is HCPI-6^T (=CCUG 75071^T=CECT 30569^T).

INTRODUCTION

The first identification of the genus *Legionella* was in 1977, following an epidemic of acute pneumonia in Philadelphia [1, 2]. *Legionella* is the single genus of the family *Legionellaceae*, with 63 species with validly published names (https://lpsn.dsmz.de/genus/legionella) [3]. The genus was originally characterized as Gram-negative rods bacterium, aerobic, amoebae-associated,

Author affiliations: [1]Biolinea Int.; Calle Sócrates, 4; ES-07007 Palma de Mallorca, Spain; [2]Public Health Institute Ostrava - National Legionella Reference Laboratory; Masarykovo nam., 16; 682 01 Vyškov, Czech Republic; [3]Culture Collection University of Gothenburg (CCUG), Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden; [4]Department of Infectious Diseases, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden; [5]Clinical Microbiology, Sahlgrenska University Hospital; Region Västra Götaland, Gothenburg, Sweden; [6]Microbiologia – Departament de Biologia, Universitat de les Illes Balears; Campus UIB - Ctra. Valldemossa, Km 7.5; ES-07122 Palma de Mallorca, Spain; [7]Department of Microbiology, Haukeland University Hospital; Bergen, Norway; [8]Laboratorio Microbiología. Clínica Rotger, Grupo Quirónsalud – Vía Roma, 3, Carrer de Santiago Rusiñol, 9: ES-07012 Palma de Mallorca, Spain; [9]Marine Biology Section, Department of Biology, University of Copenhagen, 3000 Helsingør, Denmark.

***Correspondence:** Antoni Bennasar-Figueras, toni.bennasar@uib.es

Keywords: 16S rRNA gene; average nucleotide identity; digital DNA–DNA hybridization; genome comparison; hot water; *Legionella maioricensis*; *mip*; novel species; phylogenomics; polyphasic taxonomy.

Abbreviations: ANI, average nucleotide identity; ANIb, average nucleotide identity based on BLAST; BCYEα, buffered charcoal yeast extract medium supplemented with α-ketoglutaric acid; CFA, cellular fatty acid; dDDH, digital DNA–DNA hybridization; FAME, fatty acid methyl ester; GGDC, Genome-to-Genome Distance Calculator; HSPs, high-scoring pairs; MIC, minimal inhibitory concentration.

GenBank/EMBL/DDBJ/PIR accession numbers. Genomes: HCPI-6^T (=CCUG 75071^T=CECT 30569^T) JAJKBJ000000000, EUR-108 (= CCUG 75021=CECT 30570) JAKSZS000000000, 16S rRNA genes: HCPI-6^T OL472331, EUR-108 OL472332.

One supplementary figure and four supplementary tables are available with the online version of this article.

nutritionally fastidious with a requirement for iron and L-cysteine [4]. *Legionella* is widely distributed in natural aqueous environments and can become a health concern when they colonize and grow in human-made water systems, such as hot and cold water distribution systems, cooling towers (for air conditioning and industrial cooling purposes), recreational waters and other complex water systems capable of generating aerosols [5–8]. *Legionella* can be transmitted to humans mainly through the inhalation of contaminated aerosols, causing an infection often termed legionellosis. The term legionellosis is associated with respiratory pathologies [9] and includes both Pontiac fever, which is a non-pneumonic, mild flu-like illness, and Legionnaires' disease (LD), a severe, potentially fatal form of pneumonia [6, 10, 11]. Although *L. pneumophila* causes more than 90% of legionellosis with an average fatality rate of 12% [12], at least 25 other *Legionella* species have been isolated from clinical specimens and are considered to be pathogenic or opportunistic emerging pathogens (EN ISO 11731:2017).

In this report we describe the genotypic, genomic, phylogenetic and phenotypic characteristics of two *Legionella*-like strains isolated from the hot water distribution systems of a hospital (strain HCPI-6^T) and a shopping centre (strain EUR-108) during routine sampling. Given the specific characteristics, distinct from those observed in other species of the genus *Legionella*, we propose them to be classified as a novel species with the name *Legionella maioricensis* sp. nov., for which type strain HCPI-6^T is proposed (=CCUG 75071^T=CECT 30569^T).

ISOLATION AND ECOLOGY

The samples were taken and processed according with EN ISO 11731:2007. The methods used were accredited. The isolates were cultured on Glycine Vancomycin Polymyxin Cycloheximide (GVPC) media at 36.0±2 °C for 7–10 days, and further identified by serology and sequencing of the macrophage infectivity potentiator (*mip*) gene [13]. During a routine sampling in a hospital (Mallorca, Spain) in 2012, 10 samples were collected for *Legionella* testing. Five samples from different points of the hospital hot water distribution network were positive for *Legionella*; one for *Legionella pneumophila* serogroup 3 and the other four were identified as *Legionella* species. The counts for *Legionella* species were in the range of 300–400 c.f.u. l^{-1}. The *Legionella*-like isolate HCPI-6^T was obtained from the shower of the hot water distribution system (57.3 °C). Strains of morphologically similar *Legionella* species were isolated repeatedly at later samplings in the same hospital. Strain EUR-108 was also isolated during a routine sampling from the hot water distribution system (62 °C) of a supermarket (Příbor, Czech Republic) in 2012.

Different strategies to prevent colonization and subsequent growth of *Legionella* in hot water systems have been proposed, and the main one is the maintenance of high temperatures (50–60 °C) [14]. In addition, different European guidelines and regulations for the prevention of legionellosis in public facilities include the recommendation of a periodical thermal shock in the hot water system [14]. Nevertheless, isolation of *Legionella* at elevated temperatures (i.e., >55 °C) occurs with some frequency, along with the appearance of positive samples when using a real-time PCR (qPCR) detection method; although it is necessary to keep in mind that qPCR detects cells either in viable but non-culturable (VBNC) state or free DNA and dead cells [15–18]. Also, seems that some strains of *Legionella* subjected to high temperatures usually lose culturability; possibly developing resistance to high temperatures as proved by the still high percentages of culturable cells and VBNC present even after a 30 min treatment at 70 °C [16]. This phenomenon is probably related to changes in biotic or abiotic factors that affect the ecosystem in which *Legionella* proliferates [16, 19–21]. Consequently, when considering the ecology of *Legionella* it is necessary to recognize the importance of the factors that fosters its survival within stressful environments [22–25], such as those in which *L. maioricensis* strains were isolated. The enhanced survival in hot water system facilities is probably linked to the biofilm formation capabilities of *Legionella* species, affecting negatively to the efficiency of disinfection treatments, or the fluctuation of temperatures [22–25]. Although, if in such hot water system facilities, the found temperature in the sampling point is 60 °C, it does not mean that the microorganism is exposed to such temperatures permanently. For example, the micro-organism can survive in biofilms in various parts of the system where temperatures can fluctuate for different reasons i.e., flow differences. At the time of sampling, sudden changes in pressure can facilitate the detachment of pieces of biofilm with *Legionella* that can be sampled.

16S rRNA AND *MIP* GENE PHYLOGENIES

The detection of *Legionella* species in environmental and clinical samples is frequently performed by PCR amplification of the *mip* and/or 16S rRNA genes [13]. The *mip* gene codes for the macrophage infectivity potentiator (Mip), a 24 kDa surface protein that exhibits peptidyl-prolyl-cis/trans isomerase (PPIase) activity and is required for optimal infection of macrophages, and is also required for growth in amoebae [26, 27]. The Mip protein has been identified in both, virulent and avirulent *Legionella* species [28]. Combined with DNA sequencing, these two genetic loci (*mip* and 16S rRNA genes) can be used to distinguish different species of *Legionella* and identify *L. pneumophila* [13]. Therefore, a preliminary analysis based on 16S rRNA gene and *mip* sequences was conducted to identify the strains HCPI-6^T and EUR-108.

Thus, we applied a 16S rRNA gene sequencing comparative analysis to identify both isolates, against the currently 70 species available at the List of Prokaryotic names with Standing in Nomenclature for the genus *Legionella* [3]. The pairwise sequence similarities were calculated, using the method recommended by Meier-Kolthoff *et al.* [29] for the 16S rRNA gene available via

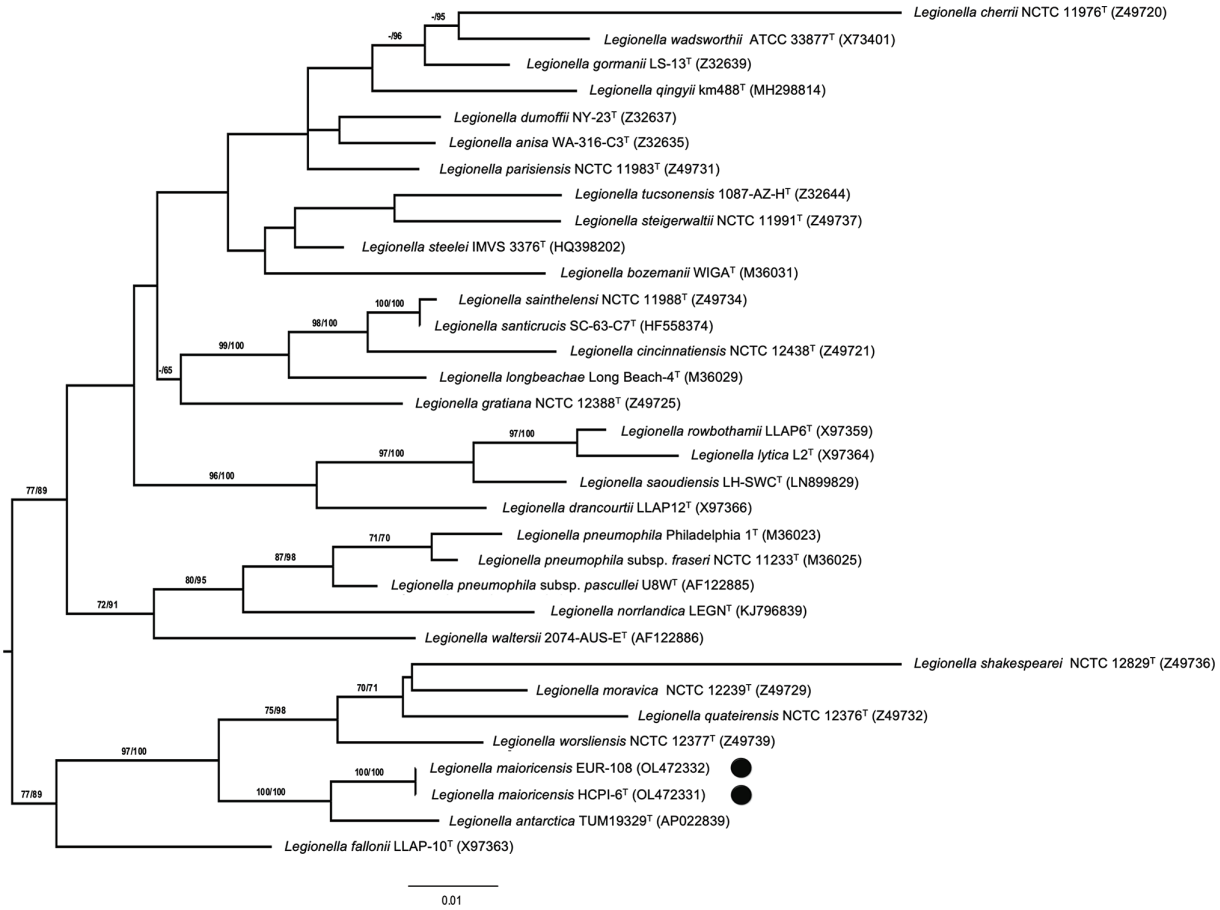

Fig. 1. 16S rRNA gene sequence-based ML phylogenetic tree inferred under the GTR+GAMMA model and rooted by midpoint-rooting [83]. The branches are scaled according to the expected number of nucleotide substitutions per site. The numbers above the branches are support values when larger than 60% from maximum-likelihood estimation (left) and maximum-parsimony (right) bootstrapping. The novel isolates are highlighted (•).

the Genome-to-Genome Distance Calculator (GGDC) web server available at http://ggdc.dsmz.de/ [30]. Phylogenies were inferred by the GGDC web server, using the DSMZ phylogenomics pipeline adapted to single genes [31]. A multiple sequence alignment was created with MUSCLE [32]. Maximum-likelihood (ML) and maximum-parsimony (MP) trees were inferred from the alignment with RAxML [33] and TNT [34], respectively. For the ML tree, rapid bootstrapping in conjunction with the autoMRE bootstopping criterion [35] and subsequent searches for the best tree were used; for the MP tree, 1000 bootstrapping replicates were used in conjunction with tree-bisection-and-reconnection branch swapping and ten random sequence addition replicates. The sequences were checked for a compositional bias, using the X^2 test as implemented in PAUP* [36]. As a result, the input nucleotide matrix comprised 33 operational taxonomic units and 1615 characters, 228 of which were variable and 122 of which were parsimony informative. The base-frequency check indicated no compositional bias (P=1.00, α=0.05). ML analysis under the GTR+GAMMA model yielded the highest log likelihood of −5710.07, whereas the estimated alpha parameter was 0.11. The ML bootstrapping did not converge; hence 1000 replicates were conducted; the average support was 63.3%. MP analysis yielded a best score of 615 (consistency index 0.48, retention index 0.61) and 11 best trees. The MP bootstrapping average support was 70.8%.

The 16S rRNA gene sequence similarities with species of *Legionella*, using 1539 nt, ranged between 98.8–90.1% (Table S1, available in the online version of this article). The results obtained from both strains (HCPI-6^T and EUR-108) are depicted in Fig. 1, and clearly indicated that the two strains formed a separate phylogenetic cluster inside the '*L. pneumophila*' clade, together with *L. fallonii* LLAP-10^T (X97363), *L. shakespearei* DSM 23087^T (Z49736), *L. moravica* NCTC 12239^T (Z49729), *L. quateirensis* NCTC 12376^T (Z49732) and *L. worsleiensis* NCTC 12377^T (Z49739); and having the closest phylogenetic affinity to a species with a recently validly published name, *Legionella antarctica* TUM19329^T (AP022839) [37, 38].

A classification scheme targeting the *mip* sequence to obtain a higher resolution identification of new *Legionella* environmental isolates was applied [39]. Accordingly, the sequence similarity values for a *mip* PCR amplicon (574 nt) among the strains HCPI-6^T and EUR-108 and other related type strains within the genus *Legionella* was determined through pairwise comparisons (Table

Table 1. Main features of *Legionella maioricensis* sp. nov. genomes

Features	HCPI-6^T	EUR-108
Contigs	68	79
Genome size (Mbp)	3.96	3.97
Plasmid*	1	1
Completeness (%)	100.0	100.0
Contamination (%)	2.19	2.19
G+C content (mol%)	39.3	39.3
Total genes	3502	3506
Genes coding for proteins	3423	3422
RNA genes	48	51
Ribosomal operons	1	1
tRNA	41	42
ncRNA	4	4

*Similarity evidences for genetic mobility elements, including a RepB family plasmid replication initiator, were detected in the genomes of HCPI-6^T (contig 8, 160.4 Kbp) and EUR-108 (contig 7, 157.7 Kbp) for plasmids pLPP of *Legionella pneumophila* str. Paris (131.9 Kbp) and pLLO of *Legionella longbeachae* NSW150 (71.8 Kbp) [84].

S1). The analysis based on *mip* [28, 39, 40] sequences indicated that the similarities among strains HCPI-6^T and EUR-108 (100%) and all 63 type strains of *Legionella* ranged between 91.5% (*L. antarctica* TUM19329^T) and 68.9% (*L. geestiana* ATCC 49504^T); i.e., the interspecies variation of *mip* sequences in *Legionella* ranged between 8.5–31.1%. According to Ratcliff *et al.* [39], the least *mip* sequence interspecies (between species) variation for a new taxon of the genus *Legionella* is >3.6%.

GENOME FEATURES

Genomic DNA of strain HCPI-6^T was obtained with the EZ1 Virus Mini Kit v2.0 (Qiagen) from pure cultures grown in buffered charcoal yeast extract medium supplemented with α-ketoglutaric acid (BCYEα) and following the manufacturer's instructions. Genomic DNA of strain EUR-108 was obtained from fresh biomass, using a modified version [41] of a previously described protocol [42]. The genome sequence of HCPI-6^T was obtained by using the Nextera XT DNA Library Preparation Kit and running on Illumina MiSeq sequencing technology with 300 bp paired-end reads. The confirmed high quality-filtered read sets were then assembled with Unicycler version 0.4.7 [43] at a 65.0× genome coverage. Similarly, for strain EUR-108 a standard genomic library preparation was performed, and sequencing was done on an Illumina NovaSeq system (read mode NovaSeq 6000 S4 PE150), obtaining 151 bp paired-end reads. High-quality Illumina paired-end reads were used for *de novo* assembly, using SPAdes v3.13.1 [44], with a total genome coverage of 256×.

The Table 1 summarizes the main features for the genomes of the *L. maioricensis* sequenced strains. The genome completeness and contamination degree were assessed by CheckM [45]. The CheckM completeness and contamination obtained for both genomes (Table 1) fall within the range of ≥95% complete with ≤% contamination that makes excellent reference genome for analysis [45, 46]. The annotation was done by the NCBI Prokaryotic Genome Annotation Pipeline [47], software revision 5.3.

A comparative genomics analysis of strains HCPI-6^T and EUR-108 with 63 species within the genus *Legionella* was carried out through identification of homologous genes to reconstruct a single-copy core genome tree to infer *phylogenomic* relations. Briefly, a pangenome based analysis was performed, applying an in-house pipeline (UHGene) designed for annotation [48], orthologous identification, gene clustering, determination of the genomic subsets of the of the core- and pangenome [49], and estimation of ML phylogenies. The core-genome based tree was calculated from a protein-concatemer, based on 418 single-copy orthologous protein sequences determined through different clustering algorithms (Fig. 2), i.e., Bi-Directional Best-Hits (BDBH) [49], COGtriangle (COG) [50], and OrthoMCL (OMCL) [51], alignment of individual protein clusters [52], concatenation of resulting alignments, and elimination of poorly aligned positions and divergent regions of protein sequences [53]. The resulting amino acids alignment contained 127 008 positions (80% of the original) and was used as input data for PhyML version 3.0 [54, 55]. An ML phylogenetic model was first estimated, using simultaneous nearest-neighbour interchange moves to improve a BioNJ [56] initial tree under the LG amino acids substitution model [57]. The shape parameter (α) of a gamma distribution with four categories was estimated. The support of the data for each internal branch of the phylogeny was estimated using non-parametric bootstrap analysis (100

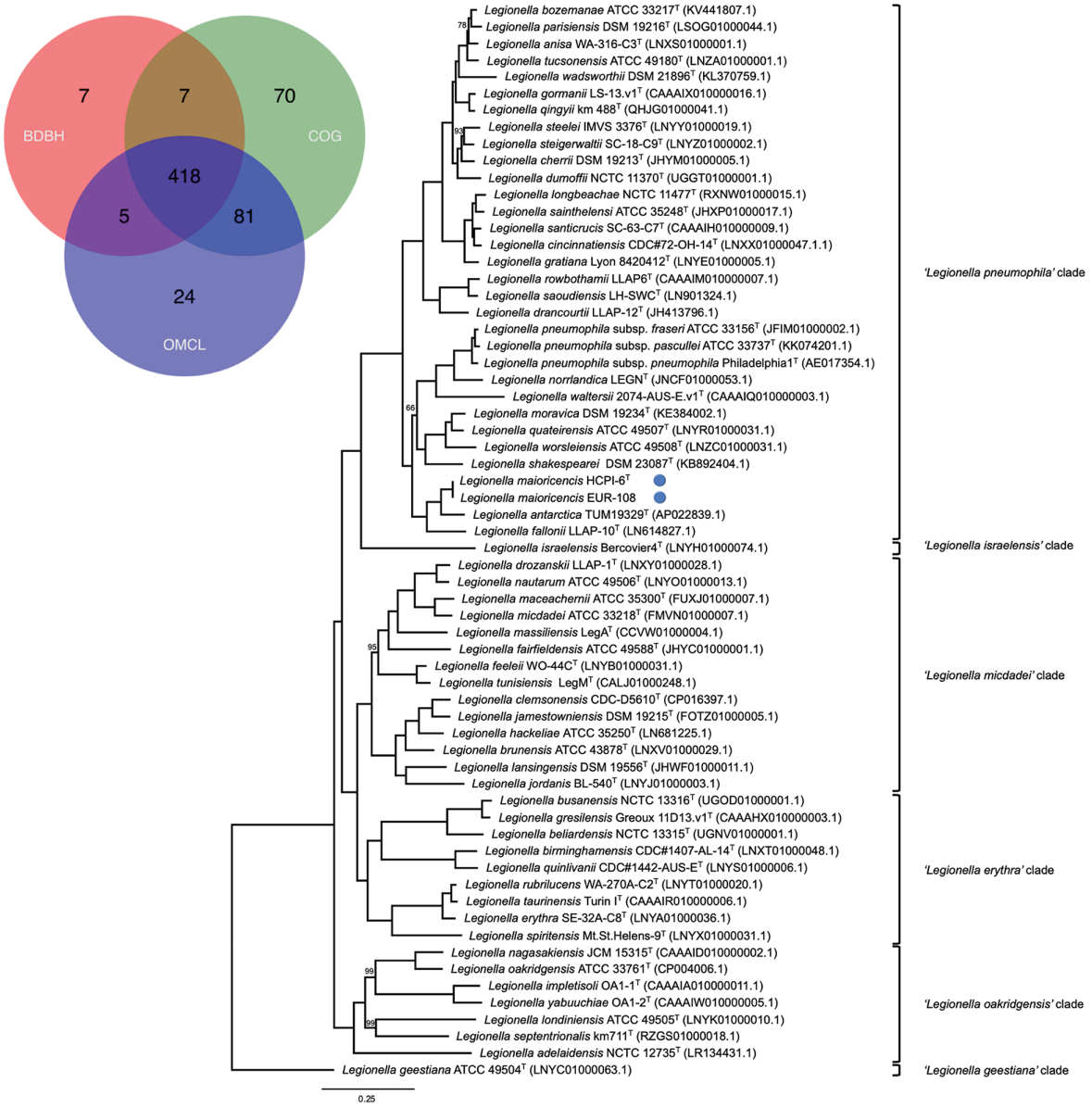

Fig. 2. Comparative genomic relationships of strains HCPI-6^T and EUR-108 with the type strains of 63 species of the genus *Legionella* [30], based on the resulting protein-concatemer of 418 single-copy orthologous sequences; i.e., those forming in the Venn diagram the consensus cluster intersection calculated by the three clustering strategies applied: BDBH, COG and OMCL. The bootstrap values are 100, except where indicated at the nodes. The genome of *L. geestiana* ATCC 49504^T (LNYC01000063.1) was used as an outgroup. The clade distributions are as described previously [72]. Bar, 0.25 substitutions per nucleotide position. The novel isolates are highlighted (•).

replicates). The estimated phylogeny is presented in Fig. 2. Similarly, to the ML tree inferred from 16S rRNA gene sequences, both strains formed a separate phylogenetic branch cluster inside the '*L. pneumophila*' clade. Again, the genomes of the type strains of *L. antarctica*, *L. fallonii*, *L. shakespearei*, *L. moravica*, *L. quateirensis* and *L. worsleiensis*, together with *L. norrlandica*, *L. waltersii* and *L. pneumophila*, shaped the most immediate proximity within the phylogenomic tree (Fig. 2).

The phylogenomic inferences based on pairwise comparisons among the genome sequences of HCPI-6^T, EUR-108 and the genome sequences of type strains available in databases for *Legionella* were computed, using overall genome relatedness indices [58] useful for species delineation as pairwise average nucleotide identity (ANI) based on BLAST (ANIb) [59] and digital DNA–DNA hybridization (dDDH). The ANIb percentages were calculated using JSpeciesWS [60, 61]. The dDDH values and confidence intervals and G+C content differences were determined at the GGDC server (see above), using the recommended settings of the GGDC 3.0 [30, 62].

The resulting values for ANIb and dDDH were below the accepted thresholds for new species (i.e., <95 and <70% for ANI and dDDH, respectively). The ANIb values between HCPI-6^T and EUR-108 and related type strains of the genus *Legionella* indicate that *L. antarctica* is the most closely related species (82.8%), followed by *L. fallonii* (77.1%) (Table S2) [63, 64].

Attending to the dDDH results obtained by applying the formula *d4* (also known as GGDC formula 2), which is independent of genome length and is, thus, robust when using incomplete draft genomes (Table S2); besides the value between strains HCPI-6^T and EUR-108 (99.9%), the highest values obtained for the species more closely related, from the trees inferred from 16S rRNA gene sequences and phylogenomics, are for *L. antarctica* (28.0%) and *L. fallonii* (24.0%). However, some genomes of other species showed values even higher than the one obtained for *L. antarctica,* like *L. steigerwaltii* ATCC 35302^T (34.3%), *L. maceachernii* ATCC 35300^T (32.4%), *L. saoudiensis* LH-SWC^T (31.0%) or *L. hackeliae* ATCC 35250^T (30.0%), which is a species belonging to the *'Legionella micdadei'* clade. Therefore, considering >70% dDDH similarity value (formula *d4*) as the 'gold standard' threshold for species boundaries, and using a 79% dDDH threshold for subspecies clustering, as previously introduced [31, 65], the dDDH results obtained also confirmed that both strains, i.e., HCPI-6^T and EUR-108, are a clearly a distinct species (Table S2). Similarly, a >1 mol% G+C content difference indicates a potentially unreliable identification result for the provided query strains; considering that, if computed from genome sequences, within-species G+C content varies no more than 1 mol% [66]. From the set of species more closely related, *L. antarctica* (0.15 mol%) and *L. fallonii* (0.96 mol%) are below 1 mol% G+C content. Furthermore, other species included in the *'L. pneumophila*' clade, with the exception of *L. hackeliae* (0.32 mol%), have values ranging from 0.04 mol% (*L. drancourtii*) and 0.72 mol% (*L. moravica*); including *L. saoudiensis* (0.14 mol%), *L. waltersii* (0.22 mol%), *L. quateirensis* (0.28 mol%), *L. dumoffii* (0.35 mol%) and *L. steelei* (0.52 mol%).

The genome mining of HCPI-6^T and EUR-108 to gain insights into the genome features related to pathogenicity raised up to 101 sequences at the VFDB [67] and 176 at TrueBac ID [68]. The putative virulence factors detected in *L. maioricensis* comprised genes for adherence, intracellular survival (*mip*), iron uptake and regulation, chemotaxis and motility, nutrient acquisition, genes coding for stress proteins, invasion, regulation virulence-related genes (e.g., LetA/S two-component system and RelA) and a type II toxin-antitoxin system PrlF family detected in a putative plasmid [69]. Moreover, the secretion systems identified included a type II Lsp, an Lvh system homologous to type IVA (T4ASS) and type IVB Icm/Dot together with effector proteins associated to this secretion complex; such as LepA and LepB which are related to nonlytic release of *L. pneumophila* from protozoa for the pathogen's dissemination [70, 71]. Finally, a class A β-lactamase gene was detected in both strains, HCPI-6^T and EUR-108. Consequently, the pathogenetic analysis of HCPI-6^T and EURO-108 for genes and combinations related to virulence factors predicted this micro-organism as a potential human pathogen.

PHYSIOLOGY AND CHEMOTAXONOMY

The growth characteristics of strains HCPI-6^T and EUR-108 were tested, using BCYEα agar, BCYEα agar without L-cysteine and Columbia blood agar. The optimal growth for both strains was at 30 °C on BCYEα for 5 days. These cultivation conditions were used to perform the phenotypic tests. Neither of the two strains grew on blood agar. Furthermore, no growth was observed at 4 °C after 10 days. Although at slower rates and with lower biomass production, both strains grew at 37 °C. Growth under microaerophilic (from 5.5–12% O_2 after 24 h) and anaerobic conditions (< 0.1% O_2 after 2.5 h and >15% CO_2 after 24 h) was tested on BCYEα agar at 30 °C for 7 days with the corresponding GENbag systems (bioMérieux). The Gram and acid-fast staining's were carried out by conventional methods. Cell morphology was investigated by transmission electron microscopy (TEM) (Fig. 3). Colonies were examined for autofluorescence under long-wavelength ultraviolet (Wood's Lamp). Phase-contrast microscopy was used to assess spore formation, after incubating the strains on BCYEα $MnSO_4{\cdot}H_2O$ agar (BCYEα supplemented with 5 mg $MnSO4{\cdot}H_2O$ l^{-1}) for 3 days [72]. Cell motility was assessed under light microscopy, using the hanging-drop method. Colony morphology was observed macroscopically, using cells grown on BCYEα agar. Bromocresol purple assay was performed on BCYEα-BCP agar (BCYEα supplemented with 10 µM bromocresol purple) [73]. The biochemical tests for catalase, urease, nitrate reduction, indole tests, glucose fermentation, starch hydrolysis and hippurate hydrolysis were determined by using cultures grown at 30 °C on BCYEα agar, as described previously [72, 74]. Oxidase was tested by using a test paper containing tetramethyl-phenylenediamine dihydrochloride (bioMérieux). A modified Gots test [75] was applied to determine the β-lactamase production of *L. maioricensis* strains. Briefly, single colonies of *L. maioricensis* strains under study were inoculated on BCYEα by a single streak from the plate centre outward, and incubated for 3 days at 30 °C. To 4 ml of melted Mueller–Hinton agar 0.7% (w/v) at 45 °C, 0.1 ml of a 24 h broth culture of a penicillin sensitive *Micrococcus luteus* strain were added, mixed and immediately poured on the surface of the BCYEα agar with *Legionella* streaks. Plates were allowed to harden at room temperature and a penicillin susceptibility disc (6 µg/10 IU) was added to the centre of the plate and incubated for 24 h more at 30 °C.

The two strains were Gram-negative (HCPI-6^T cells were Gram-negative, rod-shaped with an average size of 2.4±0.5 µm cell length and 0.6 ±0.1 µm width), acid-fast-staining negative, microaerobic, presented cell mobility, non-spore forming with growth dependent on L-cysteine. Both strains grew well on BCYEα agar after incubation for 3–5 days, but not on BCYEα agar without L-cysteine or Columbia blood agar. Colonies on BCYEα agar were grey, glistening, convex and circular with an entire edge (Fig. S1). The two strains grew at temperature ranging from 25 to 37 °C, with optimum growth at 30 °C. Both strains were positive

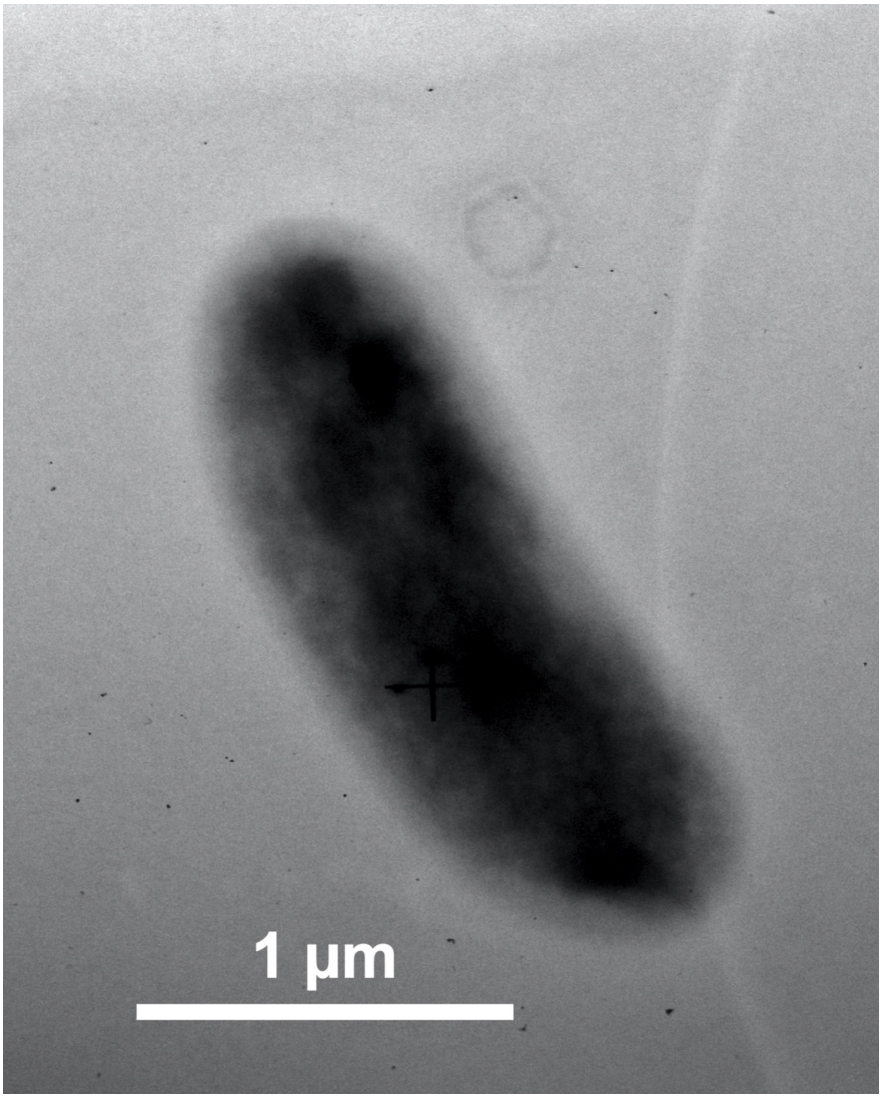

Fig. 3. TEM image of *Legionella maioricensis* HCPI-6^T.

for catalase; negative for oxidase, urease, gelatinase, glucose fermentation, hippurate hydrolysis, nitrate reduction, indole and starch tests. Neither of the two strains exhibited blue-white or red autofluorescence under long-wavelength ultraviolet. Browning of tyrosine-containing media tests were initially performed on yeast extract agar with tyrosin, following a reference from 1978 [76], but none of the *L. maioricensis* strains grew, even after 10 days of incubation at 30 °C. An alternative method using BCYEα supplemented with tyrosine was applied [77]. The bromocresol purple assay was negative since the colonies of *L. maioricensis* strains growing on the BCYE*α*-BCP agar were grey rather than purple. Differential characteristics of HCPI-6^T and EUR-108 with closely related type strains are listed in Table 2, while the detailed characteristics of strain HCPI-6^T are given in the species description. Attending to the phenotypic characteristics of the most closely related species, the psychrotolerant Antarctic strain TUM19329^T, which is able to grow at 4 °C and up to 25 °C but not above 30 °C [37], while the opposite occurs for the two novel strains.

The analysis of isoprenoid quinone composition was carried out by the Identification Service, Leibniz-Institut DSMZ - Deusche Sammlung von Mikroorganismen und Zellkulturen GmbH (Braunschweig, Germany). The analysis of respiratory quinones results (Table S3) concluded that the major ubiquinones in both strains were Q12 and Q13; these features were shared with the type strains of the closest related species, *Legionella antarctica* TUM19329^T and *L. fallonii* LLAP-10^T.

The predominant cellular fatty acids (CFAs) of strain HCPI-6^T and EUR-108 were determined after growth on BCYEα agar, at 48 h and 30 °C. The bacterial biomass was harvested (100 mg for HCPI-6^T and 60 mg for EUR-108) and the fatty acid methyl ester (FAME) extraction was performed as described by Sasser [78]. The CFA-FAME profiles were determined with a gas chromatograph

Table 2. Selected characteristics of *Legionella mairocensis* sp. nov. and selected closely related species of the genus *Legionella*

Strains*: 1, HCPI-6^T; 2, EUR-108; 3, *L. antarctica* [37]; 4, *L. fallonii* [79]; 5, *L. moravica*; 6, *L. norrlandica* LEGN^T [85]; 7, *L. quateirensis*; 8, *L. shakespearei*; 9, *L. waltersii* ATCC 51914^T; 10, *L. worsleiensis* ATCC 49508^T; 11, *L. pneumophila* subsp. *pneumophila* Philadelphia 1^T. Characteristics based on varying numbers of strains. +, Positive; –, negative; ±, strain dependent; w, weakly positive; ND, which means the information for the related species was not determined.

Characteristics	1	2	3	4	5	6	7	8	9	10	11
Gram-staining reaction	–	–	–	–	–	–	–	–	–	–	–
Acid-fastness	–	–		ND	–	ND	–	–	–	–	–
Spore-forming	–	–	–	ND	ND	ND	ND	ND	ND	ND	–
Growth requirement for L-cysteine	+	+	+	+	+	+	+	+	+	+	+
Browning of tyrosine-containing media	–	–	ND	ND	w	ND	+	–	–	+	+
Acid from carbohydrates	–	–	–	–	–	ND	–	–	–	–	–
Nitrate to nitrite	–	–		ND	–	ND	–	–	–	–	–
Hippurate hydrolysis	–	–	–	ND	–	ND	–	–	+	–	+
Gelatin liquefaction (gelatinase)	–	–	–	ND	+	ND	+	+	+	+	+
β-Lactamase production	+	+	ND	ND	+	ND	+	+	+	+	+
Catalase	+	+	+	+	+	ND	+	+	+	–	+
Oxidase	–	–	–	w	– or w	–	–	w	+	–	+ or ±
Motility	+	+	ND	ND	+	ND	+	+	+	+	+
Autofluorescence	–	–	ND	–	–	–	–	–	–	–	–

* Except where indicated, data were obtained from [4].

(HP 5890, Hewlett-Packard) and a variation of the MIDI Sherlock Microbial Identification System (MIS) (http://www.ccug.se/identification/chemotyping). The CFA-FAME profiles obtained for *L. maioricensis* strains are shown in Table 3; which also includes the profiles of the type strains of the most closely related species, *L. antarctica* TUM19329^T [37] and *L. fallonii* LLAP-10^T [79]. With the exception of *L. antarctica* TUM19329^T, which profile was obtained at 25 °C, all other strains were grown at 30 °C and makes profiles comparison feasible with those obtained for *L. maioricensis*. Since each bacterial strain possesses a unique fatty acid profile when grown on a given culture conditions, including at a certain temperature, each bacterial strain possesses a unique fatty acid profile and comparing CFA-FAME profiles of *L. antarctica* with the rest of strains is not realistic.

Strain HCPI-6^T presented the straight-chain monounsaturated $C_{16:1}$ ω7*c* as the predominant CFA (28.4%), followed by methyl branched fatty acids $C_{16:0}$ iso (22.4%) and $C_{15:0}$ anteiso (19.5%) in similar proportions; lower amounts of $C_{16:0}$ (6.4%), $C_{17:0}$ cyclo (5.1%) and $C_{14:0}$ iso (5.1%) were observed (Table 3). Strain EUR-108 showed the methyl branched fatty acid $C_{15:0}$ anteiso (23.3%) as the predominant CFA; followed by an unsaturated CFA $C_{16:1}$ ω7*c* (15.9%); and the methyl-branched FA $C_{16:0}$ iso (12.9%), $C_{17:0}$ anteiso (7.3%), $C_{14:0}$ iso (6.5%) and $C_{15:0}$ (5.2%) (Table 3).

Chemotyping has been shown to be useful for differentiating *Legionella* species from other Gram-negative bacteria [80, 81]. The species of *Legionella* are usually characterized relative amounts of methyl-branched fatty acids (iso and anteiso) and inconsequential amounts of hydroxy fatty acids [80, 82]. Both novel isolates showed a high relative proportion (%) of methyl-branched CFA, 47.0% (HCPI-6^T) to 50.0% (EUR-108), and lower amounts of hydroxy CFA, two saturated branched-chain hydroxy CFA ($C_{14:0}$ iso 3OH and $C_{15:0}$ iso 3OH) ranging from 2.8% (HCPI-6^T) to 6.6% (EUR-108), approximately.

Although *L. pneumophila* subsp. *pneumophila* CCUG 9568^T, *L. nautarum* CCUG 44900^T and *L. worsleisensis* CCUG 44923^T were not analysed in the present study and are not present in Table 3, the CFA profiles for these species were previously analysed at CCUG (http://www.ccug.se) and were observed to be similar to those reported previously [81]. Strain HCPI-6^T has a similar CFA profile pattern to the type strain of the type species of the genus, *L. pneumophila* subsp. *pneumophila* CCUG 9568^T, and *L. worsleiensis*, with $C_{16:1}$ ω7*c* and $C_{16:0}$ iso as the predominant CFAs. However, EUR-108 has a similar profile pattern to that of *L. nautarum* [80, 81], with $C_{15:0}$ anteiso as the predominant CFA, followed by high amounts of $C_{16:1}$ ω7*c* and $C_{16:0}$ iso, wherein the $C_{15:0}$ anteiso relative amount is approximately double that of $C_{16:0}$ iso (23.3 and 12.9%, respectively) (Table 3). Both strains differed from the type species of the genus in that, $C_{15:0}$ anteiso is one of the predominant CFA (19.5 and 23.3% for HCPI-6^T and EUR-108 respectively), whereas $C_{15:0}$ anteiso in *L. pneumophila* subsp. *pneumophila* is 9.4%. $C_{16:0}$ was observed as relatively high amounts

Table 3. Cellular fatty acid (CFA) compositions of *Legionella maioricensis* sp. nov. strains and the most closely related species

Strains: 1, HCPI-6^T; 2, EUR-108; 3, *L. antarctica* TUM19329^T (at 25 °C)* [37]; 4, *L. fallonii* LLAP-10^T [79]; 5, *L. pneumophila* subsp. *pneumophila* Philadelphia 1^T (=CCUG 9568^T), type species of the genus. Values are percentages of total fatty acids. Relative amounts are shown for CFAs ≥1%; CFAs <1% are indicated as trace (TR); ND, not detected. Mean value±SD of duplicates of HCPI-6^T [1] and EUR-108 [2] are presented in the table when meaningful (i.e., >1.0%). The three main CFAs for each species are highlighted in bold.

Fatty acid (%)	***1***	***2***	***3***	***4***	***5***
Saturated:					
$C_{14:0}$	TR	2.5	TR	2–3	ND
$C_{15:0}$	2.5	5.2±1.6	ND	2–3	TR
$C_{16:0}$	6.4	3.9±2.5	1.7	**15–19**	**12.0**
$C_{17:0}$	4.4	1.5	TR	2	1.0
$C_{18:0}$	T	3.5	1.0	6–7	3.0
$C_{20:0}$	ND	1.2	TR	1–2	2
Unsaturated:					
$C_{14:1}$ ω5*c*	2.0	2.2	1.2	ND	TR
$C_{15:1}$ ω6*c*	ND	TR	1.6	3–4	1.5
$C_{16:1}$ ω7*c*	**28.4±1.2**	**15.9±5.7**	**27.2**†	**21–27**	**29.6**
Cyclopropane:					
$C_{17:0}$ cyclo	5.1	4.8	ND	2–6	ND
Methyl branched (iso, anteiso):					
$C_{15:0}$ anteiso	**19.5**	**23.3±1.3**	**36.2**	12–15	9.4
$C_{17:0}$ anteiso	TR	7.3	4.1	2–3	7.4
$C_{14:0}$ iso	5.1	6.5	3.6	4–5	2.4
$C_{16:0}$ iso	**22.4**	**12.9±2.4**	**15.2**	**11–13**	**23.4**
Branched-chain hydroxy:					
$C_{14:0}$ iso 3OH	2.8	4.3	TR	ND	TR
$C_{15:0}$ iso 3OH	ND	2.3	ND	ND	ND

*Values obtained from literature: TUM19329^T was grown at 25 °C, all other strains were grown at 30 °C.
†This is given as the $C_{16:1}$ *ω7c*/$C_{16:1}$ *ω6c* summed feature 3 value in the Shimada *et al.* [37] study., the summed feature could be $C_{16:1}$ *ω7c* and/or $C_{16:1}$ *ω6c*.

of the CFA profile of *L. pneumophila* subp. *pneumophila* profile; however, straight-chain CFAs were not present in significant amounts in the CFA profiles of the novel strains (≤6.5%) (Table 3).

Antimicrobial susceptibility testing for the two *Legionella* strains (HCPI-6^T and EUR-108) using the gradient test method was performed according to EUCAST guidance on AST of *L. pneumophila*, dated May 2021 (http://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Guidance_documents/Legionella_guidance_note_-_20210528.pdf). Briefly, 0.5 McFarland suspensions of the two strains were swabbed onto BCYEα agar plates, and a single minimal inhibitory concentration (MIC) test strip (Liofilchem) was applied to each plate. A total of nine antibiotics were tested. Plates were incubated for 72 h at 35 °C in a humidified atmosphere. An MIC value was read at the point of intersection of the growth ellipse and the gradient strip. The two test strains did grow on the BCYEα agar with sufficient growth to determine the MIC values after 72 h incubation. Both strains were tested in three biological replicates and the final MIC for each antibiotic is the median of the three individual values (Table S4). For each setup a QC-strain (*L. pneumophila* ATCC 33152^T) was included to assure the method. The MICs of the QC-strain were within the tentative wild-type for every setup.

We applied a polyphasic approach to characterize two novel strains isolated from the hot water distribution systems of a hospital and a shopping centre during routine samplings. Based upon the genotypic, genomic and phenotypic analyses, the two strains,

HCPI-6^T and EUR-108, were determined to represent a distinct and novel species within the genus *Legionella*. We propose the name, *L. maioricensis* sp. nov., for the species, with the strain HCPI-6^T designated to be the type strain of the species.

DESCRIPTION OF *LEGIONELLA MAIORICENSIS* SP. NOV.

Legionella maioricensis (mai.or.i.cen'cis. M.L. fem. Adj. maioricensis, pertaining to Mallorca, where the type strain of the species was isolated).

The bacteria grow optimally at 30 °C (temperature range for growth: 25–37 °C) on BCYE agar medium, which contains L-cysteine; they do not grow on blood agar. Cells are Gram-stain negative, are negative for acid-fast staining, do not form spores and are motile; they are rod-shaped, microaerobic, oxidase-negative and catalase-positive. Cells do not autofluoresce under long-wavelength ultraviolet light. The bacteria require L-cysteine in the media for growth. The bacteria do not produce acid from carbohydrates, they do not generate nitrite from nitrate and do not hydrolyse hippurate or liquefy gelatin. Methyl-branched CFAs ($C_{16:0}$ iso, $C_{15:0}$ anteiso and $C_{14:0}$ iso) comprise approximately 50% of the CFA profile with unsaturated CFA ($_{C16:1}$ ω7c) comprising approximately 20% of the CFA profile. Characteristic dominant fatty acids comprised $C_{16:1}$ ω7c, $C_{16:0}$ iso and $C_{15:0}$ anteiso. The major respiratory lipo-quinones were determined to be ubiquinones Q12, Q13 and Q11.

Strain HCPI-6^T was obtained from the shower of a hospital hot water distribution system (57.3 °C). The bacterium clusters phylogenetically within the spectrum of *Legionella* species, with a 16S rRNA gene sequence similarity of 98.8% to the most closely related species, *L. antarctica*, and with 92.8% sequence similarity to the type species of *Legionella*, *L. pneumophila*. Genomic DNA G+C content is 39.3mol%, determined from genome sequences. The genomes were estimated to be 3.9 Mbp long, with approximately 3500 genes (approximately 3420 protein-coding genes).

The Whole Genome Shotgun (WGS) project of strain HCPI-6^T (=CCUG 75071^T=CECT 30569^T) has been deposited at DDBJ/ENA/GenBank under the genome accession number JAJKBJ000000000. Similarly, the WGS project of strain EUR-108 (=CCUG 75021=CECT 30570) has been deposited at DDBJ/ENA/GenBank under the genome accession number JAKSZS000000000. The DDBJ/ENA/GenBank accession numbers for the 16S rRNA gene sequences of strains HCPI-6^T and EUR-108 are OL472331 and OL472332, respectively. All the versions described in this paper are the first versions.

Funding information
This research was supported by Biolinea Int.

Acknowledgements
The authors acknowledge the staff at the Culture Collection University of Gothenburg (CCUG, Gothenburg, Sweden) for technical assistance with cultivation and characterization of the strains. The Substrate Unit of the Department of Clinical Microbiology (Sahlgrenska University Hospital, Gothenburg, Sweden) is acknowledged for providing the culture media. The CCUG is supported by the Department of Clinical Microbiology, Sahlgrenska University Hospital.

Author contribution
S.C. and V.D.: performed initial isolation and characterization of strains. F.S.S.: genome comparisons, deposition and identification, DNA preparation. D.J.L.: genome comparisons, database analyses, strains preparation for ubiquinones analysis. B.P.I.: FAME analysis, phenotypic tests. P.C.L.: antibiotic susceptibility testing. F.A.: DNA preparation, molecular gene amplification, gene sequencing and analysis. V.F.J.: genome comparison and database analyses. G.C.: microscopy. E.R.B.M.: taxonomic and phenotypic characterization, performed deposition and identification of strains. A.B.F.: design of the research and project outline, genomic and pangenomic based characterization, strain and genome sequence deposit, drafted the manuscript. All authors read and approved the final manuscript.

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

References

1. Fraser DW, Tsai TR, Orenstein W, Parkin WE, Beecham HJ, *et al*. Legionnaires' disease: description of an epidemic of pneumonia. *N Engl J Med* 1977;297:1189–1197.
2. McDade JE, Shepard CC, Fraser DW, Tsai TR, Redus MA, *et al*. Legionnaires' disease: isolation of a bacterium and demonstration of its role in other respiratory disease. *N Engl J Med* 1977;297:1197–1203.
3. 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.
4. Winn WC*et al*. Jr. *Legionella*. In: William Barnaby W, DeVos P, Chun J, Dedysh S, Hedlund B, et al. (eds). *Bergey's Manual of Systematic Bacteriology*. Hoboken, New Jersey: John Wiley & Sons, Inc., in association with Bergey's Manual Trust; 2015.
5. Fields BS, Benson RF, Besser RE. *Legionella* and Legionnaires' disease: 25 years of investigation. *Clin Microbiol Rev* 2002;15:506–526.
6. Mercante JW, Winchell JM. Current and emerging *Legionella* diagnostics for laboratory and outbreak investigations. *Clin Microbiol Rev* 2015;28:95–133.
7. Mondino S, Schmidt S, Rolando M, Escoll P, Gomez-Valero L, *et al*. Legionnaires' disease: state of the art knowledge of pathogenesis mechanisms of *Legionella*. *Annu Rev Pathol* 2020;15:439–466.
8. Mazzotta M, Salaris S, Pascale MR, Girolamini L, Cristino S. Occurrence of *Legionella* spp. in man-made water sources: isolates distribution and phylogenetic characterization in the Emilia-Romagna region. *Pathogens* 2021;10:552.
9. Diederen BMW. *Legionella* spp. and Legionnaires' disease. *J Infect* 2008;56:1–12.

10. Cunha BA, Burillo A, Bouza E. Legionnaires' disease. *Lancet* 2016;387:376–385.
11. Oliva G, Sahr T, Buchrieser C. The life cycle of *L. pneumophila*: cellular differentiation is linked to virulence and metabolism. *Front Cell Infect Microbiol* 2018;8:3.
12. Edelstein PH, Lück C, Jorgensen JH, Carroll KC, Funke G, *et al*. *Legionella*. In: *Manual of Clinical Microbiology*, 11th edn. ASM Press, 2015.
13. Yong SFY, Tan SH, Wee J, Tee JJ, Sansom FM, *et al*. Molecular detection of *Legionella*: moving on from mip. *Front Microbiol* 2010;1:123.
14. Molina JJ, Bennassar M, Palacio E, Crespi S. Low efficacy of periodical thermal shock for long-term control of *Legionella* spp. in hot water system of hotels. *Pathogens* 2022;11:152.
15. 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.
16. Allegra S, Grattard F, Girardot F, Riffard S, Pozzetto B, *et al*. Longitudinal evaluation of the efficacy of heat treatment procedures against *Legionella* spp. in hospital water systems by using a flow cytometric assay. *Appl Environ Microbiol* 2011;77:1268–1275.
17. Kirschner AKT. Determination of viable legionellae in engineered water systems: do we find what we are looking for? *Water Res* 2016;93:276–288.
18. Rasheduzzaman M, Singh R, Haas CN, Gurian PL. Required water temperature in hotel plumbing to control *Legionella* growth. *Water Res* 2020;182:115943.
19. Steinert M, Ockert G, Lück C, Hacker J. Regrowth of *Legionella pneumophila* in a heat-disinfected plumbing system. *Zentralbl Bakteriol* 1998;288:331–342.
20. García MT, Jones S, Pelaz C, Millar RD, Abu Kwaik Y. Acanthamoeba polyphaga resuscitates viable non-culturable *Legionella pneumophila* after disinfection. *Environ Microbiol* 2007;9:1267–1277.
21. Allegra SV, Berger FO, Berthelot P, Grattard F, Pozzetto B, *et al*. Use of flow cytometry to monitor *Legionella* viability. *Appl Environ Microbiol* 2008;74:7813–7816.
22. Murga R, Forster TS, Brown E, Pruckler JM, Fields BS, *et al*. Role of biofilms in the survival of *Legionella pneumophila* in a model potable-water system. *Microbiology* 2001;147:3121–3126.
23. Saby S, Vidal A, Suty H. Resistance of *Legionella* to disinfection in hot water distribution systems. *Water Sci Technol* 2005;52:15–28.
24. Tachikawa M, Tezuka M, Morita M, Isogai K, Okada S. Evaluation of some halogen biocides using a microbial biofilm system. *Water Res* 2005;39:4126–4132.
25. Chang CW, Hwang YH, Cheng WY, Chang CP. Effects of chlorination and heat disinfection on long-term starved *Legionella pneumophila* in warm water. *J Appl Microbiol* 2007;102:1636–1644.
26. Cianciotto NP, Fields BS. *Legionella pneumophila* mip gene potentiates intracellular infection of protozoa and human macrophages. *Proc Natl Acad Sci* 1992;89:5188–5191.
27. Swanson MS, Hammer BK. *Legionella pneumophila* pathogesesis: a fateful journey from amoebae to macrophages. *Annu Rev Microbiol* 2000;54:567–613.
28. 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.
29. Meier-Kolthoff JP, Göker M, Spröer C, Klenk H-P. When should a DDH experiment be mandatory in microbial taxonomy? *Arch Microbiol* 2013;195:413–418.
30. Meier-Kolthoff JP, Carbasse JS, Peinado-Olarte RL, Göker M. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res* 2022;50:D801–D807.
31. Meier-Kolthoff JP, Hahnke RL, Petersen J, Scheuner C, Michael V, *et al*. Complete genome sequence of DSM 30083^T, the type strain (U5/41^T) of *Escherichia coli*, and a proposal for delineating subspecies in microbial taxonomy. *Stand Genomic Sci* 2014;9:2.
32. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;32:1792–1797.
33. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;30:1312–1313.
34. Goloboff PA, Farris JS, Nixon KC. TNT, a free program for phylogenetic analysis. *Cladistics* 2008;24:774–786.
35. Pattengale ND, Alipour M, Bininda-Emonds ORP, Moret BME, Stamatakis A. How many bootstrap replicates are necessary? *J Comput Biol* 2010;17:337–354.
36. Swofford DP. *Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0 b*. 10 edn. Sunderland: Sinauer Associates; 2002.
37. Shimada S, Nakai R, Aoki K, Kudoh S, Imura S, *et al*. Characterization of the first cultured psychrotolerant representative of *Legionella* from Antarctica reveals its unique genome structure. *Microbiol Spectr* 2021;9:e0042421.
38. Oren A, Garrity GM. Valid publication of new names and new combinations effectively published outside the IJSEM. *Int J Syst Evol Microbiol* 2022;72.
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. Pascale MR, Salaris S, Mazzotta M, Girolamini L, Fregni Serpini G, *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:e0116121.
41. Salvà-Serra F, Svensson-Stadler L, Busquets A, Jaén-Luchoro D, *et al*. A protocol for extraction and purification of high-quality and quantity bacterial DNA applicable for genome sequencing: A modified version of the Marmur procedure. *Protocol Exchange* 2018.
42. Marmur J. A procedure for the isolation of deoxyribonucleic acid from micro-organisms. *J Mol Biol* 1961;3:208–IN1.
43. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 2017;13:e1005595.
44. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, *et al*. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19:455–477.
45. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 2015;25:1043–1055.
46. Brady A, Salzberg SL. Phymm and PhymmBL: metagenomic phylogenetic classification with interpolated Markov models. *Nat Methods* 2009;6:673–676.
47. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, *et al*. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 2016;44:6614–6624.
48. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;30:2068–2069.
49. Contreras-Moreira B, Vinuesa P. GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Appl Environ Microbiol* 2013;79:7696–7701.
50. Kristensen DM, Kannan L, Coleman MK, Wolf YI, Sorokin A, *et al*. A low-polynomial algorithm for assembling clusters of orthologous groups from intergenomic symmetric best matches. *Bioinformatics* 2010;26:1481–1487.
51. Li L, Stoeckert CJ, Roos DS. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* 2003;13:2178–2189.
52. Löytynoja A, Goldman N. Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. *Science* 2008;320:1632–1635.

53. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 2000;17:540–552.
54. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, *et al*. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 2010;59:307–321.
55. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 2003;52:696–704.
56. Gascuel O. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Mol Biol Evol* 1997;14:685–695.
57. Le SQ, Gascuel O. An improved general amino acid replacement matrix. *Mol Biol Evol* 2008;25:1307–1320.
58. Chun J, Rainey FA. Integrating genomics into the taxonomy and systematics of the *Bacteria* and *Archaea*. *Int J Syst Evol Microbiol* 2014;64:316–324.
59. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, *et al*. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 2007;57:81–91.
60. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci* 2009;106:19126–19131.
61. 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.
62. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14:60.
63. Thompson CC, Chimetto L, Edwards RA, Swings J, Stackebrandt E, *et al*. Microbial genomic taxonomy. *BMC Genomics* 2013;14:913.
64. Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, *et al*. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 2018;68:461–466.
65. Meier-Kolthoff JP, Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* 2019;10:2182.
66. Meier-Kolthoff JP, Klenk H-P, Göker M. Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age. *Int J Syst Evol Microbiol* 2014;64:352–356.
67. Liu B, Zheng D, Jin Q, Chen L, Yang J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res* 2019;47:D687–D692.
68. Ha SM, Kim CK, Roh J, Byun JH, Yang SJ, *et al*. Application of the whole genome-based bacterial identification system, TrueBac ID, using clinical isolates that were not identified with three matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) systems. *Ann Lab Med* 2019;39:530–536.
69. Schmidt O, Schuenemann VJ, Hand NJ, Silhavy TJ, Martin J, *et al*. *prlF* and *yhaV* encode a new toxin-antitoxin system in *Escherichia coli*. *J Mol Biol* 2007;372:894–905.
70. Chen J, de Felipe KS, Clarke M, Lu H, Anderson OR, *et al*. *Legionella* effectors that promote nonlytic release from protozoa. *Science* 2004;303:1358–1361.
71. Voth DE, Broederdorf LJ, Graham JG. Bacterial type IV secretion systems: versatile virulence machines. *Future Microbiol* 2012;7:241–257.
72. Li L-H, Zhang L, Wu H-Y, Qu P-H, Chen J-C, *et al*. *Legionella septentrionalis* sp. nov., isolated from aquatic environments in the northern PR China. *Int J Syst Evol Microbiol* 2021;71.
73. Zheng BX, Ibrahim M, Zhang DP, Bi QF, Li HZ, *et al*. Identification and characterization of inorganic-phosphate-solubilizing bacteria from agricultural fields with a rapid isolation method. *AMB Express* 2018;8:47.
74. Brenner DJ, Steigerwalt AG, Gorman GW, Wilkinson HW, Bibb WF, *et al*. Ten new species of *Legionella*. *Int J Syst Bacteriol* 1985;35:50–59.
75. Gots JS. The detection of penicillinase production properties of microorganisms. *Science* 1945;102:309.
76. Baine WB, Rasheed JK, Feeley JC, Gorman GW, Casida LE. Effect of supplementall-tyrosine on pigment production in cultures of the Legionnaires' disease bacterium. *Current Microbiology* 1978;1:93–94.
77. Vickers RM, Yu VL. Clinical laboratory differentiation of *Legionellaceae* family members with pigment production and fluorescence on media supplemented with aromatic substrates. *J Clin Microbiol* 1984;19:583–587.
78. Sasser M. *Identification of bacteria by gas chromatography of cellular fatty acids. Technical Note #101*. Newark, DE: MIDI Inc; 2001.
79. Adeleke AA, Fields BS, Benson RF, Daneshvar MI, Pruckler JM, *et al*. *Legionella drozanskii* sp. nov., *Legionella rowbothamii* sp. nov. and *Legionella fallonii* sp. nov.: three unusual new *Legionella* species. *Int J Syst Evol Microbiol* 2001;51:1151–1160.
80. Lambert MA, Moss CW. Cellular fatty acid compositions and isoprenoid quinone contents of 23 *Legionella* species. *J Clin Microbiol* 1989;27:465–473.
81. Diogo A, Veríssimo A, Nobre MF, da Costa MS. Usefulness of fatty acid composition for differentiation of *Legionella* species. *J Clin Microbiol* 1999;37:2248–2254.
82. Kowalczyk B, Chmiel E, Palusinska-Szysz M. The role of lipids in *Legionella*-host interaction. *IJMS* 2021;22:1487.
83. Hess PN, DE Moraes Russo CA. An empirical test of the midpoint rooting method. *Biol J Linn Soc Lond* 2007;92:669–674.
84. Cazalet C, Rusniok C, Brüggemann H, Zidane N, Magnier A, *et al*. Evidence in the *Legionella pneumophila* genome for exploitation of host cell functions and high genome plasticity. *Nat Genet* 2004;36:1165–1173.
85. Rizzardi K, Winiecka-Krusnell J, Ramliden M, Alm E, Andersson S, *et al*. *Legionella norrlandica* sp. nov., isolated from the biopurification systems of wood processing plants. *Int J Syst Evol Microbiol* 2015;65:598–603.