



Genomic Investigation of a Legionellosis Outbreak in a Persistently Colonized Hotel

Leonor Sánchez-Busó^{1,2}, Silvia Guiral^{1,3}, Sebastián Crespi^{4,5}, Víctor Moya⁶, María L. Camaró⁷, María P. Olmos⁷, Francisco Adrián⁶, Vicente Morera⁸, Francisco González-Morán^{1,3}, Hermelinda Vanaclocha^{1,3} and Fernando González-Candelas^{1,2*}

¹ Unidad Mixta "Infección y Salud Pública" Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana-Salud Pública, Universitat de València, Valencia, Spain, ² Centros de Investigación Biomédica en Red Epidemiología y Salud Pública, Valencia, Spain, ³ Dirección General de Salud Pública, Subdirección de Epidemiología y Vigilancia de la Salud, Valencia, Spain, ⁴ Biolinea Int, Palma de Mallorca, Spain, ⁵ Policlínica Miramar, Palma de Mallorca, Spain, ⁶ Legioprev, Valencia, Spain, ⁷ Dirección General de Salud Pública, Laboratorio de Salud Pública, Valencia, Spain, ⁸ Centro de Salud Pública Denia, Denia, Spain

OPEN ACCESS

Edited by:

Rekha Seshadri,
Department of Energy Joint Genome
Institute, USA

Reviewed by:

Awdhesh Kalia,
University of Texas MD Anderson
Cancer Center, USA
Emmanuel F. Mongodin,
University of Maryland School of
Medicine, USA

*Correspondence:

Fernando González-Candelas
fernando.gonzalez@uv.es

Specialty section:

This article was submitted to
Evolutionary and Genomic
Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 05 August 2015

Accepted: 22 December 2015

Published: 21 January 2016

Citation:

Sánchez-Busó L, Guiral S, Crespi S,
Moya V, Camaró ML, Olmos MP,
Adrián F, Morera V, González-Morán F,
Vanaclocha H and
González-Candelas F (2016) Genomic
Investigation of a Legionellosis
Outbreak in a Persistently Colonized
Hotel. *Front. Microbiol.* 6:1556.
doi: 10.3389/fmicb.2015.01556

Objectives: A long-lasting legionellosis outbreak was reported between November 2011 and July 2012 in a hotel in Calpe (Spain) affecting 44 patients including six deaths. Intensive epidemiological and microbiological investigations were performed in order to detect the reservoirs.

Methods: Clinical and environmental samples were tested for the presence and genetic characterization of *Legionella pneumophila*. Six of the isolates were subjected to whole-genome sequencing.

Results: Sequencing of 14 clinical and 260 environmental samples revealed sequence type (ST) 23 as the main responsible strain for the infections. This ST was found in the spa pool, from where it spread to other hotel public spaces, explaining the ST23 clinical cases, including guests who had not visited the spa. Uncultured clinical specimens showed profiles compatible with ST23, ST578, and mixed patterns. Profiles compatible with ST578 were obtained by direct sequencing from biofilm samples collected from the domestic water system, which provided evidence for the source of infection for non ST23 patients. Whole genome data from five ST23 strains and the identification of different STs and *Legionella* species showed that different hotel premises were likely colonized since the hotel opening thus explaining how different patients had been infected by distinct STs.

Conclusions: Both epidemiological and molecular data are essential in the investigation of legionellosis outbreaks. Whole-genome sequencing data revealed significant intra-ST variability and allowed to make further inference on the short-term evolution of a local colonization of *L. pneumophila*.

Keywords: Legionnaire's disease (LD), outbreaks, typing, complete genome, mixed infections

INTRODUCTION

Legionella infections are opportunistic and the inhalation of aerosols with enough bacterial loads can cause a severe form of pneumonia, known as Legionnaires' disease (Fraser et al., 1977), or a milder flu-like condition, denoted as Pontiac fever (Glick et al., 1978). After its first identification in 1976 in a Legionnaires' convention in a hotel in Philadelphia, many legionellosis outbreaks have involved travel-associated clusters (Benin et al., 2002; Burnsed et al., 2007; Rota et al., 2011).

Legionellosis outbreaks are usually studied by typing pure *Legionella* cultures from infected patients with different molecular techniques, assuming that only one strain is causing the disease. However, co-infection with different *Legionella* species has also been reported (Meyer et al., 1980; Muder et al., 1983; Buchbinder et al., 2004; Matsui et al., 2010; McAdam et al., 2014; Wewalka et al., 2014), as well as with different serogroups of *L. pneumophila* (Meyer et al., 1980; Yiallourous et al., 2013; Wewalka et al., 2014). The introduction of new typing methods based on direct amplification and sequencing of *Legionella* from clinical samples (Coscollá and González-Candelas, 2009; Ginevra et al., 2009; Mentasti et al., 2012) has also revealed that outbreak patients can be infected simultaneously by more than one *Legionella* strain (Coscollá et al., 2014), even from the same serogroup. Dual or multiple infections pose an additional difficulty for the identification and subsequent control of outbreak sources.

In November 2011, the first cases of a legionellosis outbreak that lasted 33 weeks were reported in a hotel in the locality of Calpe (Comunidad Valenciana, Spain). The outbreak comprised four different clusters with two suspected and 42 confirmed patients that matched the European case definition including six fatalities. The hotel had a previous history of six legionellosis cases in 2006, just 3 months after it was opened to the public and one case in 2007. However, no additional cases had been reported until the outbreak analyzed in this work, in which both tourists and workers were affected.

Upon notification of the first cases, public health officials started an environmental investigation and the implementation of control measures (Vanaclocha et al., 2012). These included the disinfection of the water distribution system of the hotel and its closure. Here, we present the molecular analyses performed during the investigation of this outbreak using both cultured and uncultured clinical and environmental samples that helped unveil the source and route of infection for all the patients. Whole-genome sequencing (WGS) was performed for some of the isolates to get higher discrimination than the traditional DNA Sequence-Based Typing (SBT) and to infer when *Legionella pneumophila* might have colonized the facility.

MATERIALS AND METHODS

Case Definition and Epidemiological Description of the Outbreak

Confirmed cases were defined as patients who, having stayed or worked at the hotel between 2 and 10 days before the onset

of symptoms, showed a clinical diagnosis of pneumonia with laboratory findings which confirmed infection by *Legionella*, including a positive urine test for *L. pneumophila* antigen or a positive culture isolation of the bacteria from respiratory secretions.

The outbreak was epidemiologically divided into four temporal clusters that involved 44 cases, 38 tourists (average age 71.5) and six hotel workers (average age 49.5) (Figure 1). Visitors had stayed at 28 different rooms in the hotel, distributed in 11 floors. Six of the rooms accommodated two cases and other three patients had also used the same room, two of them simultaneously. The outbreak involved six deaths (average age 77.2), and deceased patients had stayed at five different rooms: two of those patients had stayed in the same room at different times. Only five tourists and two workers had ever been in the facility of the local spa, which is located below the main hall from where it is easily visible through a glass dome.

Clinical Samples

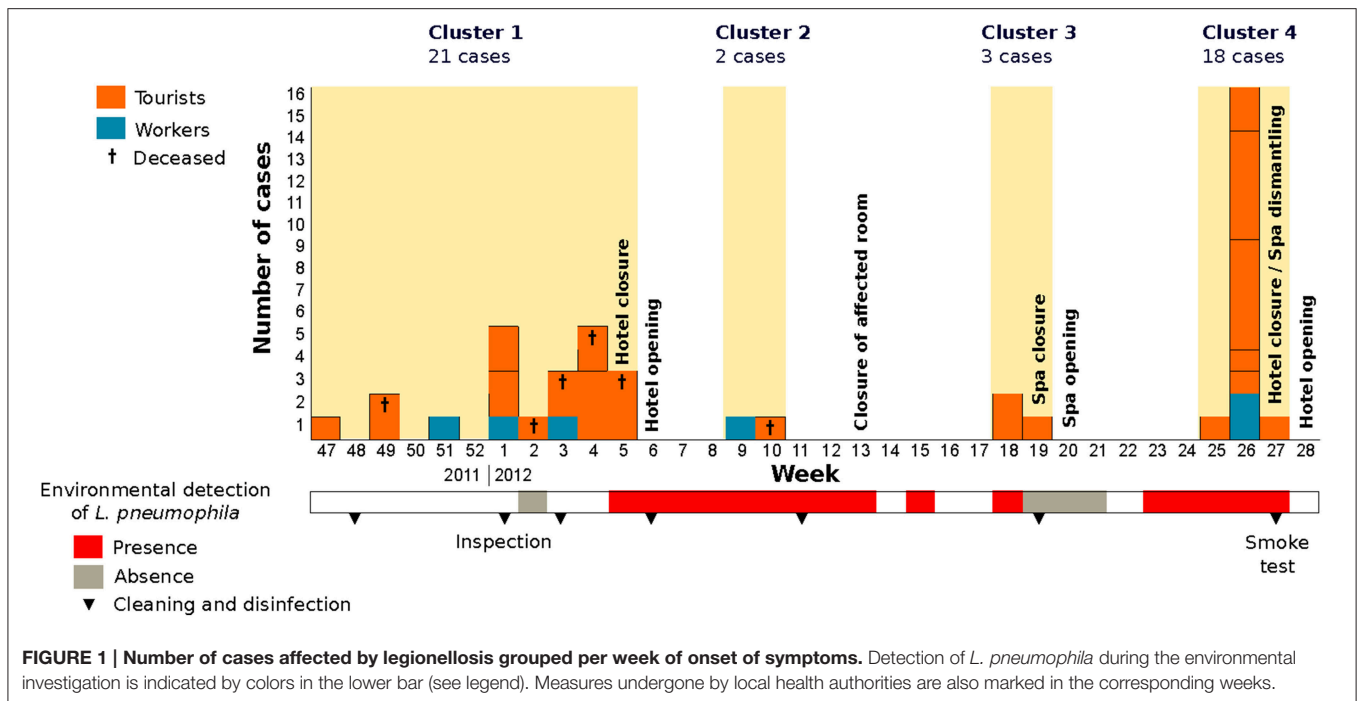
Clinical samples from 14 outbreak patients were received for study in our local reference laboratory for *L. pneumophila* infections. Specifically, eight sputum, two broncho-alveolar aspirates (BAS), one lung puncture aspirate and three cultures obtained in the Microbiology Service of the corresponding hospitals, two of them from autopsies, were collected and transported at 4°C for genetic testing. Total DNA was extracted from uncultured clinical samples using UltraClean® BloodSpin® DNA Isolation Kit (MoBio). DNA from the three isolates was extracted using a thermal shock as described below for environmental isolates.

Environmental Sampling

A total of 632 two-liter water samples and 164 biofilm swabs (per duplicate) were taken during the environmental investigation of the outbreak from the different water systems of the hotel including the spa pool. Water samples were tested for biocides, pH and temperature. Pieces of air filters of the air-conditioning system connecting the spa and the hall area were also collected for testing. All samples were transported immediately to the lab for *L. pneumophila* detection and genetic characterization.

Water filtering and culturing were performed following the AFNOR NF T 90-431 2 (2003) and UNE-EN ISO 11731-2:2008 regulations on water quality and standard culture-based detection of *Legionella* spp. Isolates were serogrouped using Oxoid *Legionella* Latex Test (Thermo Scientific), subtyped using monoclonal antibodies (Joly et al., 1983, 1986), and subsequently characterized genetically. Different colonies from the culture plates were analyzed, when available, to check for strains of different genetic profile.

Biofilm swabs and air filters were submerged into lysis buffer (100 mM Tris HCl, 100 mM NaCl, 1 mM EDTA, 10% SDS and ultrapure water, final pH 8.1) and incubated with 1 mg/mL proteinase K during 10 min at 65°C for biofilm swabs and 30 min at the same temperature for the other samples. Cotton swabs were introduced into 0.7 mL microtubes with an opening at the bottom placed inside a 2-mL microtube (Suenaga and Nakamura, 2005) and centrifuged at maximum speed for 10 min to recover the



maximum amount of lysate. Final retrieved volumes were used for direct DNA purification using the spin filter protocol of the UltraClean[®] BloodSpin[®] DNA Isolation Kit (MoBio).

DNA from *L. pneumophila* culture isolates was extracted by thermal shock. A small amount of bacteria was suspended in 200 μ L of autoclaved, ultrapure water and subjected to two cycles of a 5-min incubation step at 99°C followed by 5 min at 4°C. Cell debris was pelleted by centrifuging during 3 min at 12,000 rpm, and the DNA-containing supernatant was used for molecular amplification.

L. pneumophila Detection and Sequence-Based Typing (SBT)

Touchdown PCR (TD-PCR; Don et al., 1991, targeting two regions of the Sequence-Based Typing (SBT), Gaia et al., 2005; Farhat et al., 2011) scheme (*pilE* and *asd*) was performed on the DNA samples extracted from biofilms for *L. pneumophila* detection. Positive samples were subsequently used for amplifying the five remaining loci (*fliC*, *mip*, *mompS*, *proA*, and *neuA*) in the SBT scheme for this species. Amplification was performed using 2.5 U FastStart High Fidelity PCR System (Roche Applied Science), 1X Buffer with MgCl₂ 1.8 mM, 200 μ M of each dNTP, 4% DMSO, 0.2 μ M of the corresponding pair of primers, an amount of DNA suspension depending on its concentration (measured three times by NanoDrop 1000TM, Thermo Scientific) and autoclaved and filtered ultrapure water to a final volume of 25 μ L. The thermal profile of the TD-PCR included an initial denaturation step of 5 min at 95°C followed by 10 cycles of denaturation (30 s at 95°C), annealing (30 s at a primer-dependent temperature, Sánchez-Busó et al., 2015) and at a ramp rate of $-1^{\circ}\text{C}/\text{cycle}$ and extension (30 s at 72°C). Immediately after the TD step, 35 cycles of standard PCR were applied using the same conditions but fixing the

annealing temperature at the optimal T_m for each region; a final extension step of 8 min at 72°C was also used to ensure complete amplification.

DNA retrieved from *L. pneumophila* pure cultures and uncultured clinical samples from confirmed patients were used for amplification of the 7 loci in the SBT scheme by traditional (Gaia et al., 2003, 2005; Coscollá and González-Candelas, 2007) and seminested PCR amplification (Coscollá and González-Candelas, 2009). PCR products were purified and Sanger-sequenced as previously described in Coscollá and González-Candelas (2007, 2009).

Cloning Of PCR Products

PCR products from five positive biofilm samples obtained in hotel rooms associated to clinical cases and that showed evidence of multiple *L. pneumophila* variants in *fliC*, *pilE*, and *asd* (Table 2) were cloned into *Escherichia coli* JM109 Competent Cells (Promega) using pGEM[®]-T Easy vector System II (Promega). In order to check whether the minor variants detected could be considered actual polymorphisms and not artifacts resulting from the amplification and/or cloning process, TD-PCR products from pure cultures of these three regions were cloned and sequenced. PCR products from the same regions obtained from four sputum/BAS samples were also cloned (Table 2) in order to check for intra-patient variability. Ligation and transformation reactions were performed following the instructions of the manufacturer.

Phylogenetic Reconstruction of SBT Data

The sequences of the alleles involved in the STs found in the analyzed samples were downloaded from the EWGLI SBT database (http://www.hpa-bioinformatics.org.uk/legionella/legionella_sbt/php/sbt_homepage.php) and concatenated in

order to create a haplotype sequence for each strain. Complete and incomplete profiles (with at least two sequenced loci) were used for maximum likelihood phylogenetic reconstruction using RAxML v7.2.8 (Stamatakis, 2006) with the GTRGAMMA model of substitution and 1000 bootstrap replicates along with reference STs.

Whole-Genome Sequencing on Outbreak Isolates

Four clinical ST23 and 2 environmental isolates (ST23 and ST1236) were selected because of their DNA quality and quantity for whole-genome sequencing using the SOLiD 5500XL platform. Single-end reads of 75 bp with an average per nucleotide coverage of 90X were mapped against the assembly of one of the ST23 strains (VelvetOptimiser v2.2.5, Zerbino, 2010), ID_125 (resulting scaffold of 3307, 11 bp), following the pipeline described in Sánchez-Busó et al. (2014). Filtered SNPs from the core genome of the six strains and nine reference genomes from GenBank [Paris—NC_006368 (Cazalet et al., 2004), Lens—NC_006369 (Cazalet et al., 2004), Philadelphia1—NC_002942 (Chien et al., 2004), Alcoy—NC_014125 (D'Auria et al., 2010), Corby—NC_009494 (Steinert et al., 2007), Lorraine—NC_018139 (Gomez-Valero et al., 2011)], 130b—FR687201 (Schroeder et al., 2010), ATCC43290—NC_016811 and HL06041035—NC_018140) were used for phylogenetic reconstruction under the minimum evolution criterion (Rzhetsky and Nei, 1993). To do so, pairwise distances and 1000 bootstrap replicates were applied with MEGA5 (Tamura et al., 2011). SOLiD genomic data have been deposited in the European Nucleotide Archive (ENA) under project accession PRJEB5990.

Bayesian Estimation of Colonization Time

In order to study the time at which ST23 could have started proliferating within the hotel, we used the ST23 non-homoplastic

core alignment (3,267,949 bp) for Bayesian estimation with BEAST v1.7.5 (Drummond et al., 2012). The HKY model of nucleotide substitution and an uncorrelated lognormal molecular clock were selected using the substitution rate previously estimated for ST578 (Sánchez-Busó et al., 2014) and the isolation dates as priors. Four parallel Markov chains using the same parameters were run for 10,000,000 generations with samples taken every 10,000 steps. Convergence of the chains was tested with Tracer v1.5 when the ESS for all the parameters was >200.

RESULTS

Clinical *L. Pneumophila* SBT Profiles

L. pneumophila could be cultured from only four (28.57%) of 14 the clinical samples studied in our laboratory, two from pulmonary tissue retrieved during autopsy and the other two from sputums. All of them were typed as ST23 (Table 1), in agreement with the results obtained by Public Health England (PHE) from the first travel-associated cases related to the hotel that were reported after returning to their country of origin (Vanaclocha et al., 2012). Eleven uncultured clinical samples (one of them was also culture-positive) were analyzed by direct amplification and sequencing as described. Three of them provided a complete genetic profile (ST23) and an additional sample (163655) resulted in a new profile that differed from ST23 in two loci (*fliC* and *mompS*; Table 1). Other two samples (191712 and 192147) presented very divergent profiles from ST23. However, some of the sequencing reads from these and two other samples (192061 and 163655) presented unresolved nucleotides at several positions, which led us to seek determining their sequence with better precision.

Cloning and sequencing of PCR products from three loci (*fliC*, *pilE*, and *asd*; Table 2) from the 4 uncultured sputums detailed above revealed the presence of two variants (*asd*-1 and *asd*-15) in patient 191712, providing evidence for the presence of

TABLE 1 | Sequence-based Typing of *L. pneumophila* strains from clinical samples.

Sample	Type of sample	<i>fliC</i>	<i>pilE</i>	<i>asd</i>	<i>mip</i>	<i>mompS</i>	<i>proA</i>	<i>neuA</i>	ST
125	Culture	2	3	9	10	2	1	6	23
191712	Sputum	6	10	1,15	13	–	14	6	(578)
192091	Culture	2	3	9	10	2	1	6	23
	Sputum	–	3	9	10	2	1	6	23
163655	Sputum	6	3	9	10	9	1	6	A*
192147	Sputum	6	10	1	13	2	–	6	–
50291	Culture	2	3	9	10	2	1	6	23
50726	Culture	2	3	9	10	2	1	6	23
191620	Lung puncture	2	3	9	10	2	1	6	23
193975	Sputum	6	3	–	–	2	1	6	–
9138	Sputum	–	–	9	–	–	–	–	–
160613	Sputum	2	3	9	10	2	1	6	23
18484	Sputum	2	3	9	–	2	1	6	–
160717	BAS	2	3	9	10	2	1	6	23

Shadowed bold numbers correspond to PCR products that were cloned for further genetic testing. BAS, Broncho-alveolar aspirate.

*ST-A corresponds to a new profile that could not be assigned a ESGLI profile because the results were derived from an uncultured sample.

TABLE 2 | List of clinical and environmental biofilm samples subjected to cloning of PCR products.

Sample	C/E ^a	Allele pattern ^b	ST	Implication
191712	C	6,10,1/15 ,13,X,14,6	578 ^c	Patient
192147	C	6,10,1 ,13,2,X,6	–	Patient
163655	C	6,3,9 ,10,9,1,6	B	Patient
192091	C	X,3,9 ,10,2,1,6	23 ^c	Patient
cal12_18	E	+,+,+,+,+,+,+	+	Hotel room with 2 cases (1 death)
cal12_40	E	6,10,15 ,X,9,14,6	578 ^c	Hotel room with 1 case (1 death)
cal12_46	E	6,10,15 ,13,9,14,6	578	Hotel room with 2 cases (1 death)
cal12_97	E	6,10,+,+,9,14,+	578 ^c	Hotel room with 1 case
cal12_151	E	6,10,15 ,13,9,14,6	578	Hotel room with 2 cases (2 deaths)
F3	E	2/6,3,9 ,10,2,1,6	23	Air filter

Underlined bold regions were used (*flhC*, *pilE*, *asd*, *mip*, *mompS*, *proA*, *neuA*). Selected biofilm samples were taken from different rooms.

^aC, clinical; E, environmental.

^bX, PCR amplification failed; +, positive amplification but allele assignment failed.

^cIncomplete allele pattern but compatible with ST23 or ST578.

an infecting strain that matched that found in patient 192147. It is worth mentioning that these two patients were a couple who had stayed in the same hotel room and that although the corresponding profiles missed one locus each, their combination corresponded to ST578 (Table 1).

Environmental *L. pneumophila* SBT Profiles

From the sampled biofilms, 110 (67.1%) resulted positive for at least one of the two SBT loci initially tested, *pilE* and *asd*. PCR and further sequencing of the remaining other loci revealed that a profile compatible with ST578 was frequently found (28.2%, 31/110) in samples from most of the rooms studied and other common hotel facilities (Table 3). Two additional STs, ST1136, and ST23, could be identified by SBT from the uncultured samples, with two and one representatives respectively. These two STs differ in only one nucleotide in locus *flhC*. Only one biofilm swab was taken from the spa at the end of January 2012 and it tested negative.

Twenty positive cultures (3.2%) were obtained from water samples (Table 3), revealing mostly *L. pneumophila* serogroup 1 MAAb type Allentown/France ST23 ($n = 9$) and *L. micdadei* ($n = 4$). Two additional STs were identified from these cultured samples, ST1236 and ST1358. ST1236 has a SBT profile which is genetically close to that of ST578, with a difference in only 5 nucleotide sites (Figure 2). In addition, three samples (4039, 4857, and A147) presented a mixture of different *L. pneumophila* STs and *L. micdadei* when several colonies from the same plate were sequenced (Table 3).

The link between the environmental strain ST23 found in the spa pool and most clinical cases was firmly established by the corresponding SBT sequencing results. However, only the direct SBT analysis of uncultured environmental and clinical samples provided evidence for the existence of a second infecting

strain found only in the domestic water system, compatible with ST578, which can be linked reliably to at least one clinical case (Tables 1, 2). In general, the phylogenetic tree derived from all these sequence data revealed links between all the clinical cases and environmental samples from the hotel (Figure 2).

Smoke-tracing studies and airflow dynamics models showed that aerosols from the spa area could reach easily the hotel hall by different ways, including the air-conditioning system. Pieces of the air filters collected from the air-handling unit for the hall area were further analyzed by direct SBT. Five of the six filter samples analyzed were positive for two loci of the bacteria (*pilE* and *asd*) by TD-PCR. A mixed pattern was detected for locus *flhC* in one of the samples, which prompted us to use the same procedure for cloning and sequencing of PCR products described previously. The analysis of sequences derived from filter samples revealed the presence of strain ST23, thus providing clear evidence of the presence of this ST in the air-conditioning system of the hotel, and also a compatible profile (ST1164) which differs from ST23 in only one polymorphic site (Tables 2, 3). The sequences obtained from these samples were also included in the phylogenetic tree shown in Figure 2.

ST23 Colonization History Analyzed by WGS

Five ST23 isolates (four clinical and one obtained from the spa pool) were subjected to whole-genome sequencing (WGS) to assess the intra-ST23 variability and to gain insight about its colonization history. Additionally, one ST1236 isolate was included to check its relatedness to ST578, from which it is genetically close according to the SBT data (Table 4).

Seventy-one polymorphic sites were detected in the core genome of the five ST23 strains (3,267,949 bp). Twenty-four were estimated as homoplastic and removed from further analyses to account for the effect of recombination, leaving 47 non-homoplastic SNPs. The average number of core SNPs between ST23 isolates was estimated to be 22.4 (range 9–34; 6.85E-06 substitutions per site, *s/s*) (Figure 3). Strain 918 (ST1236) showed a high divergence from the ST23 strains, with approximately 38,810 SNPs between both STs (core genome size 2,844,379 bp; 0.0136 *s/s*). SBT data showed just 5 nucleotide differences from ST578 (Figure 2). However, at the genome level, these two strains differ in more than 21,441 positions (0.0083 *s/s*).

Using a Bayesian inference approach, the substitution rate for the five ST23 strains was estimated to be 7.76E-07 *s/s/y* (95% Highest Posterior Density (HPD) interval 2.32E-06–2.09E-08), corresponding to an accumulation of approximately 2.5 nucleotide changes in the core genome per year. Hence, these isolates started to diverge from their most recent common ancestor by approximately year 1994 (95% HPD, 1956–2011), much earlier than the hotel opening (2006).

DISCUSSION

We have described the results of the epidemiological, environmental and genetic investigations of a complex and long-lasting legionellosis outbreak that occurred between November 2011 and July 2012 in a hotel (Calpe, Spain) with

TABLE 3 | Results of the molecular investigation of environmental samples.

Date	Sample	Location	N	Type of sample	fljC	pilE	asd	mip	mompS	proA	neuA	ST/Species	Serogroup/MAB type
Jan-March 2012	Several	Hotel rooms, hall, restrooms	5	Biofilm	6	10	15	13	9	14	6	578	-
Jan-March 2012	Several	Hotel rooms, hall, restrooms	26	Biofilm	(6)	(10)	(15)	(13)	(9)	(14)	(6)	(578) (at least 5 loci)	-
Jan-March	Several	Hotel rooms, hall, restrooms	74	Biofilm	+	+	+	+	+	+	+	Incomplete STs (<5 loci)	-
07/07/2012	PF2	Air filter	1	Biofilm	6	3	-	-	-	-	-	-	-
07/07/2012	PF3	Air filter	1	Biofilm	-	3	9	-	-	1	6	-	-
07/07/2012	F1	Air filter	1	Biofilm	6	3	9	10	2	1	6	1164	-
07/07/2012	F2	Air filter	1	Biofilm	6	3	9	10	2	1	6	1164	-
07/07/2012	F3	Air filter	1	Biofilm	2	3	9	10	2	1	6	23	-
31/01/2012	918	Spa pool	2	Culture	6	10	15	10	21	40	6	1236	1
11/02/2012	12392	Spa pool	1	Culture	2	3	9	10	2	1	6	23	1
23/02/2012	2084	Spa pool	1	Culture	-	-	-	-	-	-	-	<i>L. micdadei</i>	-
02/05/2012	4855	Spa pool	9	Culture	-	-	-	-	-	-	-	<i>L. micdadei</i>	-
02/05/2012	4856	Spa pool	12	Culture	2	3	9	10	2	1	6	23	1
26/06/2012	8298	Room, cold water	9	Culture	-	-	-	-	-	-	-	-	2-15
04/07/2012	Al657	90 m3 water tank	10	Culture	2	3	9	10	2	1	6	23	1
09/07/2012	8965	Osmosis (in)	7	Culture	2	3	9	10	2	1	6	23	1
09/07/2012	8966	Osmosis (out)	5	Culture	2	3	9	10	2	1	6	23	1
09/07/2012	8968	Spa (fountain)	10	Culture	2	3	9	10	2	1	6	23	1
18/07/2012	69135	Spa pool	1	Culture	-	-	-	-	-	-	-	<i>L. micdadei</i>	-
18/07/2012	69147	Spa pool	1	Culture	-	-	-	-	-	-	-	<i>L. micdadei</i>	-
18/07/2012	69148	Osmosis device	1	Culture	2	3	9	10	2	1	6	23	1 Allentown/France
18/07/2012	69153	Spa pool	1	Culture	2	3	9	10	2	1	6	23	1 Allentown/France
18/07/2012	69183	Spa pool drainer	1	Culture	2	3	9	10	2	1	6	23	1 Allentown/France
27/07/2012	Al763	Watering network (hotel entrance)	1	Culture	5	2	22	10	6	25	203	1358	8
23/08/2012	10636	Water deposit	1	Culture	-	-	-	-	-	-	-	<i>L. anisa</i>	-
12/04/2012	4029	Spa pool	4	Culture	2	3	9	10	2	1	6	23	1
12/04/2012	4029	Spa pool	1	Culture	6	10	15	10	21	40	6	1236	1
12/04/2012	4029	Spa pool	1	Culture	-	-	-	-	-	-	-	<i>L. micdadei</i>	-
02/05/2012	4857	Spa pool	3	Culture	2	3	9	10	2	1	6	23	1 Allentown/France
02/05/2012	4857	Spa pool	1	Culture	6	10	15	10	21	40	6	1236	1 Allentown/France
02/05/2012	4857	Spa pool	1	Culture	5	2	22	10	6	25	203	1358	10
02/05/2012	4857	Spa pool	2	Culture	-	-	-	-	-	-	-	<i>L. micdadei</i>	-
31/01/2012	Al47	Spa pool	3	Culture	6	10	15	10	21	40	6	1236	1
31/01/2012	Al47	Spa pool	3	Culture	-	-	-	-	-	-	-	<i>L. micdadei</i>	-

Biofilm samples from hotel rooms, hall and restrooms (n = 164) were analyzed directly as none produced positive cultures. The results of those samples in which *L. pneumophila* was detected (n = 105) are shown in the first three rows depending on whether they yielded a complete ST (at least five alleles determined to be identical to those defining the ST), or an incomplete ST (less than five alleles determined). For some cultures, several colonies were sequenced, as detailed (N when type of sample is "Culture"). The last three cultured samples revealed a complex composition with two or more STs and/or another *Legionella* species.

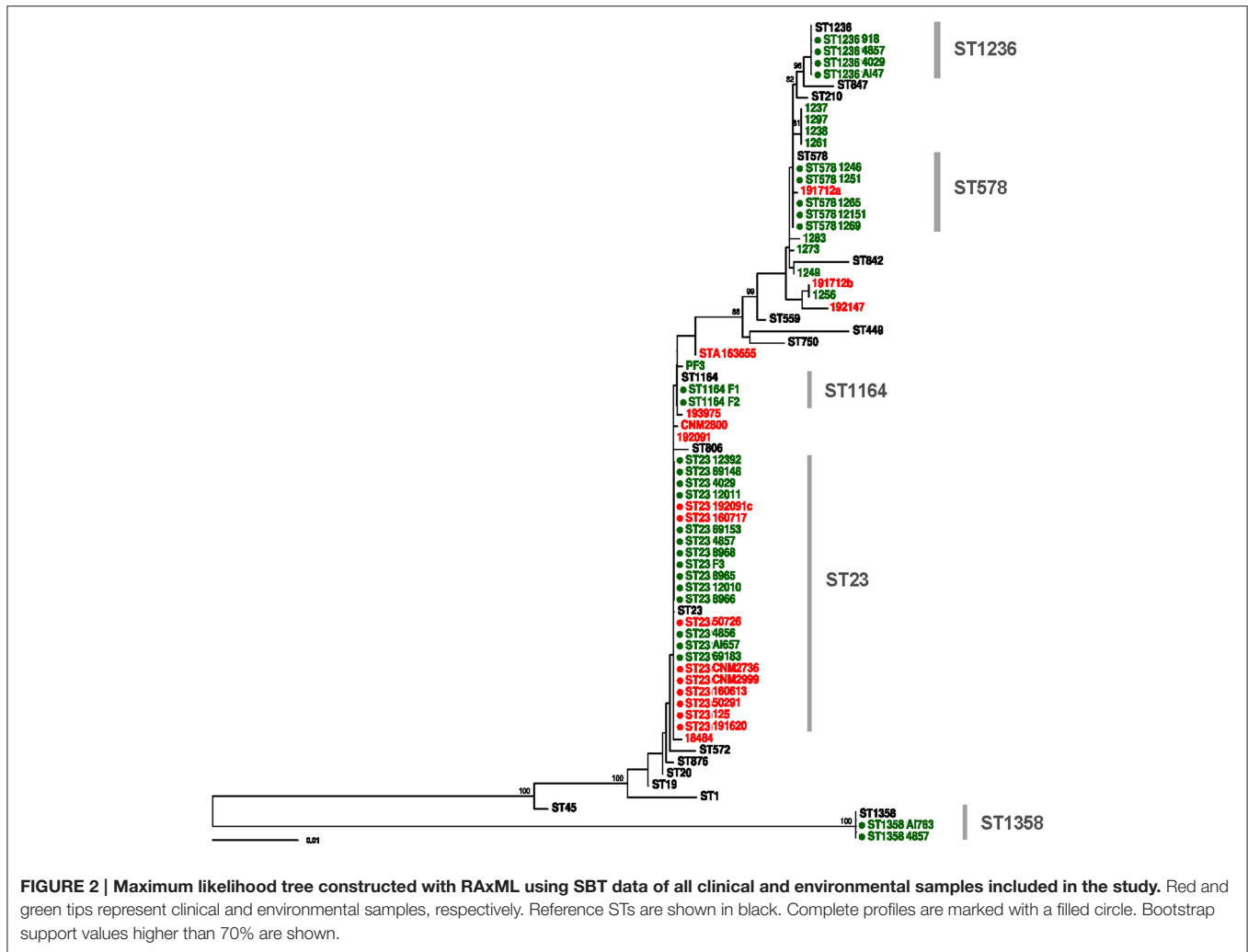


TABLE 4 | Accession numbers to the fastq files of the 6 *L. pneumophila* strains sequenced using SOLiD 5500XL during the outbreak investigation deposited in the European Nucleotide Archive (ENA) under project accession PRJEB5990.

Sample	Strain	Experiment	Run	ST	Source
125	ERS497776	ERX516043	ERR556974	23	Clinical
192091	ERS497781	ERX516048	ERR556979	23	Clinical
4029	ERS497807	ERX516074	ERR557005	23	Environmental
50291	ERS497828	ERX516095	ERR557026	23	Clinical
50726	ERS497830	ERX516097	ERR557028	23	Clinical
918	ERS497849	ERX516116	ERR557047	1236	Environmental

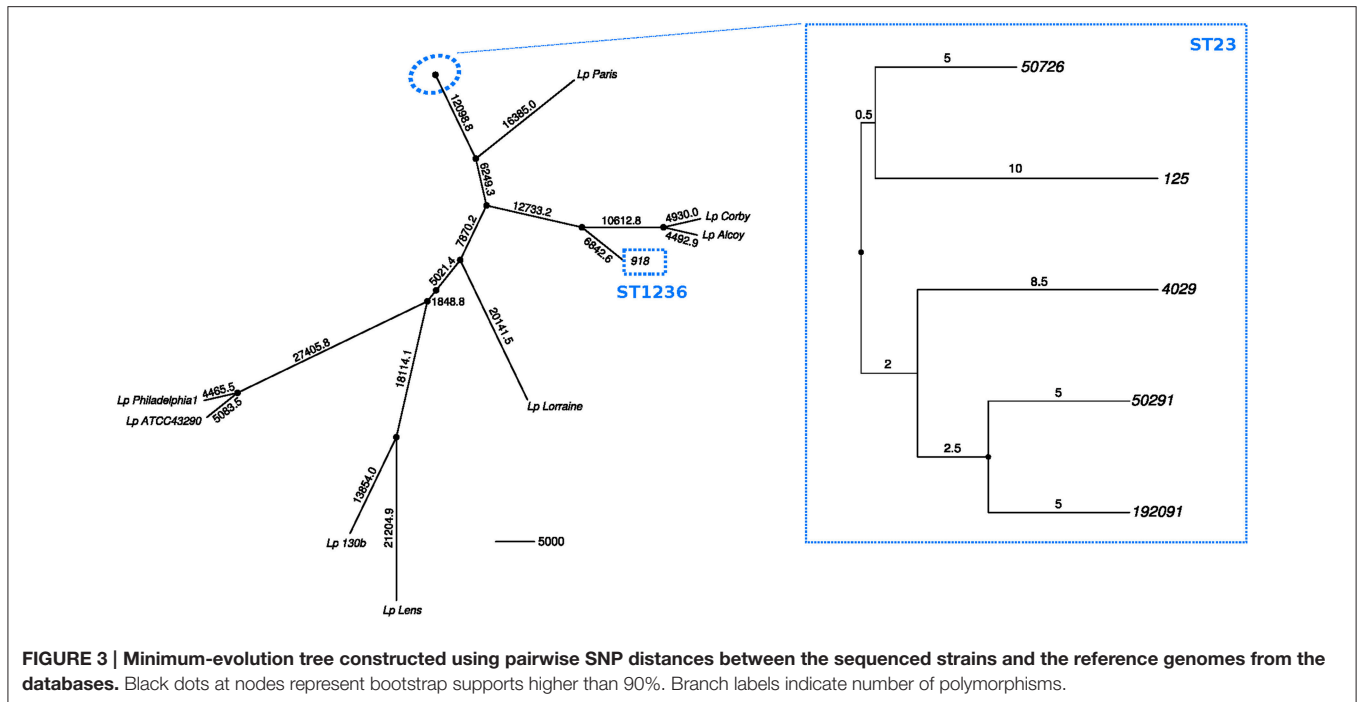
Further details on the sequenced samples are given in **Tables 1–3**.

44 cases and six fatal outcomes. During the outbreak, different control measures were adopted, which included repeated cleaning and disinfections of the different water systems of the hotel and two temporary hotel closures (**Figure 1**). The outbreak was controlled only when the hotel spa pool was permanently closed and completely dismantled. A new spa pool was built with

renewed air-conditioning ducts. No new cases of legionellosis related to this hotel have appeared since then.

The initial epidemiological investigation revealed a clear link to the hotel that prompted the first control measures. The first signs of the presence of this bacterium in the hotel were obtained by the direct amplification of two loci from biofilm samples in several hotel rooms only 10 days after the cleaning and disinfection and led to the first closure of the hotel (Vanaclocha et al., 2012). Remarkably, the corresponding water samples were negative for *Legionella* by culture. Repeated samples from the same spots revealed that biofilm-containing *Legionella* recolonized most facilities only a few days after their apparently removal by chlorine treatment.

Most of the sequenced clinical isolates and direct samples corresponded to ST23 whereas a profile compatible with ST578 was repeatedly identified in biofilm samples of the domestic water system. The only exception corresponded to a couple (samples 192712 and 192147) who shared the same room and whose incompletely determined ST differed in five of the six sequenced loci from ST23 but were compatible with ST578 (**Table 1**). Additional cloning and sequencing of PCR products from these



patients revealed the presence of common strains in both of them, thus confirming the existence of a second infecting strain along the outbreak which, in turn, suggested that the domestic water system was another source of infection. Several recent reports (Yiallourous et al., 2013; Coscollá et al., 2014; McAdam et al., 2014; Wewalka et al., 2014) have shown the presence of mixed *Legionella* infections in outbreak patients. Here, we have shown the presence of more than one strain co-existing in the same location and simultaneously producing infections that led to an outbreak.

After the first cluster of cases, ST23 was isolated for the first time in the spa pool, located at the hotel basement. However, the initial epidemiological investigation failed to find a link between many of the patients infected with *L. pneumophila* ST23 and the spa area. In July 2012, a thorough environmental investigation of the spa structure led to the discovery of multiple hidden cavities with stagnant water under the pool vessel that had connections to the bathing water. Subsequent environmental studies suggested the dissemination of ST23 from the spa pool to the hall area through different ways including the air-conditioning system. The identification of *Legionella* profiles compatible with ST23 in the air filters of the air-handling unit for the hall area confirmed the implication of the air-conditioning system in the dissemination of this strain. This helped to explain the detection of this ST in samples from patients who had not visited the spa. Sequencing of isolates from water samples revealed mainly ST23, ST1236, and *Legionella micdadei* as colonizers in different parts of the spa pool, a reverse osmosis plant, a drinking fountain at the spa and a pressure vessel of a non-potable water system.

Genomic data can be used to estimate the evolutionary rates of living organisms (Harris et al., 2010; Bryant et al., 2013; Chewapreecha et al., 2014; Sánchez-Busó et al., 2014), and even

the time at which these organisms could have colonized a specific location (Sánchez-Busó et al., 2014). Only three works up to date have used high throughput sequencing to the retrospective study of legionellosis outbreaks (Reuter et al., 2013; Lévesque et al., 2014; McAdam et al., 2014). In the latter case, *L. pneumophila* isolation was possible only from 15/91 patients. In our case, isolation was only successful in 4/14 of the clinical samples. The difficulty in obtaining *L. pneumophila* pure cultures from all the samples is reflected in the subsequent poor representation of the outbreak strains in NGS.

Using Bayesian inference, we estimated the substitution rate of the non-recombinant core genome of the ST23 strains within the hotel as $7.76E-07$ s/s/y, higher but of the same order of magnitude than that estimated for the persistent ST578 in the locality of Alcoy ($1.39E-07$ s/s/y; Sánchez-Busó et al., 2014). The ST23 strains found in the spa were not identical, as would be expected from a single clone colonizing the spa and causing an outbreak. Besides, the time at which ST23 could have started proliferating was estimated to be approximately year 1994. Despite the long HPD intervals from the Bayesian analysis (1956–2011), the finding of 71 polymorphic sites (47 non-homoplasic) reveals a higher diversity than expected. These results, as well as the finding of other STs and *Legionella* species colonizing the hotel, provide evidence that there was a complex community of *L. pneumophila* inhabiting the area previous to the building of the hotel, including a population of ST23. Thus, the hotel could have been colonized by different STs during its construction and before its opening to the public in 2006. Moreover, there was a precedent of legionellosis cases in the period 2006–2007 related to the hotel, but no genetic information about the causing strain was retrieved at that moment, so further comparisons between both clusters cannot be performed.

The results derived from the investigation of this outbreak revealed that information from both cultured and uncultured samples collected during the epidemiological investigation of a complex legionellosis outbreak can be essential for its clarification. Moreover, WGS data from an outbreak investigation can help, not only to the epidemiological investigation in order to detect the main reservoir, but it also allows the application of inference tools to trace back the local evolution of the pathogen. The information retrieved from this type of analysis can be important for public health officials to better understand the local behavior of a pathogen such as *L. pneumophila*, capable of colonizing rapidly several facilities in a complex building and remain silently evolving until conditions favor its multiplication and spread, resulting in the infection of exposed, susceptible persons.

ACKNOWLEDGMENTS

This work was supported by Ministerio de Economía y Competitividad (MINECO), Spanish Government (project BFU2011-24112 and BFU2014-58656-R) and Dirección General de Salud Pública (Generalitat Valenciana, Spain). LSB was recipient of a Formación de Profesorado Universitario (FPU) predoctoral fellowship from Ministerio de Educación, Spanish Government. The funders had no role in study design, data

collection and analysis, decision to publish, or preparation of the manuscript. The authors would like to acknowledge the contributions of all members of the investigation team: R Carbo, M Castellanos, C Marin, E Perez (Surveillance and Epidemiologic Control Unit. Public Health Directorate, Regional Government of Valencia, Spain); J Añó, C del Hierro (Environmental Health Unit. Public Health Directorate, Regional Government of Valencia, Spain); J Miralles, J Miravet, E Rubiales (Environmental Health Unit, Public Health Center, Denia, Spain); R Gonzalez (Epidemiological Unit, Public Health Center, Denia, Spain); J Calafat (Environmental Health Unit, Public Health Center, Alcoi, Spain); C Gonzalez (Microservices, Benidorm, Spain); J del Barrio, B Santandreu (Biolinea Int.); P Santa-Olalla Peralta (Coordinating Centre for Health Alerts and Emergencies, Ministry of Health, Social Services and Equity, Madrid, Spain) and V Drasar (National Legionella Reference Laboratory, Public Health Institute Ostrava, Czech Republic). We would also like to thank the members of the visiting ECDC team (B. de Jong, C.L.R. Bartlett, and J. Walker) for their suggestions and recommendations at the closing of the investigation as well as the Legionella labs of Public Health England (PHE) and Instituto de Salud Carlos III (Madrid, Spain) for sharing the results of their typing of clinical samples related to this outbreak.

REFERENCES

- Benin, A. L., Benson, R. F., Arnold, K. E., Fiore, A. E., Cook, P. G., Williams, L., et al. (2002). An outbreak of travel-associated Legionnaires disease and Pontiac fever: the need for enhanced surveillance of travel-associated legionellosis in the United States. *J. Infect. Dis.* 185, 237–243. doi: 10.1086/338060
- Bryant, J. M., Grogono, D. M., Greaves, D., Foweraker, J., Roddick, I., Inns, T., et al. (2013). Whole-genome sequencing to identify transmission of *Mycobacterium abscessus* between patients with cystic fibrosis: a retrospective cohort study. *Lancet* 381, 1551–1560. doi: 10.1016/S0140-6736(13)60632-7
- Buchbinder, S., Leitritz, L., Trebesius, K., Banas, B., and Heesemann, J. (2004). Mixed lung infection by *Legionella pneumophila* and *Legionella gormanii* detected by fluorescent *in situ* hybridization. *Infection* 32, 242–245. doi: 10.1007/s15010-004-2127-z
- Burnsed, L. J., Hicks, L. A., Smithee, L. M. K., Fields, B. S., Bradley, K. K., Pascoe, N., et al. (2007). A large, travel-associated outbreak of Legionellosis among hotel guests: utility of the Urine antigen assay in confirming Pontiac fever. *Clin. Infect. Dis.* 44, 228. doi: 10.1086/510387
- Cazalet, C., Rusniok, C., Brüggemann, H., Zidane, N., Magnier, A., Ma, L., et al. (2004). Evidence in the *Legionella pneumophila* genome for exploitation of host cell functions and high genome plasticity. *Nat. Genet.* 36, 1165–1173. doi: 10.1038/ng1447
- Chewapreecha, C., Harris, S. R., Croucher, N. J., Turner, C., Marttinen, P., Cheng, L., et al. (2014). Dense genomic sampling identifies highways of pneumococcal recombination. *Nat. Genet.* 46, 305–309. doi: 10.1038/ng.2895
- Chien, M., Morozova, I., Shi, S., Sheng, H., Chen, J., Gomez, S. M., et al. (2004). The genomic sequence of the accidental pathogen *Legionella pneumophila*. *Science* 305, 1966–1968. doi: 10.1126/science.1099776
- Coscollá, M., Fernández, C., Colomina, J., Sánchez-Busó, L., and González-Candelas, F. (2014). Mixed infection by *Legionella pneumophila* in outbreak patients. *Int. J. Med. Microbiol.* 304, 307–313. doi: 10.1016/j.ijmm.2013.11.002
- Coscollá, M., and González-Candelas, F. (2007). Population structure and recombination in environmental isolates of *Legionella pneumophila*. *Environ. Microbiol.* 9, 643–656. doi: 10.1111/j.1462-2920.2006.01184.x
- Coscollá, M., and González-Candelas, F. (2009). Direct sequencing of *Legionella pneumophila* from respiratory samples for sequence-based typing analysis. *J. Clin. Microbiol.* 47, 2901–2905. doi: 10.1128/JCM.00268-09
- D'Auria, G., Jiménez-Hernández, N., Peris-Bondia, F., Moya, A., and Latorre, A. (2010). *Legionella pneumophila* pangenome reveals strain-specific virulence factors. *BMC Genomics* 11:181. doi: 10.1186/1471-2164-11-181
- Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K., and Mattick, J. S. (1991). "Touchdown" PCR to circumvent spurious priming during gene amplification. *Nucl. Acids Res.* 19, 4008.
- Drummond, A. J., Suchard, M. A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969–1973. doi: 10.1093/molbev/mss075
- Farhat, C., Mentasti, M., Jacobs, E., Fry, N. K., and Lück, C. (2011). The N-acetylneuraminase cytidyltransferase gene, *neuA*, is heterogenous in *Legionella pneumophila* strains but can be used as a marker for epidemiological typing in the consensus sequence-based typing scheme. *J. Clin. Microbiol.* 49, 4052–4058. doi: 10.1128/JCM.00687-11
- Fraser, D. W., Tsai, T. R., Orenstein, W., Parkin, W. E., Beecham, H. J., Sharrar, R. G., et al. (1977). Legionnaires' disease: description of an epidemic of pneumonia. *N. Engl. J. Med.* 297, 1189–1197. doi: 10.1056/NEJM197712012972201
- Gaia, V., Fry, N. K., Afshar, B., Lück, P. C., Meugnier, H., Etienne, J., et al. (2005). Consensus sequence-based scheme for epidemiological typing of clinical and environmental isolates of *Legionella pneumophila*. *J. Clin. Microbiol.* 43, 2047–2052. doi: 10.1128/JCM.43.5.2047-2052.2005
- Gaia, V., Fry, N. K., Harrison, T. G., and Peduzzi, R. (2003). Sequence-based typing of *Legionella pneumophila* serogroup 1 offers the potential for true portability in legionellosis outbreak investigation. *J. Clin. Microbiol.* 41, 2932–2939. doi: 10.1128/JCM.41.7.2932-2939.2003
- Ginevra, C., Lopez, M., Forey, F., Reyrolle, M., Meugnier, H., Vandenesch, F., et al. (2009). Evaluation of a nested-PCR-derived sequence-based typing method applied directly to respiratory samples from patients with Legionnaires' disease. *J. Clin. Microbiol.* 47, 981–987. doi: 10.1128/JCM.02071-08
- Glick, T. H., Gregg, M. B., Berman, B., Mallison, G., Rhodes, W. W. Jr., and Kassinoff, I. (1978). Pontiac fever. An epidemic of unknown etiology in a

- health department: I. Clinical and epidemiologic aspects. *Am. J. Epidemiol.* 107, 149–160.
- Gomez-Valero, L., Rusniok, C., Jarraud, S., Vacherie, B., Rouy, Z., Barbe, V., et al. (2011). Extensive recombination events and horizontal gene transfer shaped the *Legionella pneumophila* genomes. *BMC Genomics* 12:536. doi: 10.1186/1471-2164-12-536
- Harris, S. R., Feil, E. J., Holden, M. T. G., Quail, M. A., Nickerson, E. K., Chantratita, N., et al. (2010). Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 327, 469–474. doi: 10.1126/science.1182395
- Joly, J. R., Chen, Y. Y., and Ramsay, D. (1983). Serogrouping and subtyping of *Legionella pneumophila* with monoclonal antibodies. *J. Clin. Microbiol.* 18, 1040–1046.
- Joly, J. R., McKinney, R. M., Tobin, J. O., Bibb, W. F., Watkins, I. D., and Ramsay, D. (1986). Development of a standardized subgrouping scheme for *Legionella pneumophila* serogroup 1 using monoclonal antibodies. *J. Clin. Microbiol.* 23, 768–771.
- Lévesque, S., Plante, P. L., Mendis, N., Cantin, P., Marchand, G., Charest, H., et al. (2014). Genomic characterization of a large outbreak of *Legionella pneumophila* serogroup 1 strains in Quebec City, 2012. *PLoS ONE* 9:e103852. doi: 10.1371/journal.pone.0103852
- Matsui, M., Fujii, S. I., Shirowa, R., Amemura-Maekawa, J., Chang, B., Kura, F., et al. (2010). Isolation of *Legionella rubrilucens* from a pneumonia patient co-infected with *Legionella pneumophila*. *J. Med. Microbiol.* 59, 1242–1246. doi: 10.1099/jmm.0.016089-0
- McAdam, P. R., vander broek, C. W., Lindsay, D. S. J., Ward, M. J., Hanson, M. F., Gillies, M., et al. (2014). Gene flow in environmental *Legionella pneumophila* leads to genetic and pathogenic heterogeneity within a Legionnaires' disease outbreak. *Genome Biol.* 15, 504. doi: 10.1186/s13059-014-0504-1
- Mentasti, M., Fry, N. K., Afshar, B., Palepou-Foxley, C., Naik, F. C., and Harrison, T. G. (2012). 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.* 31, 2017–2028. doi: 10.1007/s10096-011-1535-0
- Meyer, R. D., Edelstein, P. H., Kirby, B. D., Louie, M. H., Mulligan, M. E., Morgenstein, A. A., et al. (1980). Legionnaires' disease: unusual clinical and laboratory features. *Ann. Intern. Med.* 93, 240–243. doi: 10.7326/0003-4819-93-2-240
- Muder, R. R., Yu, V. L., Vickers, R. M., Rihs, J., and Shonnard, J. (1983). Simultaneous infection with *Legionella pneumophila* and Pittsburgh pneumonia agent. Clinical features and epidemiologic implications. *Am. J. Med.* 74, 609–614. doi: 10.1016/0002-9343(83)91018-5
- Reuter, S., Harrison, T. G., Köser, C. U., Ellington, M. J., Smith, G. P., Parkhill, J., et al. (2013). A pilot study of rapid whole-genome sequencing for the investigation of a *Legionella* outbreak. *BMJ Open* 3:e002175. doi: 10.1136/bmjopen-2012-002175
- Rota, M. C., Scaturro, M., Fontana, S., Foroni, M., Boschetto, G., Trentin, L., et al. (2011). Cluster of travel-associated Legionnaires' disease in Lazise, Italy, July to August 2011. *Eurosurveillance* 16:19982. Available online at: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19982>
- Rzhetsky, A., and Nei, M. (1993). Theoretical foundation of the minimum-evolution method of phylogenetic inference. *Mol. Biol. Evol.* 10, 1073–1095.
- Sánchez-Busó, L., Comas, I., Jorques, G., and González-Candelas, F. (2014). Recombination drives genome evolution in outbreak-related *Legionella pneumophila* isolates. *Nat. Genet.* 46, 1205–1211. doi: 10.1038/ng.3114
- Sánchez-Busó, L., Olmos, M. P., Camaró, M. L., Adrián, F., Calafat, J. M., and González-Candelas, F. (2015). Phylogenetic analysis of environmental *Legionella pneumophila* isolates from an endemic area (Alcoy, Spain). *Infect. Genet. Evol.* 30, 45–54. doi: 10.1016/j.meegid.2014.12.008
- Schroeder, G. N., Petty, N. K., Mousnier, A., Harding, C. R., Vogrin, A. J., Wee, B., et al. (2010). *Legionella pneumophila* strain 130b possesses a unique combination of type IV secretion systems and novel Dot/Icm secretion system effector proteins. *J. Bacteriol.* 192, 6001–6016. doi: 10.1128/JB.00778-10
- Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690. doi: 10.1093/bioinformatics/btl446
- Steinert, M., Heuner, K., Buchrieser, C., Albert-Weissenberger, C., and Glöckner, G. (2007). *Legionella* pathogenicity: genome structure, regulatory networks and the host cell response. *Int. J. Med. Microbiol.* 297, 577–587. doi: 10.1016/j.ijmm.2007.03.009
- Suenaga, E., and Nakamura, H. (2005). Evaluation of three methods for effective extraction of DNA from human hair. *J. Chromatogr. B* 820, 137–141. doi: 10.1016/j.jchromb.2004.11.028
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121
- Vanaclocha, H., Guiral, S., Morera, V., Calatayud, M. A., Castellanos, M., Moya, V., et al. (2012). Preliminary report: outbreak of Legionnaires disease in a hotel in Calp, Spain, update on 22 February 2012. *Eurosurveillance* 17:20093. Available online at: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20093>
- Wewalka, G., Schmid, D., Harrison, T.G., Uldum, S. A., Lück, C., and the European Society of Clinical Microbiology Infectious Diseases Study Group for Legionella Infections (ESGLI) (2014). Dual infections with different *Legionella* strains. *Clin. Microbiol. Infect.* 20, O13–O19. doi: 10.1111/1469-0691.12311
- Yiallourou, P. K., Papadouri, T., Karaoli, C., Papamichael, E., Zeniou, M., Pieridou-Bagatzouni, D., et al. (2013). First outbreak of nosocomial *Legionella* infection in term neonates caused by a cold mist ultrasonic humidifier. *Clin. Infect. Dis.* 57, 48–56. doi: 10.1093/cid/cit176
- Zerbino, D. R. (2010). Using the Velvet *de novo* assembler for short-read sequencing technologies. *Curr. Protoc. Bioinformatics* 31, 11.5.1–11.5.12. doi: 10.1002/0471250953.bi1105s31

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Sánchez-Busó, Guiral, Crespi, Moya, Camaró, Olmos, Adrián, Morera, González-Morán, Vanaclocha and González-Candelas. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.