Distribution of Sequence-Based Types of *Legionella pneumophila* Serogroup 1 Strains Isolated from Cooling Towers, Hot Springs, and Potable Water Systems in China

Tian Qin,a,b Haijian Zhou,a,b Hongyu Ren,a,b Hong Guan,a Machao Li,a,b Bingqing Zhu,a,b Zhujun Shao,a,b

National Institute for Communicable Disease Control and Prevention and State Key Laboratory for Infectious Disease Prevention and Control, Chinese Centre for Disease Control and Prevention, Beijing, China; Collaborative Innovation Centre for Diagnosis and Treatment of Infectious Diseases, Hangzhou, China

*Legionella pneumophila* serogroup 1 causes Legionnaires’ disease. Water systems contaminated with *Legionella* are the implicated sources of Legionnaires’ disease. This study analyzed *L. pneumophila* serogroup 1 strains in China using sequence-based typing. Strains were isolated from cooling towers (n = 96), hot springs (n = 42), and potable water systems (n = 26). Isolates from cooling towers, hot springs, and potable water systems were divided into 25 sequence types (STs; index of discrimination [IOD], 0.711), 19 STs (IOD, 0.934), and 3 STs (IOD, 0.151), respectively. The genetic variation among the potable water isolates was lower than that among cooling tower and hot spring isolates. ST1 was the predominant type, accounting for 49.4% of analyzed strains (n = 81), followed by ST154. With the exception of two strains, all potable water isolates (92.3%) belonged to ST1. In contrast, 53.1% (51/96) and only 14.3% (6/42) of cooling tower and hot spring, respectively, isolates belonged to ST1. There were differences in the distributions of clone groups among the water sources. The comparisons among *L. pneumophila* strains isolated in China, Japan, and South Korea revealed that similar clones (ST1 complex and ST154 complex) exist in these countries. In conclusion, in China, STs had several unique allelic profiles, and ST1 was the most prevalent sequence type of environmental *L. pneumophila* serogroup 1 isolates, similar to its prevalence in Japan and South Korea.

Bacteria of the genus *Legionella*, widely present in water, contribute to human diseases. To date, more than 50 *Legionella* species have been characterized, and 25 species are known to cause human disease (1, 2). *Legionella* was first reported as a pathogen following a pneumonia outbreak in 1976 (3). Since then, several community-, hospital-, and travel-associated *Legionella* outbreaks have been reported. The most severe form of pneumonia caused by *Legionella* is known as Legionnaires’ disease; however, there is also a milder, flu-like form known as Pontiac fever (4). *Legionella pneumophila* is responsible for approximately 90% of human infections (5). *L. pneumophila* is divided into 15 serogroups, of which serogroup 1 is the most prevalent disease-causing variant. Few outbreaks by other *Legionella* species have been reported (5, 6). Transmission of bacteria from the environment to humans occurs via inhalation or aspiration of *Legionella*-containing aerosols (7, 8). Cooling towers (9, 10), hot springs (11, 12), and potable water systems (13, 14) in large facilities, hotels, hospitals, and public baths that are contaminated with *Legionella* are the implicated sources of outbreaks and sporadic cases of Legionnaires’ disease.

The water systems of several countries are highly contaminated with *Legionella*. In our previous study that compared the efficacy of different *Legionella* detection methods, we isolated *Legionella* at rates of 26.39%, 54.44%, and 18.94% from cooling towers, hot springs, and piped water systems, respectively, with *Legionella* concentrations of >1,000 CFU/liter (15). Studies have reported that *L. pneumophila* serogroup 1 is the most frequently isolated species in cooling towers and hot springs (9, 11).

Several subtyping techniques have been used to identify and characterize *L. pneumophila* strains. Sequence-based typing (SBT), which is based on multilocus sequence typing (MLST), is currently used for *L. pneumophila*. According to the European Working Group on Legionella Infections (EWGLI; http://www.ewgli.org/), *L. pneumophila* isolates can be characterized by SBT using seven loci, including five genes associated with virulence (*flIC*, *pilE*, *mip*, *proA*, and *mompS*) and two housekeeping genes (*asd* and *neuA*) (16, 17). SBT is a highly discriminatory method that allows the rapid identification of isolates that are closely related. In this study, we used SBT to assess the genetic characteristics of *L. pneumophila* strains isolated from cooling towers, hot springs, and potable water systems in China. Furthermore, the SBT results of Chinese strains were compared with those of Japan and South Korea.

**MATERIALS AND METHODS**

**Bacterial isolates and culture conditions.** A total of 164 environmental *L. pneumophila* serogroup 1 strains were included in this study. These strains were isolated during regular surveillance events performed in the cities of Shanghai, Beijing, Shenzhen, Wuxi, Jinan, and Shijiazhuang between 2005 and 2012. In this surveillance, water samples from central air conditioning cooling towers and potable water systems in public buildings and hospitals and from indoor and outdoor hot springs of public baths were collected for a *Legionella* survey. The cooling towers were closed, and the hot springs were the sulfuric acid type. The temperature of the water samples from cooling towers and potable water systems was about 25°C, and that of the hot spring samples was between 31°C and 49°C. The pH of the water systems ranged from 6.0 to 8.0. Sampling was conducted via...
identical sampling protocols in all geographic regions. A total of 500 ml of
each water sample from cooling towers, hot springs, or faucets was
collected in sterile, screw-cap containers. All strains used in this study
belonged to the L. pneumophila serogroup 1. Strains were stored at −80°C in
charcoal yeast extract (CYE) with 20% sterile glycerol. The water sources
consisted of cooling towers (96 isolates), hot springs (42 isolates), and
potable water systems (26 isolates) of public buildings, hospitals, and
public baths (Table 1). The bacteria were streaked onto buffered charcoal
yeast extract (BCYE) agar plates. Colonies were identified by serotyping
after they were inoculated onto BCYE agar plates and incubated at 35°C in
2.5% CO₂ for 48 h.

DNA extraction, PCR amplification, and sequencing. Genomic DNA was extracted using a QIAamp DNA minikit (Qiagen, Düsseldorf,
Germany). PCR was performed in a T100 Thermal Cycler (Bio-Rad Labora-
tories, CA, USA) using the primers for seven genes (flaA, pilE, asd, mip,
mompS, proA, and neuA) in SBT analysis (16, 17). The PCR products were
prepared in a reaction volume of 50 μl with 5 μl of 10× PCR buffer (TaKaRa,
Dalian, China), 1 unit of Taq polymerase (TaKaRa), 200 μM concentra-
tion of the deoxynucleosides (dNTPs; TaKaRa), 0.4 μM each primer set,
20 ng of the DNA template, and filtered sterile water. PCR was performed
with an initial denaturation step at 94°C for 5 min followed by 30 cycles,
each consisting of an initial denaturation at 94°C for 40 s followed by
annealing and extension steps. The primers and corresponding annealing
temperatures used were previously described (16, 17). DNA sequencing
was performed using an ABI BigDye Terminator cycle sequencing kit
(Applied Biosystems, CA, USA); the products were analyzed on a 3130xl
ABI Prism genetic analyzer (Applied Biosystems, CA, USA). The SBT database,
available on the EWGLI website, was used for nucleotide analysis. The
sequences were compared with those in the SBT database, which were also
available on the EWGLI website (http://www.hpa-bioinformatics.org.uk/
legionella/legionella_sbt/php/sbt_homepage.php). The index of dis-
crimination (IOD) was calculated using the following formula: IOD =
1 − Σ(ni − nj − 1) |N[N − 1]|, where ni is the number of strains belonging
to the jth pattern, and N is the number of strains in the population (18).
Diversity of SBT loci was calculated by Nei’s index: 1 − Σ ni |N (allele fre-
cency)² (19).

BioNumerics software (version 5.10; Applied Maths, Kortrijk,
Belgium) was used to create minimum spanning trees (MST). In MST,
the founder sequence type (ST) is defined as the ST with the highest
number of single-locus variants. Clusters of related STs that descend from a
common ancestor are defined as clone groups (complexes); single genotypes
that do not correspond to any clone groups are defined as singletons.
Types are represented by circles; the size of a circle indicates the number
of strains of this particular type. Heavy solid lines connect two types
that differ within a single locus, light solid lines connect double-locus variants,
heavy dotted lines connect triple-locus variants, light dotted lines connect
quadruple-locus variants (see Fig. 3 and 4).

Statistical analyses. All calculations were conducted using SPSS,
version 13.0, software (SPSS, Inc., Chicago, IL). A chi-square test was used to
compare the proportional distributions of STs according to different re-

gen regions and environmental water source types.

Nucleotide sequence accession numbers. The sequences of the DNA
fragments encompassing the seven genes flaA, pilE, asd, mip, mompS,
proA, and neuA as determined by SBT analysis were deposited in the
GenBank database under accession numbers KJ160781 to KJ161159.

RESULTS
SBT analysis. In this study, 164 Legionella isolates were analyzed
by SBT and divided into 42 STs (IOD, 0.749); 16 STs were associ-
ated with more than one strain, and 26 STs were identified with
one single strain (Fig. 1). Most of the STs were city specific; how-
ever, ST1, ST154, ST630, ST752, and ST1471 were present in more
than one city. Querying the EWGLI SBT database (available at
http://www.ewgli.org) found that 23 of the 42 STs were unique
to China and 11 (ST1556 to ST1566) were identified for the first
time. Furthermore, of the 16 STs that were identified with more
than one strain, 5 STs (ST149, ST155, ST160, ST1471, and
ST1479) had been previously identified in China, and the other 11
STs were prevalent throughout the world, according to the
EWGLI SBT database. Only one ST (ST59) had been isolated in
clinical cases in China, and 10 STs had been isolated in clinical
cases abroad, according to the EWGLI SBT database.

ST1 (1,4,3,1,1,1,1) (the sequence of numbers represents the
allelic profile, respectively, of the seven genes flaA, pilE, asd, mip,
mompS, proA, and neuA, as shown in Fig. 1), which was identified
with 49.4% of the analyzed strains (n = 81), was present in all
cities. Based on the results, 53.1% (51/96), 14.3% (6/42), and
92.3% (24/26) of cooling tower, hot spring, and potable water
isolates belonged to ST1. There were no regional differences in
the distribution of ST1 in potable water systems (75.0% [6/8] from
Jinan and 100% [18/18] from Shanghai; χ² = 4.875, P = 0.086);
however, the distribution of ST1 in cooling towers was regionally
different (Fig. 2). For cooling tower water isolates, 12.5% (2/16),
20.0% (4/20), 40.0% (2/5), 65.4% (17/26), 75.0% (9/12), and
100% (17/17) of the isolates from Wuxi, Shenzhen, Jinan, Shang-
hai, Beijing, and Shijiazhuang, respectively, belonged to ST1 (χ² =
38.687, P < 0.001). Additionally, 40% (6/15) of the hot spring
isolates from Shenzhen belonged to ST1; however, no ST1 strains
were isolated from hot springs in Beijing (χ² = 6.618, P = 0.018).

ST154 was the second most predominant ST. In this study,
ST154 isolates were from two water sources: the cooling tower
in Shenzhen and the hot spring in Beijing. According to the
EWGLI SBT database, accessed on 8 January 2014, there were 33 isolates
belonging to ST154. These strains were isolated from seven coun-
tries between 1982 and 2013. Among them, 10 strains were clin-
ical, and the others were environmental.

Strains from different sources showed different levels of diver-
sity. The 42 hot spring isolates were divided into 19 STs (IOD,
0.934), the 96 cooling tower isolates were divided into 25 STs
(IOD, 0.711), and the 26 potable water isolates were divided into
three STs (IOD, 0.151). With the exception of two strains, all
potable water isolates (92.3%) belonged to ST1. In contrast,
53.1% (51/96) and 14.3% (6/42) of cooling tower and hot spring
isolates, respectively, belonged to ST1 (χ² = 40.401, P < 0.001).
For the seven SBT loci, Nei’s index values were 0.739 to 0.847,
0.482 to 0.572, and 0.148 to 0.151 for hot spring, cooling tower,

<table>
<thead>
<tr>
<th>Year</th>
<th>City</th>
<th>Facility type</th>
<th>Water source</th>
<th>No. of strains analyzed</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>Beijing</td>
<td>Building</td>
<td>Cooling tower</td>
<td>12</td>
</tr>
<tr>
<td>2005</td>
<td>Shenzhen</td>
<td>Building</td>
<td>Cooling tower</td>
<td>20</td>
</tr>
<tr>
<td>2005</td>
<td>Beijing</td>
<td>Public bath</td>
<td>Hot spring</td>
<td>13</td>
</tr>
<tr>
<td>2008</td>
<td>Shanghai</td>
<td>Hospital</td>
<td>Cooling tower</td>
<td>26</td>
</tr>
<tr>
<td>2008</td>
<td>Shijiazhuang</td>
<td>Hospital</td>
<td>Cooling tower</td>
<td>17</td>
</tr>
<tr>
<td>2008</td>
<td>Shanghai</td>
<td>Hospital</td>
<td>Potable water</td>
<td>18</td>
</tr>
<tr>
<td>2009</td>
<td>Jinan</td>
<td>Building</td>
<td>Potable water</td>
<td>8</td>
</tr>
<tr>
<td>2009</td>
<td>Jinan</td>
<td>Building</td>
<td>Cooling tower</td>
<td>5</td>
</tr>
<tr>
<td>2010</td>
<td>WuXi</td>
<td>Hospital</td>
<td>Cooling tower</td>
<td>16</td>
</tr>
<tr>
<td>2011</td>
<td>Beijing</td>
<td>Public bath</td>
<td>Hot spring</td>
<td>14</td>
</tr>
<tr>
<td>2012</td>
<td>Shenzhen</td>
<td>Public bath</td>
<td>Hot spring</td>
<td>15</td>
</tr>
</tbody>
</table>
Clustering results of data obtained by SBT analysis of 164 L. pneumophila serogroup 1 isolates. These strains were isolated during regular surveillance in different water systems of China between 2005 and 2012. Clustering was created using the unweighted pair group method with average linkages. The ST, city, surveillance year, number of isolates, ST complex, and alleles used in the SBT analysis are shown. The asterisks indicate STs that were identified for the first time according to the EWGLI SBT database (accessed 5 September 2013).
and potable water isolates, respectively (Table 2), which revealed a high diversity of hot spring isolates and a low diversity of potable water isolates.

Population structure analyses. MST illustrates the distribution of STs. As shown in Fig. 3, 28 STs were predicted to form five clone groups (complexes), whereas 14 STs, which differed from every other ST in four or more genes, were identified as singletons. The biggest clone group was the ST1 complex, containing 91 isolates belonging to six STs. In the ST1 complex, ST1 was predicted as the putative ancestor and contained the highest number of iso-

### TABLE 2 Diversity of seven SBT loci of L. pneumophila serogroup 1 isolates from three water sources

<table>
<thead>
<tr>
<th>Locus</th>
<th>Cooling tower (n = 96)*</th>
<th>Hot spring (n = 42)</th>
<th>Potable water (n = 26)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of types</td>
<td>No. of isolates/type</td>
<td>Nei’s index</td>
</tr>
<tr>
<td>flaA</td>
<td>7</td>
<td>13.7</td>
<td>0.352</td>
</tr>
<tr>
<td>pilE</td>
<td>4</td>
<td>24.0</td>
<td>0.547</td>
</tr>
<tr>
<td>asd</td>
<td>9</td>
<td>10.7</td>
<td>0.570</td>
</tr>
<tr>
<td>mip</td>
<td>10</td>
<td>9.6</td>
<td>0.482</td>
</tr>
<tr>
<td>mompS</td>
<td>8</td>
<td>12.0</td>
<td>0.565</td>
</tr>
<tr>
<td>proA</td>
<td>9</td>
<td>10.7</td>
<td>0.572</td>
</tr>
<tr>
<td>neuA</td>
<td>7</td>
<td>13.7</td>
<td>0.552</td>
</tr>
</tbody>
</table>

*a n, number of isolates.
lates (81/91, 89.0%). The other five STs of the ST1 complex were four single-locus variants and one double-locus variant of ST1 and were identified with one to four strains each.

The second clone group was the ST154 complex (11 STs and 33 isolates), followed by the ST149 complex (5 STs and 9 isolates), the ST454 complex (4 STs and 4 isolates), and the ST1565 complex (2 STs and 2 isolates). There were differences in the distribution of clone groups based on the water source. With the exception of one cooling tower isolate, all strains of the ST149 complex were obtained from hot springs. In contrast, all ST454 and ST27 complex strains were isolated from cooling towers. However, the ST1 complex was present in all three water sources, and the ST154 complex was present in cooling towers and hot springs.

Comparison of isolates obtained from China, Japan, and South Korea. The 164 isolates obtained in this study were compared with \textit{L. pneumophila} serogroup 1 isolates from Japan and South Korea (Fig. 4). The SBT data of 135 Japanese and 104 South Korean isolates were obtained from other studies (20, 21). These 403 isolates were divided into 107 STs. The majority of the STs (100/107, 93.5%) were exclusive to one country; four STs (ST1, ST150, ST154, and ST159) were present in all three countries, and three STs (ST22, ST45, and ST59) were present in two countries (Fig. 4 and 5). In the MST dendrogram, eight clone groups (groups I to VIII) were identified: three clone groups (groups I to III) were relatively large and contained isolates from three countries, and five clone groups (groups IV to VIII) were small and contained isolates from one country.

Group I, which consisted of the ST1 complex in this study and of groups C1 (20) and CG1 (21) in previous studies, was the largest clone group. ST1 was identified as the potential founder of group I. Group II was mainly composed of three STs: ST154, ST150, and ST159. ST150 and ST159 were single-locus and double-locus variants of ST154, respectively. Group II consisted of the ST154 complex in this study and of groups C2 and CG3 in previous studies (20, 21).

Group III contained strains of five clone groups previously identified in this study (ST149 and ST454 complexes) and other studies (groups B1, B2, and CG2). These previously identified clone groups were independently distributed in group III and linked to each other. South Korean CG2 (ST39, ST363, STK1, and STK11) and Japanese group B2 (ST86, ST128, ST603, and ST605) were clustered together and linked to the Chinese ST149 complex (ST149, ST155, ST156, ST199, and ST1476), Japanese group B1, and the Chinese ST454 complex. A few singletons identified in this study and other studies, such as ST1564, STK9, and STK10, were similarly clustered in group III.

The other five clone groups mainly contained STs and strains exclusive to one country. Group IV mainly contained Chinese ST27 complex, group V mainly contained Japanese group S1, group VI mainly contained Japanese group S2, group VII mainly

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{minimum_spanning_tree.png}
\caption{Minimum spanning tree analysis of 164 \textit{L. pneumophila} serogroup 1 strains isolated in China. STs are shown as circles. The size of each circle corresponds to the number of isolates within this particular type; the STs and the number of isolates in each ST are shown in the circles. The shading simply links STs or sets of STs within an ST complex.}
\end{figure}
contained Japanese group B3, and group VIII contained Japanese group S3.

DISCUSSION

In this study, we analyzed L. pneumophila serogroup 1 isolates by SBT from three water sources in China: cooling towers, hot springs, and potable water systems. Furthermore, isolates from China were compared with isolates from Japan and South Korea. Our findings revealed that STs had several unique allelic profiles and that ST1 of L. pneumophila serogroup 1 was the most prevalent sequence type in China. Additionally, the distributions of isolate STs differed among the water sources and cities.

Even though 23 of the 42 STs obtained in this study were unique to China, the EWGLI SBT database indicated that there were single-locus variants abroad of the majority of the STs. Therefore, a few STs might be unique to China. In our previous study, we reported four Legionnaires’ disease cases attributed to ST36 and ST346 strains (22). However, no isolates belonging to these two STs were isolated in this study. In 2011, we detected a case of Legionnaires’ disease caused by an L. pneumophila serogroup 1 strain of ST59 (unpublished data) in Beijing. The patient was a 70-year-old man. An L. pneumophila serogroup 1 strain belonging to ST59 was isolated from bronchoalveolar lavage fluid (BALF) after typical symptoms of Legionnaires’ disease emerged in him. However, no ST59 strain was found upon epidemiological investigations. In this study, two ST59 strains were isolated from hot springs, which suggests that hot springs are potential sources of legionellosis.

In this study, the distribution of STs from natural water sources (e.g., hot springs) was significantly different from that from artificial water sources (e.g., cooling towers and potable water systems). This result, which is in accordance with previous studies conducted in Japan and Canada (20, 23), might be attributed to differences in the characteristics of the water sources (e.g., temperature and pH). The characteristics of hot springs are highly variable, whereas those of cooling towers and potable water systems tend to have similar characteristics due to similar water treatments. Legionella has high prevalence, rapid intracellular growth rates, and wide genetic diversity in hot springs (11). In this study, STs of isolates from hot springs, which had higher genetic diversity, were significantly different from those from cooling towers and potable water systems. These results might be due to host amoebae, which adapt to and inhabit different environments (24, 25). It has been reported that the growth of L. pneumophila in host amoebae depends on the bacterial genetic background (26, 27). Hot spring isolates, which have unique STs adapted to amoebae, may be infectious to humans.

ST1, which is the most prevalent ST worldwide (20, 21, 23, 22),
serogroup 1 isolates belonged to ST1 (31). Additionally, this study revealed that the prevalence rates of ST1 in artificial water systems were higher than those in natural water sources. In this study, ST1 accounted for 92.3% and 53.1% of the environmental serogroup 1 isolates from potable water systems and cooling towers, respectively, whereas only 14.3% of hot spring serogroup 1 isolates belonged to ST1. This result is consistent with the results of a study conducted in Japan (20), where ST1 accounted for 74% and 12% of the environmental serogroup 1 isolates from cooling towers and hot springs, respectively, whereas no ST1 strain was isolated from soil. ST1 isolates have adapted to water environments, especially to artificial water systems such as potable water systems and cooling towers, and have been isolated worldwide. The ability of ST1 isolates to adapt to natural water sources such as hot springs and soil might be rather low.

Potable water isolates were divided among three STs, and most of them belonged to ST1. Potable water ST1 isolates had a low IOD (0.151), whereas hot spring and cooling tower isolates had high IODs (0.934 and 0.711, respectively). Two other potable water isolates belonged to ST354 and ST1556: ST354 was detected in clinical isolates in Japan and Canada, and ST1556 was a novel ST identified in this study. These STs did not have the same alleles in the seven genes used for SBT typing. Further investigations of potable water isolates may identify STs that link the three STs or that indicate similar STs with isolates from water environments. ST1 and ST354 were identified in clinical isolates; therefore, potable water systems might represent an infectious source of legionellosis.

China, Japan, and South Korea belong to East Asia. Three major clone groups (groups I through III) were identified among environmental L. pneumophila serogroup 1 isolates from these three countries. Groups I and II were identified in each country. Recombination events that shortened the distance between the groups I and II of the MST may have occurred (Fig. 4). ST150 (11,14,16,1,15,13,1), ST161 (11,4,3,1,1,1,1), and STK14 (11,14,3,1,15,13,1) are recombinants between ST1 (1,4,3,1,1,1,1) and ST154 (11,14,16,16,15,13,2), which are predicted primary founders of group I and group II, respectively. Group III consists of several clone groups previously identified in each country. Recombination STs between different clone groups linked them together and formed group III in the MST. Another five clone groups mainly contained STs and strains exclusive to one country. These results suggested that these groups diverged in the recent past and that if more isolates had been included in the analysis, all of the STs might be contained in one clone group.

In conclusion, there is a complex population structure and bias distribution of genotypes of environmental L. pneumophila serogroup 1 isolates in China. The comparative analysis among strains isolated from China, Japan, and South Korea revealed that similar clones exist in different countries and that certain clones colonized different regions. The findings of this study highlight the importance of understanding the epidemiology and ecology of L. pneumophila from public facilities in terms of public health. As a result of the potential for water sources to harbor and disseminate Legionella, improved control and prevention strategies are urgently needed.

ACKNOWLEDGMENTS

We thank Norman K. Fry (Respiratory and Systemic Infection Laboratory, Health Protection Agency Centre for Infections, London, United Kingdom) for assigning the newly identified alleles and sequence types. This study was supported by grants from the National Natural Science Foundation of China (grant number 81271251), the Science Foundation for the State Key Laboratory for Infectious Disease Prevention and Control of China (grant number 2011SKLID202), and the Priority Project on Infectious Disease Control and Prevention (grants 2012ZX10004215, 2013ZX10004221, and 2013ZX10004610) of the Ministry of Health, China.

REFERENCES


