Infection, Genetics and Evolution 29 (2015) 35–41

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Sequence types diversity of *Legionella pneumophila* isolates from environmental water sources in Guangzhou and Jiangmen, China

Jingyu Guo [a,b], Ting Liang [a,c], Chaohui Hu [b], Ruichen Lv [a], Xianwei Yang [a], Yujun Cui [a], Youtao Song [c], Ruifu Yang [a], Qingyi Zhu [b,*], Yajun Song [a,*]

[a] State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Beijing, China
[b] Guanzhou Kingmed Centers for Clinical Laboratory, Guangzhou, China
[c] Department of Microbiology, Liaoning University, Shenyang, China

ARTICLE INFO

Article history:
Received 8 July 2014
Received in revised form 17 October 2014
Accepted 25 October 2014
Available online 31 October 2014

Keywords:
Legionella pneumophila
Sequence-based typing
Recombination

ABSTRACT

In this study, 159 *Legionella pneumophila* strains isolated from various natural and artificial water sources in Guangzhou and Jiangmen, China, were subjected to genotyping by the sequence-based typing (SBT) scheme. These isolates were assigned into 53 sequence types (STs) (50 STs with seven loci data and three unidentified STs with incomplete loci profiles) with ST1 as the dominant one (14.5%), and the index of diversity (IOD) was 0.950. Eight new alleles and 34 new STs were reported here. Notably, most of the newly identified STs with seven loci data (24/34) contained no new allele, implying frequent recombination events in *L. pneumophila*. Five intragenic recombination events were identified in the concatenated sequences of seven loci. The diversity of STs in natural environmental isolates (41 STs, IOD = 0.956) is higher than that of artificial environmental ones (17 STs, IOD = 0.824). The ST patterns varied in isolates from these two sources: the most common STs from artificial water sources, ST1 and ST752 (39.2% and 13.7%), were only occasionally isolated from natural water sources (2.9% and 3.8%, respectively); while the predominant STs from natural water sources, ST1048, ST739 and ST1267 (15.2%, 6.7% and 6.7%), were less frequently seen in artificial environments (2.0%, 0% and 0%, respectively). We also found out that Legionnaires' disease associated STs might be more frequently isolated in artificial environments than in natural ones. Our data revealed remarkable genetic diversity of *L. pneumophila* isolates from environmental water systems of Guangzhou and Jiangmen, and the different ST distribution patterns between natural water and artificial water sources as well.

1. Introduction

Legionella pneumophila is the causative agent of Legionnaires' disease, an atypical pneumonia (Fraser et al., 1977), and Pontiac fever, a self-limited flu-like illness (Glick et al., 1978). Although there are 57 valid species in the *Legionella* genus already, *L. pneumophila* causes approximately 90% of all reported Legionnaires' disease cases, 84% of which are due to *L. pneumophila* serogroup 1 (Fields et al., 2002; Yu et al., 2002). As reviewed in a recent paper (Luck et al., 2013), several molecular typing methods have been applied to molecular epidemiological investigations of *L. pneumophila*, such as pulsed-field gel electrophoresis (PFGE), amplified fragment length polymorphism (AFLP) and sequence-based typing (SBT). The SBT scheme, proposed by the European Working Group for *Legionella* Infections (EWGLI, now the ESCMID Study Group for *Legionella* Infections, ESGLI), is performed by sequencing and comparing seven loci (*flaA*, *pilE*, *asd*, *mip*, *mompS*, *proA* and *neuA*) (Gaia et al., 2005; Ratzow et al., 2007). The results are reported as the combination of alleles (seven-digit allelic profiles) and serial numbered sequence types (STs). With the advantages of data robustness and ease for inter-laboratory comparisons, SBT is now globally accepted and widely used in genotyping of *L. pneumophila*.

Since human-to-human transmission has not been reported for Legionnaires' disease (LD) yet, human LD is termed as an environmental disease, which is always transmitted via the inhalation of *Legionella*-containing aerosols (Blatt et al., 1993). *L. pneumophila* is ubiquitous in aquatic environments, including cooling towers, water towers, ponds, lakes and rivers. Theoretically, these aquatic environments could all serve as potential sources of *Legionella* infection. Previous studies reported that almost all cases of Legionnaires' disease could be traced to artificial water environments instead of natural aquatic environments (Coetzee et al., 2012).

* Corresponding authors.
E-mail addresses: zqy@kingmed.com.cn (Q. Zhu), songyajun88@gmail.com (Y. Song).

http://dx.doi.org/10.1016/j.meegid.2014.10.023
1567-1348/

To determine whether there are different STs distributions patterns of *L. pneumophila* isolates between artificial environments and natural environments, we performed SBT assays for *L. pneumophila* strains isolated from two water sources from Guangzhou and Jiangmen in Guangdong Province, China.

2. Materials and methods

2.1. Sampling and processing of water samples

During Oct 2003 to Dec 2008, 249 water samples were collected from natural (163 samples) and artificial water environments (86 samples) in Guangzhou City and Jiangmen City, Guangdong Province, China. Samples from lakes, rivers, harbors, ponds, and springs were termed as natural water sources, while those from cooling towers and water towers as artificial ones. Isolation of *Legionella*-like bacteria were performed as previously described (Zhan et al., 2014). Briefly, 500 ml water was concentrated by filtration through MicroFunnel 4800 with 0.45 μm Metricel membrane filters (Pall Inc., USA). Filter membranes were aseptically transferred into 5 ml of sterile water and vortexed for 5 min to release the microorganisms.

2.2. Isolation of L. pneumophila

Fifty microliters of the processed samples were plated onto GVPC plate (Oxoid Inc., USA) after heat treatments (50 °C for 30 min) and acid treatments (0.2 mol/L HCl–KCl, pH 2.0 for 5 min) to eliminate non-*Legionella* organisms. Plates were incubated at 37 °C enriched with 5% CO_2 (v/v) for 3–7 days. Colonies exhibiting *Legionella*-like morphology were transferred to BCYEα agar (with L-cysteine), BCYEα-cys$^-$ agar (without L-cysteine), and Columbia blood agar (Oxoid Inc., USA). Colonies able to grow on BCYEα but not on BCYEα-cys$^-$ agar and Columbia blood agar plates were possibly *Legionella* species, which were then identified to the species level by sequencing partial 16S rRNA gene (Wilson et al., 2007) and *mip* gene as previously reported (Ratcliff et al., 1998).

2.3. DNA extraction from L. pneumophila isolates

The identified *L. pneumophila* isolates were grown on BCYEα agar plates at 37 °C for two days, and then the bacteria cultures were harvested from the plates. Genomic DNAs were extracted by using the conventional SDS lysis and phenol–chloroform method. DNAs were dissolved in TE buffer (10 mmol/L Tris/HCl, 1 mmol/L EDTA, pH 8.0) and stored at −20 °C after concentration measurements.

2.4. Sequence-based typing

The ST of each *L. pneumophila* isolate was determined by using the standard protocol from ESGLI with seven gene fragments (*flaA*, *pilE*, *asd*, *mip*, *mompS*, *proA* and *neuA*) (Gaia et al., 2005; Ratzow et al., 2007). For the strains that failed to get *neuA* amplification, primers targeting *neuAh* were used as suggested by ESGLI (Farhat et al., 2011). The PCR products were sent to BGI Beijing Inc. for purification and sequencing. The sequence trace files were submitted to the *L. pneumophila* SBT website to get the trimmed sequences, and then queried against existing alleles in the SBT database by using the "Sequence Quality Tool" module (http://www.hpa-bioinformatics.org.uk/legionella/legionella_sbt/php/sbt_homepage.php). Strains with known allele profiles were assigned with the corresponding ST in the database, while those with new alleles or allele profiles were assigned as new STs by the database curator.

2.5. Population genetics analysis

Hunter and Gaston's modification of Simpson's index of diversity (IOD) was calculated by using the allele profiles of all the isolates to evaluate the diversity of STs (Hunter and Gaston, 1988). The proportion of each ST was compared between the natural and artificial water isolates using Fisher exact test (PASW Statistics 17.0, SPSS Inc., USA). A maximum likelihood phylogenetic tree was obtained by using the concatenated sequences of seven loci (2501/2498 bp) and 1000 bootstrap replicates using MEGA5 package (Tamura et al., 2011). The population structure of the strain collections was reflected by a minimum spanning tree (MSTree) based on allele profiles by using Bionumerics 6.6 (Applied Maths Inc., Belgium). To compare the diversity of concatenated SBT sequences between the isolates from natural and artificial water sources, hierarchical AMOVA (Analyses of Molecular Variance) was performed with program Arlequin 3.5 (Excoffier et al., 2005).

2.6. Recombination analysis

The standardized index of association (I_A^S) was calculated with START2 software (Jolley et al., 2001) to evaluate the linkage disequilibrium by using allele profiles of the isolates. The intragenic recombination was determined with the Sawyer's test implemented in START2 (Jolley et al., 2001) by using individual single locus sequences. The intergenic recombination was screened by RDP4 software (Martin et al., 2010) using concatenated sequences of seven loci (the highest *P*-value was set at $P \leqslant 0.01$). To visualize possible recombination events in the population, a reticulate network tree was prepared by using the Neighbor-net algorithm of SplitsTree4 software (Huson and Bryant, 2006). ClonalFrame (Didelot and Falush, 2007) was used to estimate two key parameters: the ratio of recombination and substitution rates (ρ/θ), and the ratio of probabilities that a given site is altered through recombination or substitution (r/m).

3. Results and discussion

3.1. Isolation of L. pneumophila

A total of 159 *L. pneumophila* strains were isolated from environmental water sources from Guangzhou City (112, 70.4%) and Jiangmen City (47, 29.6%) from Oct 2003 to Dec 2008 (Table S1). Amongst the 159 *L. pneumophila* isolates, 106 were isolated from natural water sources, including lakes, rivers, harbors, ponds, and springs (163 water samples). The other 53 *L. pneumophila* strains were isolated from artificial water sources, including cooling towers and water towers (86 water samples). The isolation rates of *L. pneumophila* from natural and artificial water sources were similar (65.0% vs. 61.6%), which was close to that of cooling tower waters in Shanghai (58.9%, 189/321) and hot spring waters in Beijing (Lin et al., 2009; Qin et al., 2013). These high isolation rates of *Legionella* from environmental water sources in China deserved more investigations.

3.2. L. pneumophila sequence-based typing

Of these 159 isolates, 156 yielded complete SBT data (all seven loci) and were identified as 50 STs, while the remaining three isolates with only six allele profiles were designated as unidentified STs (named as STGDXX) and excluded for further analysis (Table S1). For the three strains with incomplete loci data, two failed *flaA* amplification and one failed *proA* gene, which might be caused by genetic variations within the primer-binding regions of the corresponding isolates.