Original Article

Re-characterization of an extrachromosomal circular plasmid in the pathogenic
Leptospira interrogans serovar Lai strain 56601

Lili Huang1†, Weinan Zhu1†, Ping He1, Yan Zhang1, Xuran Zhuang1, Guoping Zhao2, Xiaokui Guo1, Jinhong Qin1*, and Yongzhang Zhu1*

1Department of Microbiology and Immunology, Institute of Medical Science, Shanghai Jiao Tong University School of Medicine, 280 South Chongqing Road, Shanghai 200025, China
2CAS-Key Laboratory of Synthetic Biology, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200032, China
†These authors contributed equally to this work.
*Correspondence address. Tel: +86-21-63846590-778017; Fax: +86-21-64453285; E-mail: yzhzhu@hotmail.com (Y.Z.); jinhongqin@sjtu.edu.cn (J.Q.)

In China, Leptospira interrogans serovar Lai strain 56601 (str.56601) is one of main pathogenic strains that cause severe leptospirosis in both human and animals. The genome of this organism was completely sequenced in 2003. However, in 2011, we identified and corrected some assembly errors in the str.56601 genome due to the repeat sequences widely distributed in the Leptospira genome. In this study, we re-analyzed the previously reported mobile, phage-related genomic island in the chromosome and rectified detailed sequence information in both the plasmid and chromosome using various experimental methods. The presence of a separate circular extrachromosomal plasmid was also confirmed, and its location in the genomic region was determined relative to the genomic island reported in L. interrogans serovar Lai by a combination of pulsed-field gel electrophoresis-based and plasmid extraction-based Southern blot analysis. This report confirmed that the separate extrachromosomal circular plasmid is not integrated into the chromosome of L. interrogans str.56601 and markedly improved our understanding of the genomic organization, evolution, and pathogenesis of L. interrogans. In particular, characterization of this extrachromosomal circular plasmid will contribute to the development of genetic manipulation systems in pathogenic Leptospira species.

Keywords genomic island; Leptospira interrogans; extrachromosomal circular plasmid

Received: December 24, 2013 Accepted: March 18, 2014

Introduction

Leptospira, which forms thin coils with a hook at one or both ends, belongs to the phylum Spirochetes [1,2]. When compared with other Gram-negative and Gram-positive bacteria, Leptospira has unique biochemical, genomic, morphological, and structural characteristics. Leptospirosis can be caused by infection with a variety of pathogenic species within the genus Leptospira. Pathogenic Leptospira has been recognized as the major agent of globally important emerging and zoonotic infectious diseases in tropical and subtropical areas [3]. Human beings and a wide variety of animals can be infected through either direct contact with the urine of infected animals or indirect contact with water contaminated with such urine [2,4].

The rapid development of next generation sequencing technology has made this a routine and affordable approach for nearly all fields of life science, including medical microbiology. These techniques are particularly useful for the study of Leptospira, as tools for genetic manipulation tools are still lacking [5–7]. Increasing number of bacterial genome sequences have been released in the NCBI database. However, sequencing and assembly errors were common problems for early sequences and even for newly completed genomes [8,9]. The genomic sequence of Leptospira interrogans serovar Lai strain 56601 (str.56601) within the genus Leptospira was first finished in 2003 using classical shotgun genome sequencing. Subsequent transcriptomic and proteomic data from L. interrogans has further deepened our understanding of this strain [9–13] and facilitated the study of pathogenesis in Leptospirosis. However, some assembly errors that were caused by an abundance of repeat sequences that globally distributed in the genome have been found and finally corrected in str.56601 [9,14].

Previous studies have shown that common mobile elements in bacteria such as plasmids or prophages were not found in L. interrogans [8,15]. Comparative genomic hybridization (CGH) analysis in L. interrogans serovar Lai str.56601 and 10 other L. interrogans strains isolated from China revealed an ~83 kb genomic island (LaiGIB) [16],
the location of which is similar to the 54 kb genomic island (LaiGI) identified in \textit{L. interrogans} serovar Lai \textit{str.} 56601. However, the location was not completely consistent between these two genomic islands [14]. Interestingly, LaiGI, previously considered to be a genomic element, can either integrate into or excise from the chromosome and is absent in \textit{L. interrogans} serovar Copenhageni Fiocruz L1-130 [17]. Both of these islands were thought to be bacteriophage-related.

In the present study, we further analyzed the genomic region and re-characterized a 64 kb plasmid outside of the chromosome and an empty chromosome site. The corresponding site in the chromosome was also corrected.

**Materials and Methods**

**Bacteria**

\textit{Leptospira interrogans} serovar Lai \textit{str.} 56601 was obtained from the Institute for Infectious Disease Control and Prevention (Beijing, China). Bacterial virulence was maintained by serial passage in golden hamster. The strain was grown in liquid Ellinghausen–McCullough–Johnson–Harris medium under aerobic conditions at 28°C and collected in mid-exponential phase for the subsequent study.

**DNA manipulation**

Genomic DNA from \textit{L. interrogans} \textit{str.} 56601 was isolated using the phenol-chloroform method. Polymerase chain reaction (PCR) was performed using one cycle at 95°C for 5 min, followed by 35 cycles at 94°C for 30 s, 55°C for 30 s, and 72°C for 1–5 min. Then, a final cycle of 72°C for 10 min was performed. PCR polymerase from TaKaRa (Dalian, China) and TOYOBO (Tokyo, Japan) and restriction endonucleases from New England Biolabs (Massachusetts, USA) and TIANGEN (Shanghai, China) were used in this study. PCR fragments were purified with a QIAquick gel extraction kit (Qiagen, Hilden, Germany).

For pulsed-field gel electrophoresis (PFGE), 2 x 10^9 cells or extracted DNA were first embedded in 1% agarose plugs (SeaKem Gold agarose; Cambrex, Rockland, USA), digested with lysozyme overnight, and lysed with 0.1 mg/ml protease K for 2 h. DNA fragments were separated using a 1% pulse-field agarose gel, which was prepared and run in 0.5% tris-borate-EDTA buffer on a contour-clamped homogeneous electric field machine (CHEF-DR® III; Bio-Rad, Richmond, USA). DNA fragments were separated with ramping from 5 to 65 s for 20 h at 6 V/cm.

A classical alkaline lysis approach was used to concentrate large, circular, single-copy plasmids from \textit{L. interrogans} \textit{str.} 56601 [18]. First, \textit{Leptospira} cells were resuspended in solution 1 containing 2.5 mg/ml of lysozyme at 37°C for 1 h. Solution 2 containing 0.2 M NaOH-1% sodium dodecyl sulfate was added and incubated at 55°C for 30 min. Finally, neutralization solution was added, mixed thoroughly, and cooled on ice for 5 min. After centrifugation, the supernatant was collected and extracted with acid phenol-chloroform and neutral phenol-chloroform. Finally, DNA was precipitated with isopropanol and washed twice with 70% ethanol. The purified DNA was dissolved in sterile deionized water and stored at −20°C.

For Southern blot analysis, DNA separated with an agarose gel was transferred onto nylon membranes (NC). Digoxigenin-labeled probes were generated with PCR amplification using the oligonucleotide primers listed in Table 1 (PCR DIG Probe Synthesis Kit; Roche, Basel, Switzerland). Hybridization and detection were performed with a DIG DNA detection system (Roche) according to the manufacturer’s standard protocol.

**Results**

**Genetic organization of the Laicp plasmid in \textit{L. interrogans} Lai \textit{str.} 56601**

According to the previous CGH analysis, the genomic island (LaiGIB) which is 83 kb in length and is located in LA1747–LA1851, was determined to be a unique region in \textit{L. interrogans} serovar Lai \textit{str.} 56601 [16], while the genomic island

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence of primer (5′→3′)</th>
<th>Location (bp)a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probe1-F</td>
<td>GGAAGGTTATCTATCTACGGGA</td>
<td>1,812,770–1,812,789</td>
</tr>
<tr>
<td>Probe1-R</td>
<td>CTGTACAGTTAACCCCACT</td>
<td>1,813,229–1,813,247</td>
</tr>
<tr>
<td>Probe2-F</td>
<td>TTGTAGTAGCAGAACACGACGGGA</td>
<td>1,758,019–1,758,042</td>
</tr>
<tr>
<td>Probe2-R</td>
<td>AAATAACGACTGTTCGGGTGCTGA</td>
<td>1,758,488–1,758,512</td>
</tr>
<tr>
<td>Probe3-F</td>
<td>CTTCTGAAGCCAAGGACATG</td>
<td>1,816,087–1,816,106</td>
</tr>
<tr>
<td>Probe3-R</td>
<td>CCATGGCTCATCAACTGTG</td>
<td>1,816,592–1,816,611</td>
</tr>
<tr>
<td>flaB-probe-F</td>
<td>GACGATGAAAGCTCTGTCT</td>
<td>1,995,287–1,995,305</td>
</tr>
<tr>
<td>flaB-probe-R</td>
<td>CATTGCCGTACCCTGTG</td>
<td>1,995,989–1,996,007</td>
</tr>
</tbody>
</table>

aPrimer locations on the chromosome of \textit{L. interrogans} serovar Lai \textit{str.} 56601.

---

LaiGI located at LA1768-LA1847 was shown to integrate into and excise from the chromosome [14]. To characterize the genetic locus and determine whether or not a stable circular plasmid was present in *L. interrogans* serovar Lai str.56601, we first re-analyzed the region with a series of PCR reactions followed by DNA sequencing of the PCR products with primers listed in Table 2. As a result, a 64,215 bp region designated Laicp was identified as a circular replicon outside of the chromosome at nucleotide position 1,751,019–1,815,512. This location was clearly shifted relative to LaiGIB which was located at 1,739,130–1,821,280, and LaiGI which was located at 1,764,549–1,818,007. LaiGIB spanned from LA1747 to LA1851, and Laicp encoded 61 predicted coding sequences named LA1761a–LA1841. However, genes located in LA1768–LA1847 were consistent with the genetic organization of LaiGI (Fig. 1). PCR reactions using primers in two ends of the repeat sequence (i.e. Pa1 and Pa2, Pa3 and Pa4, Pa2 and Pa5, and Pa3 and Pa6) revealed that the left and right ends of Laicp could join together due to a 12,162 bp large-repeat sequence present in both Laicp and 56601 chromosome CI (from 1,751,019 to 1,763,459). The large-repeat sequences in both Laicp and 56601 chromosome CI were identical and included the open boxes 1 and 2 shown in Fig. 1, but not the 279 bp direct-repeat open box 3 located at nucleotide positions 1,751,893–1,752,171 in the genome of *L. interrogans* str.56601 [8]. These missing sequences may be due to an assembly error in the genomic sequences. In addition, the end of the repeat sequence in chromosome CI also contained a 208 bp sequence that was missing in the previous genome sequences.

To confirm that the Laicp replicon was present as a plasmid and exclude the possibility of plasmid integration, PCR strategies employing primer pairs targeting the repeat sequences were performed and the PCR products characterized by DNA sequencing (Table 2 and Fig. 1). PCR products were successfully obtained with primer pairs Pa1:Pa6 and Pa4:Pa5, either from Laicp or the chromosomes, while no PCR products were detected with primer pairs Pa1:Pa4 and Pa5:Pa6, which contradicted the hypothesis that Laicp could integrate into the chromosomes. These results eliminated the possibility that Laicp was integrated into the chromosome.

### Molecular determination that Laicp replicates autonomously

To determine whether Laicp can also integrate into the chromosome, genomic DNA from *L. interrogans* serovar Lai str.56601 was digested with three rare-cutting restriction enzymes, *NotI*, *SgrAI*, and *BbvCI*, each of which had a single restriction site in Laicp. Digested DNA was then subjected to in situ PFGE-based Southern blot analysis. Southern-blot with DNA probe 1 located in the gene LA1839 (Table 1) revealed DNA bands of ≏65 kb, consistent with the size predicted in silico for Laicp (Fig. 2). Surprisingly, additional bands were obtained after amplification of gene LA1765 with DNA probe 2 and digestion with each of the three restriction enzymes (Table 1). These new bands revealed the presence of the repeat sequence in both the Laicp plasmid and the chromosome CI. In addition, the absence of similar bands at the 65 kb location amplified with DNA probe 3 (Table 1), which shares sequences with LaiGI but not Laicp, further

### Table 2. Oligonucleotide primers used for Laicp region verification in *L. interrogans* serovar Lai str.56601

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence of primer (5' → 3')</th>
<th>Location (bp)a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pa.1</td>
<td>GTTCCATGGGTATTGATGACT</td>
<td>1,750,487–1,750,507</td>
</tr>
<tr>
<td>Pa.1.2</td>
<td>CTGCATTCCTTGTACCAGT</td>
<td>1,749,936–1,749,954</td>
</tr>
<tr>
<td>Pa.1.3</td>
<td>AGCATTGTGTAAGGAGAGA</td>
<td>1,748,860–1,748,878</td>
</tr>
<tr>
<td>Pa.2.1</td>
<td>CTACTACATAACGTGCTCAAAGT</td>
<td>1,752,824–1,752,844</td>
</tr>
<tr>
<td>Pa.2.2</td>
<td>GTATCCATGGAAAGCTCTGA</td>
<td>1,754,405–1,754,423</td>
</tr>
<tr>
<td>Pa.3.1</td>
<td>CTTGCAGGATATTGTACCA</td>
<td>1,762,307–1,762,326</td>
</tr>
<tr>
<td>Pa.3.2</td>
<td>GTTTCCACTCAGTGCAAAAGAC</td>
<td>1,761,193–1,761,214</td>
</tr>
<tr>
<td>Pa.4.1</td>
<td>GTTTACGGGCTCTTTTTCTACT</td>
<td>1,764,625–1,764,646</td>
</tr>
<tr>
<td>Pa.4.2</td>
<td>TTATCTCTACAGAAAGCTAGC</td>
<td>1,764,780–1,764,800</td>
</tr>
<tr>
<td>Pa.4.3</td>
<td>TCAGTAAGATTTACTGTACG</td>
<td>1,766,426–1,766,446</td>
</tr>
<tr>
<td>Pa.5.1</td>
<td>CGTTACGGTCACTCGGA</td>
<td>1,815,066–1,815,087</td>
</tr>
<tr>
<td>Pa.5.2</td>
<td>CAAGAATCTCATCATATCGGT</td>
<td>1,813,611–1,813,630</td>
</tr>
<tr>
<td>Pa.5.3</td>
<td>GTACCCGAAAATATTCGCT</td>
<td>1,812,724–1,812,743</td>
</tr>
<tr>
<td>Pa.6.1</td>
<td>CATACGATAAACCTGTATAGT</td>
<td>1,816,251–1,816,272</td>
</tr>
<tr>
<td>Pa.6.2</td>
<td>GGTTCCTTTAAAGGAACCA</td>
<td>1,816,680–1,816,699</td>
</tr>
<tr>
<td>Pa.6.3</td>
<td>GACCAGATGATAGTGTACGT</td>
<td>1,818,555–1,818,575</td>
</tr>
<tr>
<td>Pa.6.4</td>
<td>TTICAATTTATCATATACTGTCC</td>
<td>1,820,599–1,820,620</td>
</tr>
</tbody>
</table>

*aPrimer locations on the chromosome of *L. interrogans* serovar Lai str.56601.*
verified the modification of the genetic organization of Laicp and the corresponding region in chromosome CI. The DNA bands using probe 3 were identical to those that appeared with probe 2, suggesting that the DNA probe sequences in the chromosome CI were close together, confirming our results above. All of the bands revealed by ethidium-bromide staining were consistent with the in silico-predicted size of fragments from whole genomic DNA with restriction enzyme digestion, which further confirmed the correct organization of plasmid Laicp and the chromosome (Fig. 2). When the whole genome was analyzed using PFGE without digestion, no linear forms of the replicon were detected, and the small quantity of DNA bands visible likely represented the remnants of degraded genomic DNA. Finally, DNA bands showed that
the Southern blot results were correct when compared with the in silico-predicted results for the revised L. interrogans serovar Lai str.56601 genome.

To further confirm Laicp as a self-replicating plasmid, plasmid DNA of L. interrogans serovar Lai str.56601 was extracted using the classical alkaline lysis method followed by PFGE-based Southern blot analysis with DNA probe 1 and the flaB probe, which is specific for a gene from chromosome CI (Table 1). As shown in Fig. 3, we observed a linear configuration for the plasmid in PFGE after digestion of total extracts with NotI, an enzyme with a unique restriction site in Laicp, and no band was detected without digestion. Southern blot results using DNA probe 1 produced the same band size as in situ whole genome digestion, and no DNA remnants were observed. In addition, bands were absent from the plasmid extraction when using the genome probe flaB (Fig. 3) and DNA probe 3. However, these bands were present when using DNA probe 2 (Fig. 4). Taken together, these results suggested that Laicp is an extrachromosomal plasmid and confirmed the genome organization of Laicp in L. interrogans serovar Lai str.56601.

Nucleotide sequence accession number
The complete nucleotide sequence of Laicp has been deposited in GenBank with the following number: KJ586855.

Discussion

Extrachromosomal genetic elements, such as plasmids or phage-related elements, may play an important role in host–bacterial interactions by contributing extra genetic components [19–21]. In the present study, a 64-kb plasmid independent of the chromosome was first confirmed in L. interrogans serovar Lai str.56601 via a combination of sequence-directed PCR, plasmid extraction with classical alkaline lysis, and in situ PFGE-based Southern blot during genome sequence data rectification. Some sequence fragments were overlooked or were redundant in the plasmid region from the previously completed genome of str.56601. These discrepancies may be due to some errors in the assembly of genomic sequences or genetic differences between the two L. interrogans serovar Lai str.56601 strains derived from different sources that arose during in vivo and in vitro passage.

Unlike the genomic island LaiGI, the Laicp plasmid in L. interrogans serovar Lai str.56601 replicated independently of the chromosome, because an integrated plasmid intermediate could not be identified. The PFGE-based Southern blot results clearly demonstrated the presence of a single fragment and excluded the possibility of integration. In addition, the genomic locus of the Laicp plasmid we identified...
was not consistent with the genome island LaiGI which has a clear forward shift. A 12,162 bp repeat sequence containing genes from LA1761 to LA1765 was identified in both Laicp and the chromosome and formed the primary difference between the Laicp plasmid and LaiGI. To our knowledge, long sequence repeats have rarely been observed in both plasmids and the chromosome, and the function of this duplication is not clear. One of the two breaking points for the repeat located at the \textit{Rhs} element sequence within the C-terminal of the large cytoplasmic membrane protein LA1765 functioned as a rearrangement hot spot locus, and no tRNA gene was present nearby. Furthermore, we identified mobile element transposases in the flank of the repeat in both the plasmid and the chromosome CI. This information suggests that mobile elements in combination with the \textit{Rhs} element instead of any tRNA genes which are usually used as targets for plasmids integration contributed to the large duplications [22]. \textit{Rhs} elements which usually contain several large open reading frames in \textit{Escherichia coli} are unknown and any association between plasmids and \textit{Rhs} elements has not been clear [23,24]. In conclusion, lateral gene transfer mediated by mobile elements might play an important role in \textit{Leptospira} genome plasticity.

As indicated previously, the Laicp plasmid is a phage-related replicon, although it has similar GC content to the two chromosomes. We speculated that phage-mediated integration events may have been involved in the acquisition of new genetic elements in \textit{L. interrogans} and have contributed to the bacteria-specific biological functions. We will continue to rectify the genomic sequence in this important pathogenic organism. The validation of this plasmid will open the field of \textit{L. interrogans} genomics and provide more information concerning the relationship between exogenous genetic elements and \textit{L. interrogans}.

**Funding**

This work was supported in part by the grants from the National Natural Science Foundation of China (30770820, 30970125, 31070111, 31270077).
References