



HAL
open science

A Genomic Island of the Pathogen *Leptospira interrogans* Serovar Lai Can Excise from Its Chromosome

Pascale Bourhy, Laurence Salaün, Aurélie Lajus, Claudine Médigue, Caroline Boursaux-Eude, Mathieu Picardeau

► **To cite this version:**

Pascale Bourhy, Laurence Salaün, Aurélie Lajus, Claudine Médigue, Caroline Boursaux-Eude, et al.. A Genomic Island of the Pathogen *Leptospira interrogans* Serovar Lai Can Excise from Its Chromosome. *Infection and Immunity*, 2007, 75 (2), pp.677-683. 10.1128/iai.01067-06 . cea-04586169

HAL Id: cea-04586169

<https://cea.hal.science/cea-04586169>

Submitted on 24 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Genomic Island of the Pathogen *Leptospira interrogans* Serovar Lai Can Excise from Its Chromosome^{∇†}

Pascale Bourhy,¹ Laurence Salaün,¹ Aurélie Lajus,² Claudine Médigue,²
 Caroline Boursaux-Eude,³ and Mathieu Picardeau^{1*}

Laboratoire des Spirochètes, Institut Pasteur,¹ Génoscope and CNRS-UMR8030, Atelier de Génomique Comparative,² and
 Plate-forme Intégration et Analyse Génomique, Institut Pasteur,³ Paris, France

Received 7 July 2006/Returned for modification 16 August 2006/Accepted 13 November 2006

An examination of the two *Leptospira interrogans* genomes sequenced so far reveals few genetic differences, including an extra DNA region, 54 kb in length, in *L. interrogans* serovar Lai. This locus contains 103 predicted coding sequences that are absent from the genome of *L. interrogans* serovar Copenhageni, of which only 20% had significant BLASTP hits in GenBank. By analyzing the *L. interrogans* serovar Lai genome by pulsed-field gel electrophoresis, we also found that this 54-kb DNA fragment exists as a circular plasmid. This was confirmed by amplification of a DNA fragment corresponding to that of the predicted fragment if this region excised from the chromosome and its left and right ends joined together. In addition, cloning of the putative *rep* gene of this DNA region was responsible for autonomous replication in *Leptospira* spp., therefore generating a new *Escherichia coli*-*Leptospira* sp. shuttle vector. Taken together, our results show that this genomic island can excise from the chromosome and form a replicative plasmid. Analysis of the distribution of this genomic island revealed that highly related sequences exist in other *L. interrogans* virulent strains. This genomic island, containing a high proportion of novel genes, may have an important role in spreading genes, including virulence factors, among bacterial populations.

The phylum of spirochetes has a deep branching lineage, and numerous studies have shown that they differ in a number of morphological, structural, biochemical, and genetic respects from well-studied gram-negative and gram-positive bacteria. Spirochetes are the causative agents of several important human diseases such as syphilis, Lyme disease, and leptospirosis.

Leptospirosis is a worldwide-distributed zoonosis, which is endemic in tropical areas. With the occurrence of large outbreaks in the last decade in areas including Nicaragua (23), Brazil (12), and India (26), leptospirosis is now recognized as an important emerging infectious disease. In industrialized countries, the chance of contracting leptospirosis can increase depending on recreational or occupational exposure. It is estimated that more than 500,000 cases of severe leptospirosis occur annually in the world, with a mortality rate of up to 23% (25). Transmission to humans occurs through direct or indirect contact with urine of infected animals, such as small rodents (15).

Leptospira interrogans is the most frequently reported agent of leptospirosis, with *L. interrogans* serogroup Icterohaemorrhagiae representing more than half of the leptospires encountered in human infections. Recently, the completion of the genome sequences of *L. interrogans* serovar Lai and *L. interrogans* serovar Copenhageni, both belonging to serogroup Icterohaemorrhagiae, was achieved (16, 20). The Lai and Copenhageni sequenced strains were isolated from patients with

severe leptospirosis in China (20) and Brazil (16), respectively. It should be noted that while the regular reservoir host for serovar Lai is the striped field mouse (*Apodemus agrarius*), the host for serovar Copenhageni is the domestic rat (*Rattus norvegicus*) (15). The *L. interrogans* genome consists of a 4.33-Mb large circular chromosome and a 350-kb small chromosome (16, 20). The majority of predicted coding sequences (CDSs) of the *L. interrogans* genome fail to exhibit similarities to proteins of known function in other organisms. The genome is highly conserved between the serovars Lai and Copenhageni, exhibiting 95% identity at the nucleotide level (16). Comparative genomics of these two *L. interrogans* serovars reveals few genetic differences, including a 54-kb insertion specific to serovar Lai, not found in serovar Copenhageni (Fig. 1). In this study, we show that this 54-kb DNA fragment can both excise from the Lai chromosome and form a circular plasmid.

MATERIALS AND METHODS

Bacterial strains and culture conditions. We obtained reference strains belonging to *L. interrogans*, *L. kirschneri*, *L. noguchi*, and *L. borgpetersenii* from the collection maintained by the National Reference Laboratory for *Leptospira* at the Institut Pasteur (Paris, France). All strains were cultured at 30°C in EMJH liquid medium (7, 11).

DNA manipulations. Genomic DNA of *Leptospira* was isolated using the phenol-chloroform method as previously described (17). For pulsed-field gel electrophoresis (PFGE), cells were embedded in agarose plugs as previously described (5). For digestion, DNA plugs were washed in Tris-EDTA buffer, followed by equilibration in 1× restriction enzyme buffer, and then incubated overnight at 37°C in fresh 1× restriction enzyme buffer containing 30 U NotI restriction enzyme. PFGE was performed in a contour-clamped homogeneous electric field DRII apparatus (Bio-Rad Laboratories, Richmond, CA). Programs with a ramping from 5 to 100 s for 40 h at 170 V or from 10 to 100 s for 40 h at 150 V were used for the resolution of DNA fragments. The DNA was amplified using *Taq* polymerase (Amersham Biosciences, Little Chalfont, England) under the following conditions: 2 min at 95°C, followed by 35 cycles of 10 s at 95°C, 15 s at 55°C, and 5 min at 72°C, and then one cycle of 10 min at 72°C. For Southern

* Corresponding author. Mailing address: Laboratoire des Spirochètes, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France. Phone: 33 (1) 45 68 83 68. Fax: 33 (1) 40 61 30 01. E-mail: mpicard@pasteur.fr.

† Supplemental material for this article may be found at <http://iai.asm.org/>.

∇ Published ahead of print on 21 November 2006.

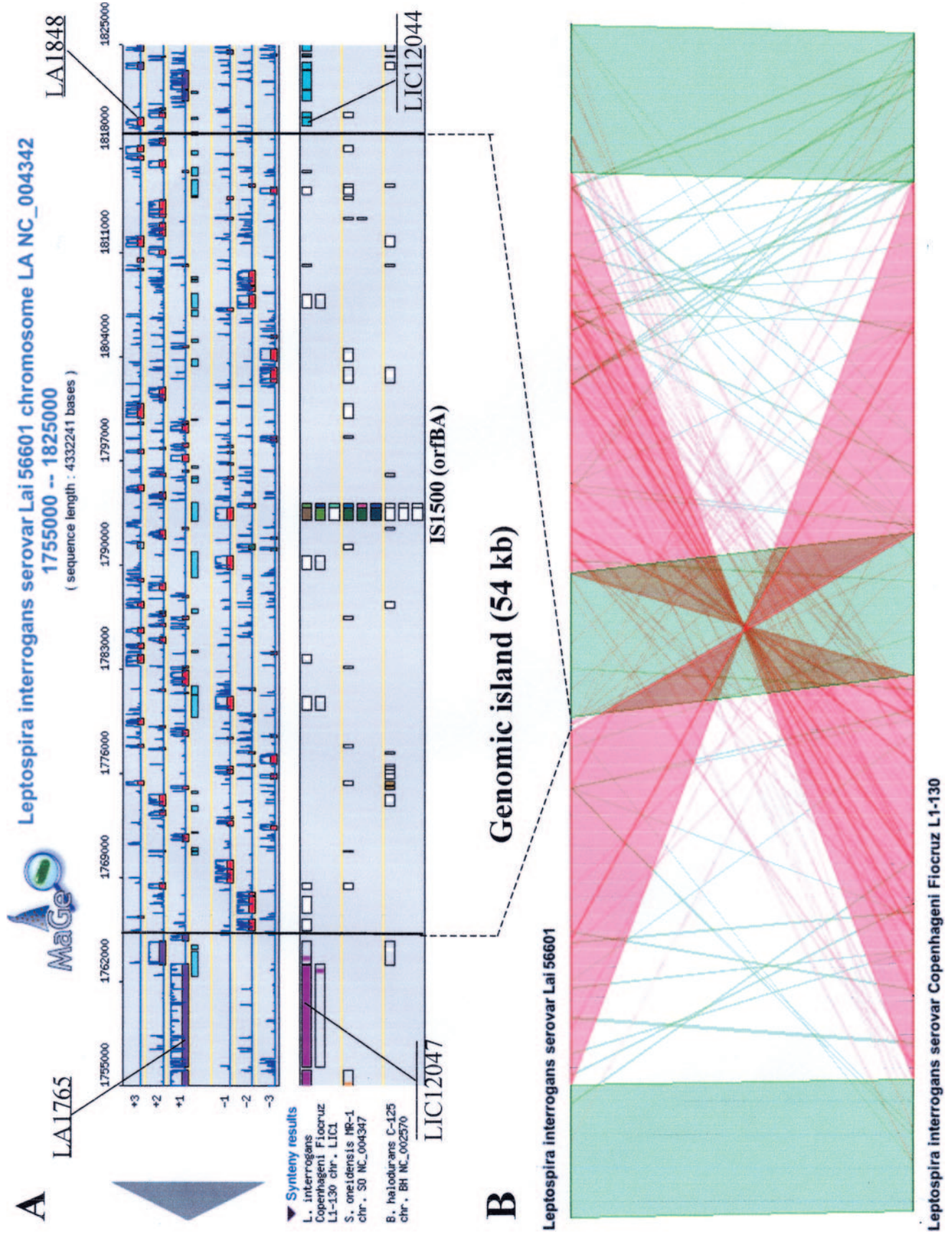


FIG. 1. Genomic island of *L. interrogans* serovar Lai in comparison with the corresponding region in *L. interrogans* serovar Copenhageni. (A) A chromosomal segment, extending between bp positions 1755000 and 1825000, is represented on this graphical map of the MaGe interface (24). Annotated CDSs are represented in the six reading frames of the sequence by red rectangles, and coding prediction curves are superimposed on the predicted CDSs (blue curves). The synteny maps, calculated on a set of selected genomes, are displayed below. A line contains the similarity results between *L. interrogans* serovar Lai and one given genome. A rectangle represents a putative ortholog between one CDS of the compared genome and one CDS of the *L. interrogans* serovar Lai genome. If, for several colocalized CDSs on the *L. interrogans* serovar Lai genome, there are several colocalized orthologs in the compared genome, the rectangles are all of the same color; otherwise, the rectangle is white. A group of rectangles of the same color thus indicates synteny between *L. interrogans* serovar Lai and the compared genome. *L. interrogans* serovar Copenhageni LIC12047 and LIC12044 proteins are similar to LA1765 (the LA1766 protein is a duplication of a 3' region of the LA1765 protein) and LA1848 proteins, respectively, in *L. interrogans* serovar Lai. Homologs of the two IS1/500 open reading frames have been found in synteny with *Shewanella oneidensis* (second line of the synteny map). The third line shows *Bacillus halodurans* synteny results in which similar integrases are found but not colocalized (white rectangles). (B) A line plot has been obtained comparing synteny results between *L. interrogans* serovar Lai and *L. interrogans* serovar Copenhageni. Putative orthologous relations between the two genomes are defined as gene couples satisfying the bidirectional best hit criterion or a BLASTP alignment threshold, a minimum of 35% sequence identity on 80% of the length of the smallest protein. These relations are subsequently used to search for conserved gene clusters, e.g., synteny groups (syntons). All possible kinds of chromosomal rearrangements are allowed (inversion, insertion, and deletion). A gap parameter, representing the maximum number of consecutive genes which are not involved in a synteny group, is set to five genes. Green, synteny groups are organized on the same strand; red, synteny groups are organized on two opposite strands. chr., chromosome.

TABLE 1. Oligonucleotide primers used in this study

Primer	Sequence of primer (5'-3')	Location (bp) ^a
L9a	TTTGATTTTCGAAAGCGATCC	1291453-1291472
L9b	CTTAACATCACTATCTATCG	1291263-1291282
P1	GTTTTCCACTGAGTGC AAAGAC	1761032-1761053
P2	TTCAGTAAAGTTACTGTTACG	1767103-1767123
P3	ACAAAGCCAAGGTGTTGATG	1817975-1817994
P4	TTCAATTTATCAATAATCTGC	1821276-1821297
P5	TTATTCTTACAAGAAAGTAGC	1765457-1765477
P6	ACCACCATGAAGTCACACCAC	1816851-1816870
P7	GATTGTTCATTAGACAAGAAACC	1817064-1817085
P8	CTGACTATTACGGTGGGAG	1764962-1764980
LA1773a	ATTTCCGTTGGGTTATGGAC	1769416-1769457
LA1773b	GCATTTAGTCTTCTCATCTC	1768812-1768831
LA1839a	TGTCGGTACCGCACTCAGGA	1813414-1813433
LA1839c	GGGAGTGGAAGTCTGAAGAG	1814879-1814898
LA1809a	GAGAAAATTACATATACTTC	1791796-1791815
LA1809b	TTCAAATGGAGTAGTGTTG	1792414-1792433
LA1850a	ATCTATTACGGCAGGGTCTC	1820711-1820730
LA1851b	CGCGATGCGGTTGACACGAG	1821894-1821913

^a The GenoList database (<http://bioinfo.hku.hk/GenoList/index.pl?database=leptolist>) was used for determining primer locations on the CI chromosome of *L. interrogans* serovar Lai.

blot analysis, genomic DNA was digested, subjected to electrophoresis in a 1% agarose gel, and transferred onto nylon membranes as previously described (17). Probes were generated by PCR and radiolabeled with [α -³²P]dATP using a commercial kit (Megaprime; Amersham Biosciences). Membranes were hybridized overnight at 55 or 60°C in rapid hybridization buffer (Amersham Biosciences) and then washed as previously described (17).

The oligonucleotides listed in Table 1 were used to amplify DNA fragments from *L. interrogans* serovar Lai.

Plasmid construction and electroporation-mediated transformation of *Leptospira* spp. A 1.5-kb fragment encompassing the LA1839 gene of *L. interrogans* serovar Lai was amplified with primers 1839a and 1839c (Table 1) and inserted into pCR2.1-TOPO by using a TOPO TA cloning kit (Invitrogen Life Technologies, Carlsbad, CA). After PvuII digestion, the DNA fragment containing LA1839 was inserted into the dephosphorylated SmaI site of the spectinomycin-resistant plasmid pGSpc, which is derived from pGEM-7Zf(+) (Promega, Madison, WI). Plasmids from *Escherichia coli* were recovered using a QIAprep spin miniprep kit (QIAGEN GmbH, Hilden, Germany). *Leptospira* cells were electrotransformed as previously described (22). Spectinomycin-resistant clones were verified by isolation of plasmid DNA and subsequent restriction fragment length polymorphism analysis.

Sequence analysis. The nucleotide sequences of the genomes of *L. interrogans* serovar Lai strain Lai 56601 (20) and *L. interrogans* serovar Copenhageni strain Fiocruz L1-130 (16) were analyzed using MaGe software (<http://www.genoscope.cns.fr/agc/mage/>) (24). PCR products were sequenced at Genome Express (Meylan, France).

Nucleotide sequence accession numbers. The accession numbers of the nucleotide sequences of the serovar Lai genomic island are EF088430 (5' flanking sequence of the chromosomal locus), DQ890383 (chromosomal empty locus), and DQ890384 and EF100899 (circular intermediates).

RESULTS AND DISCUSSION

Genetic organization of the 54-kb genomic island of *L. interrogans* serovar Lai. In bacteria, genomic islands (including those related to virulence, pathogenicity islands) represent large chromosomal regions that are present in a subset of strains of the same species. Their G+C content usually differs from that of the host chromosome. They are flanked by direct repeats and are often inserted into tRNA-encoding genes. They also encode mobility enzymes, such as integrase, allowing for excision from the host chromosome (6). Analysis of the 54-kb DNA region of *L. interrogans* serovar Lai (from LA1768 to LA1847) not found in serovar Copenhageni (Fig. 1) suggests

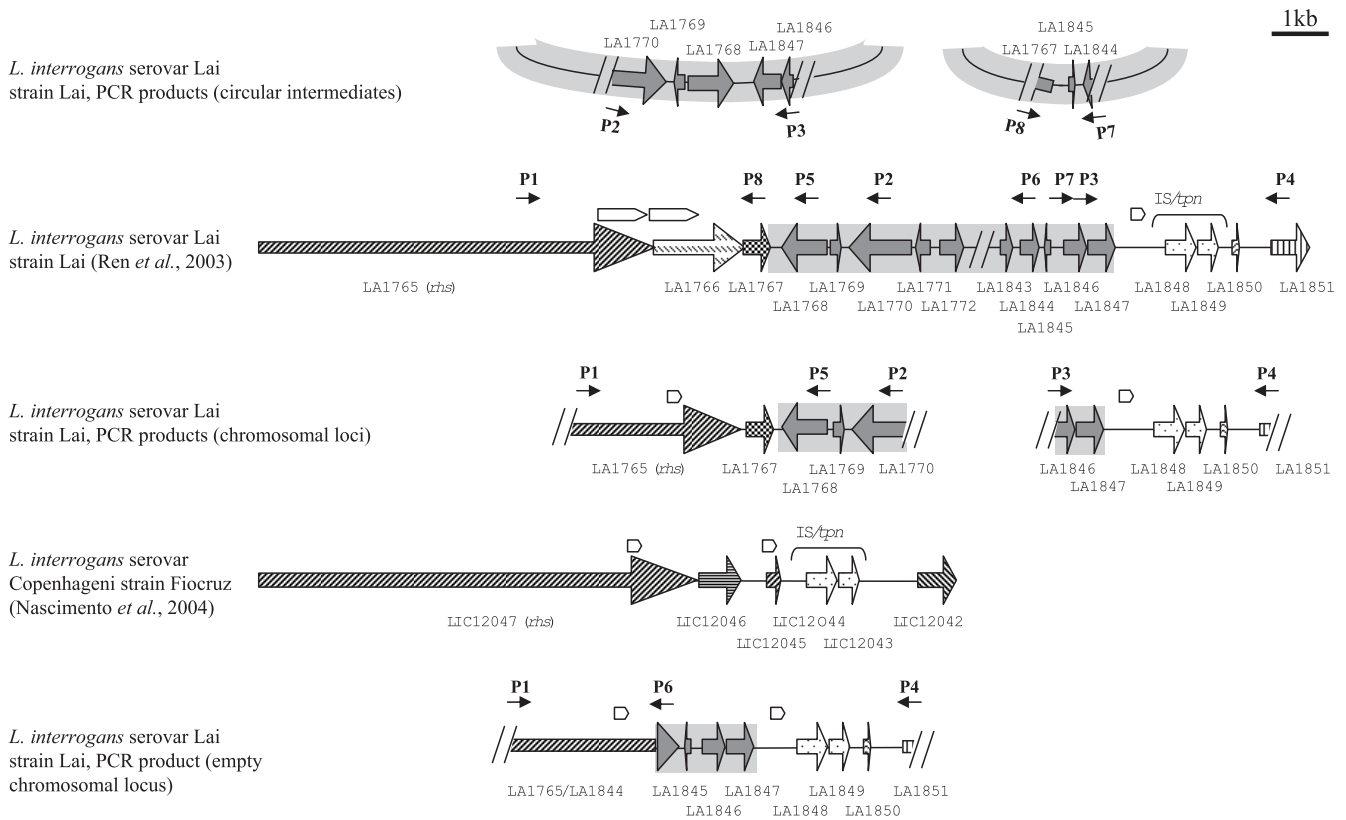


FIG. 2. Organization of the genes located in the genomic island locus of *L. interrogans* serovar Lai. Comparison of the overall organization of the genomic island of *L. interrogans* serovar Lai (including the published genome sequence and the amplified products that were sequenced) with that of the corresponding region in *L. interrogans* serovar Copenhageni. Empty boxes indicate the 812- and 170-bp direct repeats at the boundaries of the 54-kb DNA fragment. DNA fragments belonging to LaiGI I are shadowed in gray. CDSs with the same motif are homologs. The locations of primers (Table 1) used for PCR amplification are indicated. The figure is drawn to scale.

that this locus is a genomic island, further referred to as Lai genomic island I (LaiGI I).

This island is not associated with tRNA genes but is flanked by a putative *rhs* (rearrangement hot spot)-like gene and an insertion sequence (LA1848 and LA1849) belonging to the IS4 family (Fig. 1 and 2). Rhs elements belong to a set of composite elements found in the chromosome of *Escherichia coli*. These genes, the function of which remains unknown in *E. coli*, are typically long CDSs with repetitions and can be found in association with insertion sequences (10). The reason for the association between LaiGI I and the *rhs*-like gene is not clear.

The presence of LaiGI I in the large chromosome of *L. interrogans* serovar Lai was verified by PCR amplification across the putative DNA junctions (Fig. 2 and 3), followed by sequencing of the PCR fragments. Sequencing of the flanking sequences of LaiGI I indicated the absence of an 812-bp direct repeat (Fig. 2), corresponding to the 5' end of LA1766, which was previously believed to be present in the *L. interrogans* serovar Lai genome (20). This discrepancy may be due to an error in the assembly of genomic DNA sequences or to DNA rearrangements of the sequenced strain after *in vitro* passages. The 170-bp direct repeat found in *L. interrogans* serovar Copenhageni was found in the flanking regions of LaiGI I (Fig. 2). Sequence analysis of LaiGI I revealed an average G+C

content of 33%, in comparison to 36% for its chromosomal host.

By using AMIGene (Annotation of Microbial Genes) software (1), which allows the calculation of coding prediction curves, the pattern of codon usage of genes belonging to LaiGI

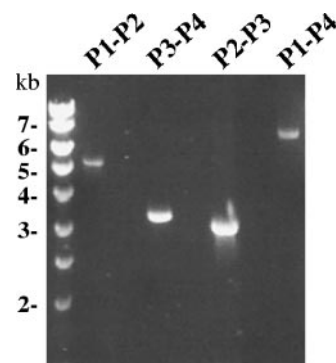


FIG. 3. PCR results showing DNA excision and integration of the 54-kb DNA fragment. PCR amplification for detection of the left end of LaiGI I into the chromosome (P1-P2), the right end of LaiGI I into the chromosome (P3-P4), the circular intermediate with junction of left and right ends of the chromosomal LaiGI I region (P2-P3), and the excision event (P1-P4) in *L. interrogans* serovar Lai.

I was not distinguishable from that of codon usage in the rest of the genome (data not shown). LaiGI I contains 103 predicted CDSs, 23 of which had significant BLASTP hits in GenBank (see the supplemental material). The first gene of LaiGI I, LA1768, encodes a putative protein which shows features related to members of the RNase H-like superfamily (Pfam 3e-11), which includes the human immunodeficiency virus type 1 integrase and the Mu transposase (21). The LA1768 protein contains the three amino acids of the DDE motif (Asp-86, Asp-152, and Glu-185) found in the catalytic domains of transposases and retroviral integrases (14). This putative protein may be required for the integration and/or excision of the 54-kb DNA fragment.

Genomic islands can excise from the chromosome and form circular intermediates (6). Interestingly, genes encoding proteins that could be involved in plasmid segregation were identified in LaiGI I. The segregation of low-copy-number plasmids in bacterial cells is an efficient process that ensures that every daughter cell receives a copy of plasmid DNA, thereby enhancing their inheritance and stability (9). LaiGI I contains putative addiction (LA1780-LA1781) and partition (LA1837-LA1838) modules (see the supplemental material). Addiction modules contain two small genes encoding a potent cell toxin and an antidote protein. In the presence of the addiction module-containing plasmid, the toxicity is neutralized by the antidote protein, which is labile. In contrast, upon loss of the plasmid, the residual stable toxin will likely kill plasmid-free cells. Other toxin-antitoxin loci were identified in the *L. interrogans* genome (18, 27). Partition modules usually consist of the *parA* and *parB* genes, whose products act in conjunction with a centromere-like locus to facilitate faithful plasmid segregation (9). The identification of homologs of partitioning and post-segregational killing genes indicates that LaiGI I could be highly stable in *L. interrogans* serovar Lai.

The origins of replication of low-copy-number plasmids are often found in the vicinity of *parAB*. For instance, the *rep* gene of the leptophage LE1 was found immediately downstream of *parAB* (2). In LaiGI I, the gene found downstream of the *parAB* genes, LA1839, encodes a putative protein (381 amino acids in length) unrelated to any other eukaryotic or prokaryotic protein. However, the C-terminal region of the LA1839 protein (from amino acids 153 to 358) shows 50% similarity to the LB376 and LIC20276 proteins, which are encoded by genes located in the putative replication origin of the small chromosome (CII) of *L. interrogans* serovar Lai and *L. interrogans* serovar Copenhageni, respectively (16, 20). A 1.5-kb fragment encompassing the LA1839 protein was amplified and then cloned into an *E. coli* plasmid to generate pORLS. After electroporation with pORLS, we observed plasmid autonomous replication in our laboratory strain *Leptospira biflexa* and in the pathogen *L. interrogans* serovar Canicola strain Hond Utrecht but not in *L. interrogans* serovar Lai and *L. interrogans* serovar Copenhageni (data not shown). However, transformation efficiency was very low (10 transformants per μg of DNA). The absence of transformants in serovars Lai and Copenhageni can therefore be due to poor transformation of the strains and not the absence of plasmid replication. The maintenance of pORLS in leptospiral cell strains suggests that LA1839 is the *rep* gene of LaiGI I. Consistent with the replication initiation protein (Rep protein), LA1839 encodes a protein containing a

predicted helix-turn-helix motif at its N-terminal end that could be involved in DNA binding activity. Finally, the origin of replication of LaiGI I coincides with the polarity switch of the GC skew of the circular plasmid (see the supplemental material). The *L. biflexa*-*E. coli* shuttle vector derived from LE1 has already proved to be useful in studies of gene function in saprophytes (2). Our results enable the construction of a new shuttle vector for the genetics of *Leptospira*, including pathogenic species.

Identification of both a circular intermediate of the genomic island and an empty chromosomal site in *L. interrogans* serovar Lai. PFGE analysis of chromosomal DNA after digestion with the rare-cutting restriction enzyme NotI revealed the presence of an additional fragment, larger than 50 kb in size, in comparison to predictions made from the whole-genome sequence of *L. interrogans* serovar Lai (Fig. 4). Southern hybridization of the blot with specific probes of LaiGI I gave a strong signal with this ≈ 50 -kb fragment. While DNA probes B and C are physically separated by a NotI restriction site when LaiGI I is inserted into the chromosome, our data indicate that these DNA probes are linked in the ≈ 50 -kb NotI-digested DNA fragment (Fig. 4). Since LaiGI I contains a single NotI restriction site, this restriction profile is consistent with the linearized form of a circular plasmid corresponding to LaiGI I. No linear replicon was detected by PFGE of undigested DNA and Southern blot hybridization (Fig. 4). Comparison of the in silico restriction map of *L. interrogans* serovar Lai (20) and PFGE-separated, NotI-digested chromosomal DNA fragments also reveals that a 22-kb fragment predicted by in silico analysis was missing in PFGE. This 22-kb restriction fragment is associated with the chromosomal LaiGI locus (Fig. 4). Consistent with the absence of this segment, the 720-kb NotI-digested DNA deduced from the published sequence should be ≈ 688 kb in size (Fig. 4). However, it remains uncertain by PFGE whether the restriction fragment is 688 kb rather than 720 kb (the restriction fragment is located between the 679- and 727-kb fragments of the molecular size standard) (Fig. 4). Evaluation of the ethidium bromide staining intensities of the chromosomal restriction fragment less than 100 kb in length (Fig. 4) also suggests that the majority of LaiGI I DNA remains as a low-copy-number circular plasmid (similar to the chromosome copy number) rather than an integrated form in *L. interrogans* serovar Lai.

The presence of DNA segregation and replication mechanisms typical of circular plasmid and visualization of an extra-chromosomal band by PFGE suggests that LaiGI I can replicate as a plasmid. A PCR strategy was therefore designed that would amplify a PCR product only if LaiGI I excised from the chromosome and formed a circular intermediate. PCR analysis with primers P2 and P3 successfully amplified a PCR product (Fig. 3). DNA sequencing of the cloned PCR product showed that the sequences were identical to that of the predicted fragment if LaiGI I excised from the chromosome and its left and right ends joined together. The occurrence of DNA junctions between distinct left and right ends of LaiGI I was also detected with primers P7 and P8 (Fig. 2), suggesting that imprecise excision of LaiGI I from the chromosome may result in distinct plasmid derivatives. By using a similar methodology, circular excision intermediates of genomic islands have been demonstrated in other organisms (8, 13, 19). We also hypoth-

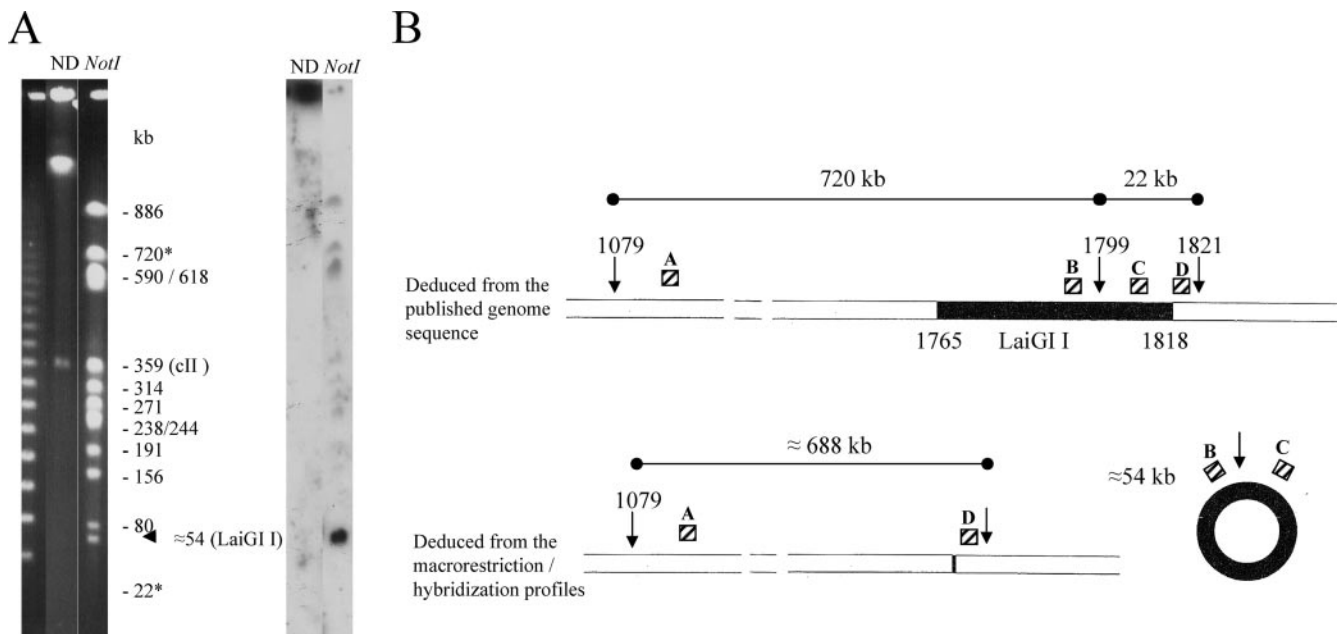


FIG. 4. Detection of the circular intermediate of LaiGI I by PFGE. (A) Whole genomic analysis of *L. interrogans* serovar Lai DNA by PFGE. ND, undigested DNA; NotI, digestion with NotI. The arrowhead indicates the additional restriction fragment when the NotI macrorestriction profile was compared with predictions made from the whole-genome sequence of *L. interrogans* serovar Lai. Asterisks indicate DNA fragments reported as absent (22-kb fragment) or possibly different in length (720-kb fragment) by PFGE compared with predictions made from the whole-genome sequence of *L. interrogans* serovar Lai (20). For Southern blot analysis, the genomic DNA of *L. interrogans* serovar Lai was blotted to a nylon membrane and then hybridized at 60°C with the radiolabeled probe B (corresponding to the *rep* gene). The size marker, on the left, is the bacteriophage λ DNA multimer (monomer = 48.5 kb). (B) Schematic representation of the macrorestriction patterns from *L. interrogans* serovar Lai. For Southern blot analysis, the genomic DNA of *L. interrogans* serovar Lai was blotted to a nylon membrane and then hybridized at 60°C with radiolabeled probes. Probes A, B, C, and D were generated by PCR of total genomic DNA with primer pairs L9a-L9b, 1809a-1809b, 1839a-1839c, and 1850a-1851b, respectively (Table 1). The probes B and C hybridized to the same fragment (located between the 48- and 97-kb fragments of the molecular size standard), indicating the linkage between these two probes. Similarly, probes A and D hybridized to a single macrorestriction fragment (located between the 679- and 727-kb fragments of the molecular size standard). The 22-kb fragment predicted from the published genome sequence (20) is reported as absent. This suggests that the majority of LaiGI I (black box) is not inserted into the chromosome but as a circular plasmid, which is linearized after NotI restriction. Arrows indicate the NotI restriction sites on the large chromosome of *L. interrogans* serovar Lai (the sizes of restriction fragments are indicated in kilobases).

esized that there should be an empty chromosomal site, and we demonstrated that to be the case by PCR (Fig. 2 and 3). However, sequencing of the PCR product suggested that imprecise excision occurs, leaving behind a remnant of LaiGI I (Fig. 2). As implied by imprecise excision, the acquisition of DNA of LaiGI I to the right end of LaiGI I is accompanied by the loss of host DNA from the left end of LaiGI I. Taken together, our data suggest that LaiGI I is preferentially in an excised state and that both normal and abnormal excisions of LaiGI I occur in *L. interrogans* serovar Lai.

LaiGI I is present in other virulent strains. We were interested in determining whether this genomic island is restricted to *L. interrogans* serovar Lai or is also found in other pathogenic members of the genus. Hybridization analysis demonstrated that sequences highly related or identical to those of this genomic island exist in other *L. interrogans* strains, i.e., serovars Australis, Bataviae, Canicola, Hebdomadis, and Pyrogenes (Table 2). Transduction, or bacteriophage-mediated gene transfer, is thought to play an important role in the dissemination of genomic islands in other organisms. However, none of the LaiGI I CDSs were found to be homologs of phage-like structural proteins. In addition, genes of LaiGI I are not clearly clustered in large transcription units of related or

TABLE 2. Distribution of the *L. interrogans* serovar Lai genomic island^a

Species	Serovar	Strain	LaiGI I
<i>L. interrogans</i>	Australis	Ballico	+
<i>L. interrogans</i>	Bataviae	Van Tienen	+
<i>L. interrogans</i>	Canicola	Hond Utrecht IV	+
<i>L. interrogans</i>	Copenhagani	Fiocruz L1-130 ^b	-
<i>L. interrogans</i>	Hardjobovis	Sponselee	-
<i>L. interrogans</i>	Hebdomadis	Hebdomadis	+
<i>L. interrogans</i>	Lai	Lai 56601 ^c	+
<i>L. interrogans</i>	Icterohaemorrhagiae	Verdun	-
<i>L. interrogans</i>	Pyrogenes	Salinem	+
<i>L. noguchi</i>	Panama	CZ 214 K	-
<i>L. kirschneri</i>	Grippotyphosa	Moskva V	-
<i>L. borgpetersenii</i>	Ballum	Castellon 3	-
<i>L. borgpetersenii</i>	Tarassovi	Mitis Johnson	-

^a Genomic DNA from a collection of strains was digested with *EcoRI*, separated on an agarose gel, transferred to a nylon membrane, and hybridized with two probes of LaiGI I (LA1773 and LA1839) at 55°C as previously described (17). +, presence; -, absence.

^b Strain sequenced by Nascimento et al. (16).

^c Strain sequenced by Ren et al. (20).

interacting proteins as in most phage genomes (see the supplemental material). It is possible that the phage genes that are essential for the viral life cycle are not fully functional, so the phage cannot be propagated. This element may also be transferred, if it can really be transferred, to a new host by another mechanism such as conjugation.

In conclusion, the 54-kb DNA region described here fulfills most of the criteria proposed for genomic islands (4): the overall G+C content is lower than that found in the rest of the *L. interrogans* chromosome, its size is >30 kb, genes encoding mobility enzymes are present, and the LaiGI I flanking sequences contain direct repeats and insertion sequences. In addition, while the genome of *L. interrogans* serovar Lai was originally described as containing two circular chromosomes (20), we show that *L. interrogans* serovar Lai also possesses a circular plasmid. It is well established that integrated phage genomes or genomic islands are responsible for conferring new traits that can allow an adaptation of the bacteria to new environment/host conditions (4). For example, LaiGI I contains genes encoding putative regulators and membrane proteins (see the supplemental material) that may play an important role in the physiology of the bacterium. Interestingly, this island contains a high proportion of novel genes. The determination of the function of these novel genes will likely provide insights into the role of this genomic island in *L. interrogans* serovar Lai.

Only recently, the first evidence of gene transfer was demonstrated in *L. interrogans* by transposition of *HimarI*, a transposon of eukaryotic origin (3). The putative integrase (encoded by LA1768) and the replication protein (encoded by LA1839) from LaiGI I can be developed as novel tools for the genetics of *Leptospira*. A better understanding of the molecular biology of this genomic island not only will allow the development of genetic tools but also may lead to improvements in our understanding of virulence and gene dissemination in *Leptospira* spp.

ACKNOWLEDGMENTS

This work was supported by the French Ministry of Research "ANR Jeunes Chercheurs" (no. 05-JCJC-0105-01).

We thank L. Frangeul for his help in the drawing of the circular map of LaiGI I and I. Saint Girons for her encouragement.

REFERENCES

1. Bocs, S., S. Cruveiller, D. Vallenet, G. Nuel, and C. Médigue. 2003. AMIGene: Annotation of Microbial Genes. *Nucleic Acids Res.* **31**:3723–3726.
2. Bourhy, P., L. Frangeul, E. Couve, P. Glaser, I. Saint Girons, and M. Picardeau. 2005. Complete nucleotide sequence of the LE1 prophage from the spirochete *Leptospira biflexa* and characterization of its replication and partition functions. *J. Bacteriol.* **187**:3931–3940.
3. Bourhy, P., H. Louvel, I. Saint Girons, and M. Picardeau. 2005. Random insertional mutagenesis of *Leptospira interrogans*, the agent of leptospirosis, using a *mariner* transposon. *J. Bacteriol.* **187**:3255–3258.
4. Canchaya, C., G. Fournous, and H. Brussow. 2004. The impact of prophages on bacterial chromosomes. *Mol. Microbiol.* **53**:9–18.
5. Davidson, B. E., J. MacDougall, and I. Saint Girons. 1992. Physical map of the linear chromosome of the bacterium *Borrelia burgdorferi* 212, a causative agent of Lyme disease, and localization of rRNA genes. *J. Bacteriol.* **174**:3766–3774.
6. Dobrindt, U., B. Hochhut, U. Hentschel, and J. Hacker. 2004. Genomic islands in pathogenic and environmental microorganisms. *Nat. Rev. Microbiol.* **2**:414–424.
7. Ellinghausen, H. C., and W. G. McCullough. 1965. Nutrition of *Leptospira pomona* and growth of 13 other serotypes: fractionation of oleic albumin complex and a medium of bovine albumin and polysorbate 80. *Am. J. Vet. Res.* **26**:45–51.
8. Fukudaa, Y., Y. Okamura, H. Takeyama, and T. Matsunaga. 2006. Dynamic analysis of a genomic island in *Magnetospirillum* sp. strain AMB-1 reveals how magnetosome synthesis developed. *FEBS Lett.* **580**:801–812.
9. Gordon, G. S., and A. Wright. 2000. DNA segregation in bacteria. *Annu. Rev. Microbiol.* **54**:681–708.
10. Hill, C. W. 1999. Large genomic sequence repetitions in bacteria: lessons from rRNA operons and Rhs elements. *Res. Microbiol.* **150**:665–674.
11. Johnson, R. C., and V. G. Harris. 1967. Differentiation of pathogenic and saprophytic leptospires. *J. Bacteriol.* **94**:27–31.
12. Ko, A. I., M. Galvao Reis, C. M. Ribeiro Dourado, W. D. J. Johnson, and L. W. Riley. 1999. Urban epidemic of severe leptospirosis in Brazil. Salvador Leptospirosis Study Group. *Lancet* **354**:820–825.
13. Lavigne, J. P., A. C. Vergunst, G. Bourg, and D. O'Callaghan. 2005. The IncP island in the genome of *Brucella suis* 1330 was acquired by site-specific integration. *Infect. Immun.* **73**:7779–7783.
14. Mahillon, J., and M. Chandler. 1998. Insertion sequences. *Microbiol. Mol. Biol. Rev.* **62**:725–774.
15. McBride, A. J., D. A. Athanazio, M. G. Reis, and A. I. Ko. 2005. Leptospirosis. *Curr. Opin. Infect. Dis.* **18**:376–386.
16. Nascimento, A. L., A. I. Ko, E. A. Martins, C. B. Monteiro-Vitorello, P. L. Ho, D. A. Haake, S. Verjovski-Almeida, R. A. Hartskeerl, M. V. Marques, M. C. Oliveira, C. F. Menck, L. C. Leite, H. Carrer, L. L. Coutinho, W. M. Degraive, O. A. Dellagostin, H. El-Dorry, E. S. Ferro, M. I. Ferro, L. R. Furlan, M. Gamberini, E. A. Giglioti, A. Goes-Neto, G. H. Goldman, M. H. Goldman, R. Harakava, S. M. Jeronimo, I. L. Junqueira-de-Azevedo, E. T. Kimura, E. E. Kuramae, E. G. Lemos, M. V. Lemos, C. L. Marino, L. R. Nunes, R. C. de Oliveira, G. G. Pereira, M. S. Reis, A. Schriefer, W. J. Siqueira, P. Sommer, S. M. Tsai, A. J. Simpson, J. A. Ferro, L. E. Camargo, J. P. Kitajima, J. C. Setubal, and M. A. Van Sluys. 2004. Comparative genomics of two *Leptospira interrogans* serovars reveals novel insights into physiology and pathogenesis. *J. Bacteriol.* **186**:2164–2172.
17. Picardeau, M., A. Brenot, and I. Saint Girons. 2001. First evidence for gene replacement in *Leptospira* spp. Inactivation of *L. biflexa flaB* results in non-motile mutants deficient in endoflagella. *Mol. Microbiol.* **40**:189–199.
18. Picardeau, M., S. Ren, and I. Saint Girons. 2001. Killing effect and antitoxic activity of the *Leptospira interrogans* toxin-antitoxin system in *Escherichia coli*. *J. Bacteriol.* **183**:6494–6497.
19. Rajanna, C., J. Wang, D. Zhang, Z. Xu, A. Ali, Y.-M. Hou, and D. K. R. Karaoilis. 2003. The *Vibrio* pathogenicity island of epidemic *Vibrio cholerae* forms precise extrachromosomal circular excision products. *J. Bacteriol.* **185**:6893–6901.
20. Ren, S., G. Fu, X. Jiang, R. Zeng, H. Xiong, G. Lu, H. Q. Jiang, Y. Miao, H. Xu, Y. Zhang, X. Guo, Y. Shen, B. Q. Qiang, X. Q., A. Danchin, I. Saint Girons, R. L. Somerville, Y. M. Weng, M. Shi, Z. Chen, J. G. Xu, and G. P. Zhao. 2003. Unique and physiological and pathogenic features of *Leptospira interrogans* revealed by whole genome sequencing. *Nature* **422**:888–893.
21. Rice, P., R. Craigie, and D. R. Davies. 1996. Retroviral integrases and their cousins. *Curr. Opin. Struct. Biol.* **6**:76–83.
22. Saint Girons, I., P. Bourhy, C. Ottone, M. Picardeau, D. Yelton, R. W. Hendrix, P. Glaser, and N. Charon. 2000. The LE1 bacteriophage replicates as a plasmid within *Leptospira biflexa*: construction of an *L. biflexa*-*Escherichia coli* shuttle vector. *J. Bacteriol.* **182**:5700–5705.
23. Trevejo, R. T., J. G. Rigau-Perez, D. A. Ashford, E. M. McClure, C. Jarquin-Gonzalez, J. J. Amador, J. O. de los Reyes, A. Gonzalez, S. R. Zaki, W. J. Shieh, R. G. McLean, R. S. Nasci, R. S. Weyant, C. A. Bolin, S. L. Bragg, B. A. Perkins, and R. A. Spiegel. 1998. Epidemic leptospirosis associated with pulmonary hemorrhage—Nicaragua, 1995. *J. Infect. Dis.* **178**:1457–1463.
24. Vallenet, D., L. Labarre, Z. Rouy, V. Barbe, S. Bocs, S. Cruveiller, A. Lajus, G. Pascal, C. Scarpelli, and C. Médigue. 2006. MaGe: a microbial genome annotation system supported by synteny results. *Nucleic Acids Res.* **34**:53–65.
25. WHO. 1999. Leptospirosis worldwide, 1999. *Wkly. Epidemiol. Rec.* **74**:237–242.
26. WHO. 2000. Leptospirosis, India. Report of the investigation of the post-cyclone outbreak in Orissa, November 1999. *Weekly Epidemiol. Rec.* **75**:217–223.
27. Zhang, Y. X., J. Li, X. K. Guo, C. Wu, B. Bi, S. X. Ren, C. F. Wu, and G. P. Zhao. 2004. Characterization of a novel toxin-antitoxin module, vapBC, encoded by *Leptospira interrogans* chromosome. *Cell Res.* **14**:208–216.